人因工程
Ergonomics
人間工学
Ergonomi
인간공학
Ergonomie
人类工效学
การยศาสตร์
ارگونومی
Ergonomia

Ergonomics for All

Celebrating PPCOE's 20 years of excellence

Editors:
Dyi-Yih Michael Lin
Hsieh-Ching Chen

Co-editors:
Chiuhsiang Joe Lin
Eric Min-yang Wang
An-Hsiang Wang

CRC Press
Taylor & Francis Group
A BALKEMA BOOK

ERGONOMICS FOR ALL

SELECTED PAPERS OF THE PAN-PACIFIC CONFERENCE ON ERGONOMICS (PPCOE 2010), KAOHSIUNG, TAIWAN, 7–10 NOVEMBER 2010

Ergonomics for All

Celebrating PPCOE's 20 years of excellence

Editors

Dyi-Yih Michael Lin
Department of Industrial Engineering and Management, I-Shou University, Kaohsiung, Taiwan

Hsieh-Ching Chen
Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung, Taiwan

Co-Editors

Chiuhsiang Joe Lin
Department of Industrial and Systems Engineering, Chun Yuan Christian University, Chungli, Taiwan

Eric Min-yang Wang
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

An-Hsiang Wang
Department of Labor Relations, National Chung Cheng University, Chiayi, Taiwan

CRC Press
Taylor & Francis Group
Boca Raton London New York Leiden

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A BALKEMA BOOK

CRC Press/Balkema is an imprint of the Taylor & Francis Group, an informa business

Typeset by Vikatan Publishing Solutions (P) Ltd., Chennai, India
Printed and bound in Great Britain by Antony Rowe (a CPI Group Company), Chippenham, Wiltshire

Published by: CRC Press/Balkema
P.O. Box 447, 2300 AK Leiden, The Netherlands
e-mail: Pub.NL@taylorandfrancis.com
www.crcpress.com – www.taylorandfrancis.co.uk – www.balkema.nl

ISBN: 978-0-415-58608-5 (Hbk)
ISBN: 978-0-203-82933-2 (eBook)

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Table of contents

Part III: Man-machine interface & cognitive ergonomics

Part VII: Ergonomics for people with special needs

Part VIII: Performance modeling & simulation

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Preface

Ergonomics for All represents a selected collection of the papers accepted for oral presentation at the 9th Pan-Pacific Conference on Ergonomics (PPCOE 2010) held in Kaohsiung city, Taiwan on November 7–10, 2010. PPCOE 2010 evolved from the series of Pan-Pacific Conference on Occupational Ergonomics that successfully took place in Kitakyushu Japan (1990), Wuhan China (1992), Seoul Korea (1994), Taipei Taiwan (1996), Kitakyushu Japan (1998), Beijing China (2001), Cairns Australia (2004), and Bangkok Thailand (2007).

PPCOE 2010 marked two important milestones. First, the event attracted from 16 countries/areas 200 submissions (147 papers and 53 posters), all of which peer-reviewed by the International Reviewing Board of the Scientific Committee with an averaged 82% acceptance rate. The numbers in terms of contribution quantity and geographical distribution are both record high. A 20th anniversary meeting, PPCOE 2010 would not have been able to achieve this level of success if there had not been the collective endeavors of the eight previous PPCOE conferences that have established globally recognized excellence in scientific and professional contribution.

Secondly, considering the pervasive influence of ergonomics on almost every walk of our lives, PPCOE 2010 commenced an omnivorous platform to diversify the horizon beyond simply occupational ergonomics. To address the prospects and challenges, we set *Ergonomics for All* as the conference theme, hoping to shape a lifestyle and working environments where ergonomics benefits everyone. Genuinely reflecting the essence towards such a mission, the book covers a fruitful spectrum of up-to-date and comprehensive issues in ergonomics, including Anthropometry, Biomechanics & Physical Ergonomics, Usability & User Experience, Man Machine Interface & Cognitive Ergonomics, Musculoskeletal Disorders & Ergonomics Intervention, Bio-signals & Ergonomic Assessment, Design & Ergonomics, Ergonomics for People with Special Needs, and Performance Modeling & Simulation.

The making of a memorable conference and quality proceedings has never been an easy task. We would like to convey our sincere gratitude to all of the authors and participants who are truly the key to the success of PPCOE 2010. We would like to honor professors Chiuhsiang Joe Lin and Eric Min-yang Wang (Conference Chairs) for their unconditional supports. Special thanks also go to the members of the International Reviewing Board, and Organizing Committee of the Ergonomics Society of Taiwan (host) for their professional devotion. Production of this monograph received financial support from the National Science Council of Taiwan under grant NSC99–2916-I-007–011-A1.

Editors
Dyi-Yih Michael Lin (*General Program Chair*)
Hsieh-Ching Chen (*Scientific Committee Chair*)

Part I: Anthropometry, biomechanics & physical ergonomics

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The category strategy of human body size with age intervals

Ming-Hsu Wang & Wen-Ko Chiou
Department of Industrial Design, Chang Gung University, Kweishan, Taoyuan, Taiwan

ABSTRACT: Sizing category strategy play an important role in the anthropometric measurement, it will affect the application of the different fields, including garment, ergonomic product, interior design, working space and nutritional status of people.Through the three-dimensional (3D) body scan process, this research collect 8463 digital human models from the Chang Gung Memorial Hospital. It categories the human body size by two body indexes and four age intervals. The result shows that Body Mass Index (BMI) could be the better sizing strategy to category the human body size. And the "age-10age" interval has the minimum Mean Absolute Difference (MAD) with the 22 body dimensions which compare with the 56 standard human models than the three age interval in both genders. The index "BMI" and "age-10" interval could be the better strategy to category the human body size. The standard human models could be the better template to analyze the digital human model; it will provide a quick and effective method to collect the body dimensions to the different applications.

Keywords: three-dimensional (3D), body dimensions, body index, age interval, mean absolute difference

1 INTRODUCTION

Sizing category strategies provide size specification for different population groups based on somebody dimension data from demographic anthropometric surveys and studies. The goal of sizing is to choose limited size groups to cover large percentage of the population. They could be applied to the garment industry, work place design, helmet design, and other operator related products. A sizing strategy is created by using various methods from trial-and-error to complicated statistical methods. Sizing approaches have been improved and optimization of sizing strategys has been developed all along (McCulloch, 1998).

Sizing category strategy play an important role in the anthropometric measurement, it will affect the application of the different fields, including garment, ergonomic product, interior design, working space and nutritional status of people. Characterizing and understanding sizing category is traditionally the subject of anthropometry-the study of human body measurement. It is essential for better ergonomic design of any product with which people interact, such as clothing, automobiles, and workstations.

Consumer products that relate to the human body rely on good fit for proper function. The design of such products relies on the availability of accurate anthropometric information describing the body characteristics.

Human anthropometry involves the measurement of body characteristics such as reach, body segment length and circumferences, widths, and heights, among others. This information can be used to inform the design of tools, equipment, workstations and clothes. 3D body-surface scanners are transforming the ability to accurately measure a person's body size, shape, and skin-surface area. Three-dimensional data can be a great aid to overcome the disadvantages of traditional category strategy since they can provide rich geometry information on the 3D surface. And it is useful for the application of the anthropometry and nutrition. The 3D human body data can have wide applications in industry, such as design of shoes and spectacle frames (Mochimaru, 2000), apparel design (Perissinotto, 2002, Santos, 2004, Wang et al, 2003).

3Danthropometric measurement strategies offer an interesting alternative to traditional methods in applications such as clothing sizing and product design. The automated strategies are attractive because of their low cost and the speed with which they can measure size and determine the best-fitting garment and products. Three-dimensional data can be a great aid to overcome the disadvantages of traditional category strategy since they can provide rich geometry information on the 3D surface (Li, 2006).

Anthropometric measurements, although limited, are the most practical way for a nutritional assessment of individuals and populations, through nutritional risk monitoring, prognosis of acute and chronic diseases, or clinical actions (Barbosa, 2005, Perissinotto, 2002, Santos, 2004).

The most widely used measurements are weight and height, which are often combined as Body Mass Index (BMI; in kg/m^2) to provide a proxy for nutritional status. BMI is used to categorize underweight (James, 1988), normal weight, overweight, and obesity (Garrow, 1985), and much research has illustrated disease (CVD) and other diseases (Hayashi, 2005, Zhou, 2002).

Body shape contrasts with size in providing information on weight distribution. BMI represents a very crude index of shape, whereas Waist Circumference (WC) gives a clearer indication of relative abdominal shape. More sophisticated information can be obtained from ratios of different body girths, such as the Waist-Hip (WHR) or Waist-Chest (WCR) Ratio, to act as a proxy for central adiposity. Such measurements are easy to perform and are often highly informative.

How useful these indicators are depends on the availability of reference data related to age, sex, and age group, specific to each population. The World Health Organization (de Onis, 1996) recommends the development of specific reference values for each country, with availability of means, standard deviations, and percentiles for every measurement or index. BMI is a useful alternative to traditional somatotyping techniques; the findings also suggest that a prisoner's somatotype is associated with criminal patterns while being a meager predictor of criminality. Methodological and theoretical implications of this study are discussed (Maddan, 2008). Body fat distribution, as measured by Waist-to-Hip Ratio (WHR), may be an ideal feature. First, WHR, unlike overall body weight, is an unambiguous indicator of age in men and women (Maddan, 2008). Prior to puberty, the sexes have similar WHR because body fat distribution is essentially similar.

However how to category the body size by the different criteria could be an important issue to the application, and could be a critical factor of the presentation of the human body size. The main aim of this research tries is to find out the strategy of category for the human body size by the different age interval and body index.

2 METHODS

Through the 3D body lasers scan process; this research collected 8463 digital human body models from the Chang Gung Memorial Hospital, Tao-Yuan, Taiwan. To capture the intricacies of the human body in one pass, the Chang Gung Whole-Body Scanner (CGWBS, as shown in Figs. 1, 2 and Fig. 3) uses four scanning instruments mounted on three vertical towers. Each tower has a linear ball-bearing rail and servo motor assembly that moves the scanning instrument vertically. When a subject stands on the scanner's platform, the scanning instrument starts at the person's head and move down to scan the entire body.

The actual measurement procedure is as follows: the examinee stands in the measuring positions, both hands down and forming a 30 degree angle with the body. The operator will execute the procedure, scanning the examinee from top to bottom, for about 12 seconds, at this time the body contour's resolution is about 2.5 mm. Next, the examinee holds on to the handle (the handle has been designed so that it does not interfere with the measuring), in order for the body to stand erect. Then, the body is scanned from top to bottom, focusing on the examinee's head, thoracic, hip, and leg shapes. At this time, the resolution is 1 mm, and the scanning time is 24 seconds. After this measurement, before the examinee leaves, the operator must review the measurement results on the computer, and if there is any error, the subject could be immediately measured again.

After the body scanning process, this research tries to build up the digital human body size database for application (as shown in Fig. 2). And it collects and analyzes the 22 human body dimensions, including circumference, width, area and volume to build up the digital human

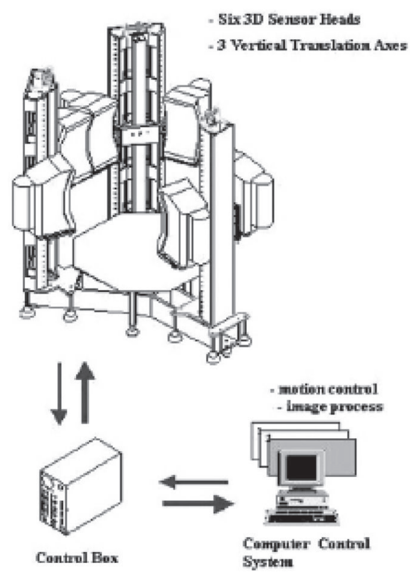


Figure 1. Chang Gung Whole-Body Scanner-a scanning illustration.

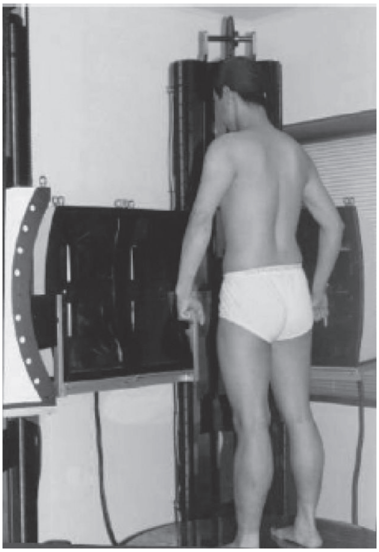


Figure 2. Chang Gung Whole-Body Scanner-a scanning illustration.

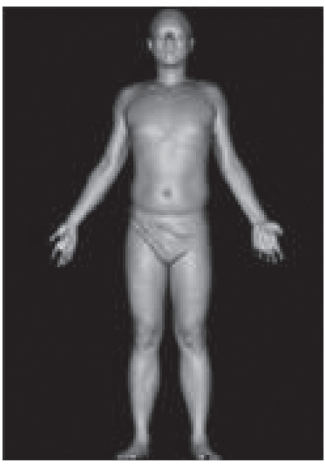


Figure 3. Chang Gung Digital Human Body-a scanning illustration.

database, the dimensions including distance, girth, circumference, volume, surface area, Body Mass Index(BMI), Waist-Hip Ratio (WHR), Waist-Hip area Ratio (WHaR), weight and height.

According the reference review result, it categories the human body size by the BMI and WHR indexes. First, we test the correlation between 22 body dimensions and two body indies. It also use the cluster analysis method to category human body size into 28 standard human models for both genders by the different age interval from age 21 to age 90, including age all, age 5-years, age 10-years and age-generation. Using correlation analysis to find out the best body index strategy and analyze the term Mean Absolute Differences (MAD) to indicate the error between the measurements of the different standard sample subjects from the 22 human body dimensions, the MAD mean that the difference between the standard samples. And divide into five categories, including circumference (C), Width (W), Profile Area (P-A), Volume (V) and Surface Area (S-A) with the standard human models from the cluster analysis, the MAD calculation.

Then we compare the total difference from the 28 standard human models by the different strategies. To compare the mean difference and find out the better strategy to category the human body size with the body index and age interval.

3 RESULTS & DISCUSSION

The result shows that BMI could be the better body index strategy to category the human body size than the WHR index. BMI has the higher correlation with the 22 body dimensions than WHR in both genders, and both genders also have the similar trend for the two body indexes. Most correlation index of weight and BMI is higher than 60%, it shows that they have high correlation with most body dimensions. And it shows that BMI could be the better strategy to category the size than WHR index (as shown in Figs. 4, 5).

To compare the indexes, weight and BMI, BMI, often referred to as the Quetelet Index, utilizes a person's height and weight to gauge the total body fat in adults; it is an indicator of optimal weight for health and different from lean mass or percent body fat calculations because it only considers height and weight. The BMI is a well established measure in the medical community; this is observed by a cursory review of the available literature (Feitosa, 2002). We have given interpretations to the first five components-some of them provide quantitative evidence to the empirical anthropometric observations (Sheldon, 1940).

Compare the MAD of the five categories of dimensions, the result shows that (as shown in Figs. 6, 7) the age interval "age-10" has the minimum Mean Absolute Difference (MAD) with the 22 body dimensions which compare with the 56 standard human models than the three age interval in both genders. In the result of males, the age interval "age-10" have the minimum MAD and the "age-5" almost the same with the "age-10", and the "age-all" got the maximum MAD between the four different age intervals.

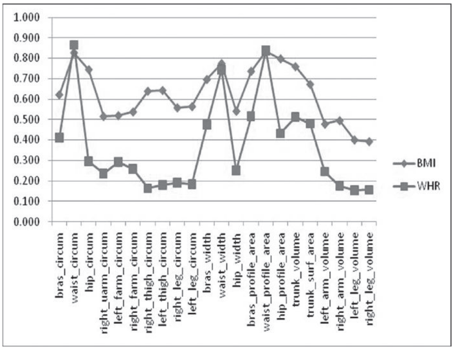


Figure 4. Correlation with body index and dimensions (Male).

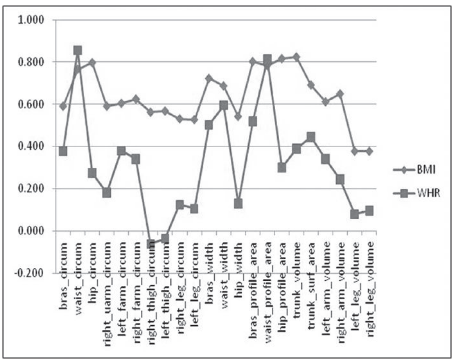


Figure 5. Correlation with body index and dimensions (Female).

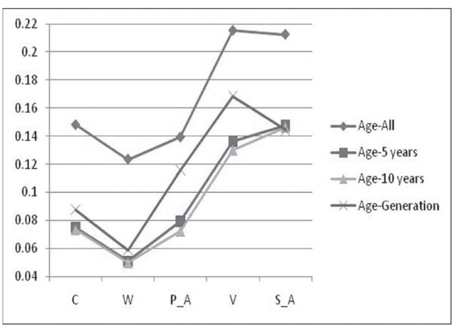


Figure 6. MAD with age interval and body dimensions (Male).

In the result of females, the age interval "age-10" also have the minimum MAD and but the "age-5" did not have the similar MAD with the "age-10", and the "age-generation" got the maximum MAD between the four different age intervals. And the "age-all" is similar to the "age-10" interval.

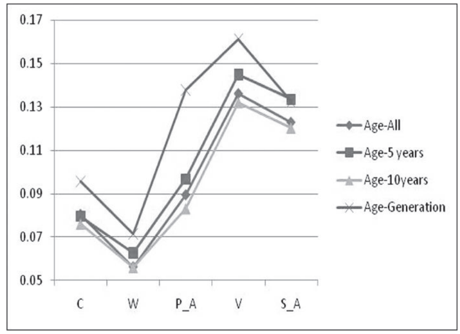


Figure 7. MAD with age interval and body dimensions (Female).

The results of both genders show that the width dimension got the minimum MAD in the all dimension categories. And for male, the volume dimension got the highest MAD than other dimensions. It is different between both genders to compare the category dimensions.

4 CONCLUSIONS

Compare the WHR and BMI indexes; BMI got the higher correlation with the 22 body dimensions. BMI index could be the better strategy for body size category. Then we try to find out the better age interval strategy with the BMI index for category strategy.

BMI is insensitive to age-associated changes in the distribution of body dimensions, which also differ markedly between the sexes. In both sexes, BMI did not increase steadily with age interval group.

Bout the age interval effect, the "age-10" could be the better way to category the body size with the lowest MAD in the all age intervals. So this research try to conclude that BMI and age-10 could be the better strategy to category the human body size, and collect the standard human models from the large digital human database. The standard human models could be the best template to analyze the digital human model; it will provide a quick and effective method to collect the body dimensions to the different applications.

REFERENCES

Barbosa, A.R., Souza, J.M.P., Lebrão, M.L., Laurenti, R. & Marucci, M.F.N. 2005. Anthropometry of elderly residents in the city of São Paulo, Brazil. *Cad Saude Publication* 21: 1929–1938.

De Onis, M. & Habicht, J.P. 1996. Anthropometric reference data for international use: Recommendations from a world health organization expert committee. *American Journal of Clinical Nutrition* 64: 650–658.

Feitosa, M.F., Borecki, I.B., Rich, S.S., Arnett, D.K., Sholinsky, P., Myers, R.H., Leppert, M. & Province, M.A. 2002. Quantitative-trait loci influencing body-mass index reside on chromosomes 7 and 13: The National Heart, Lung, and Blood Institute Family Heart Study. *American Journal Human Genettation* 70(1): 72–82.

Garrow, J.S. & Webster, J., 1985. Quetelet's index (W/H2) as a measure of fatness. *International Journal of Obesity* 9: 147–53.

Hayashi, R., Iwasaki, M., Otani, T. 2005. Body mass index and mortality in a middle-aged Japanese cohort. *Journal of Epidemiol* 15(3): 70–77.

James, W.P., Ferro-Luzzi, A. & Waterlow, J.C. 1988. Definition of chronic energy deficiency in adults: Report of a working party of the International Dietary Energy Consultative Group. *European Journal of Clinical Nutrition 42:* 969–981.

Kouchi, M. & Mochimaru, M. 2002. Analysis of 3D Human Face Forms and Spectacle Frames Based on Average Forms. In: *Proceedings of the Digital Human Modeling Conference*, Munich, 69–89.

Li, Z.Z. 2006. Anthropometric topography. In: Karwowski, Waldemar (Ed.), *International Encyclopedia of Ergonomics and Human Factors, second ed*. Taylor & Francis, London.

Maddan, S., Walker, J.T. & Miller, J.M. 2008. Does size really matter? A reexamination of Sheldon's somatotypes and criminal behavior. *The Social Science Journal* 45: 330–344.

McCulloch, C.E., Paal, B. & Ashdown, S.A. 1998. An optimization approach to apparel sizing. *Journal of the Operational Research Society* 49(5): 492–499.

Mochimaru, M., Kouchi, M. & Dohi, M. 2000. Analysis of 3-D human foot forms using the free form deformation method and its application in grading shoe lasts. *Ergonomics* 43(9): 1301–1313.

Paquette, S. 1996. 3D scanning in apparel design and human engineering. *IEEE Computer Graphics and Applications* 16 (5): 11–15.

Perissinotto, E., Pisent, C., Sergi, G., Grigoletto, F. & Enzi, G. 2002. Anthropometric measurements in the elderly: age and gender differences. *British Journal of Nutrition* 87: 177–86.

Santos, J.L., Albala, C., Lera, L., Garcia, C., Arroyo, P. & Perez-Bravo, F. 2004. Anthropometric measurements in the elderly population of Santiago, Chile. *Nutrition* 20: 452–457.

Sheldon, W., Stevens, S. & Tucker, W. 1940. *The varieties of human physique*. Harper and Brothers Publishers: New York.

Wang, M.J., Lin, Y.C. & Wu, W.Y. 2003. The development of automated tailoring strategy using 3D scanning data. *IEA 2003*, 24–29.

Zhou, B.F. 2002. Effect of body mass index on all-cause mortality and incidence of cardiovascular diseases—report for meta-analysis of prospective studies open optimal cut-off points of body mass index in Chinese adults. *Biomedical Environment Science* 15: 245–52.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluating gender differences in foot dimensions

Yu-Chi Lee, Gloria Lin & Mao-Jiun J. Wang
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

ABSTRACT: The objective of this study was to collect foot anthropometric data of Taiwanese adults and to examine the gender differences in foot dimensions. Three hundred healthy male and female students were recruited. A 3D foot scanner was used to collect 14 foot dimensions. Results showed that the average difference between males and females was about 24 mm and 10 mm in foot length and breadth, respectively. The right foot was smaller than left foot in foot length, ball of foot length, hell breadth and sphyrion fibulare height regardless of genders. Overall, body weight and stature had significant correlates to the 14 foot dimensions except for navicular height. Designing footwear should consider the difference in foot dimensions among different genders.

Keywords: gender difference, foot dimensions, anthropometry

1 INTRODUCTION

Footwear is used for protecting feet when users engage in different kind of activities. Wearing improper footwear increases the risk of foot problems, for instance, corns, ankle injury, chronic pain or blister of foot (Killian et al. 1998). Using anthropometric data for footwear design can enhance the fitness between footwear and feet.

Many studies have been investigating the anthropometry of foot dimensions and its relationship among factors such as gender, stature, and body sizes. Gender effect on foot dimension has been studied for a long time. In Africa, males had greater foot dimensions than females; the dimension difference between genders was about 30 mm and 10 mm in foot length and breadth, respectively (Agnihotri et al. 2007). To minimize the difference between subjects, some studies selected the same foot length (Anil et al. 1997, Voracek et al. 2007, Krauss et al. 2008) or the same stature (Wunderlich & Cavanagh 2001) to evaluate gender effect. The findings indicated that males' foot have greater size in breadth and girth than those of females'. Agrawal et al. (2010) collected the foot anthropometrics data in workers of Meghalaya and indicated that the gender difference was 13 mm in foot length and 6 mm in ball breadth. Moreover, Xiong et al. (2009) gathered 9 foot dimensions of 30 subjects in Hong Kong. They found that males had 22.2 mm greater in foot length and 7.3 mm greater in foot width than the females. From the results of previous study, we can conclude that males have greater foot length and breadth than females. Nevertheless, Salami (2009) indicated that foot breadth of the Nigerian females was larger than that of the males (about 4 mm), but males had greater foot length than females for about 14 mm.

In 1997, Anil et al. observed that stature had high correlation with foot length. In addition, Manna et al. (2001) indicated that the stature had higher correlation on foot length than weight, and foot length was highly correlated with stature and foot volume, especially in left foot. Furthermore, Ozden et al. (2005) evaluated the gender difference in correlation between foot dimensions. For males, a significant correlation was observed in stature and right foot length and also in stature and right foot breadth. Moreover, it was shown that stature was related to right foot length of females, but there was no significant correlation between stature and right foot breadth.

Since there is little information available on foot dimensions between genders in Taiwan, this research was aimed to investigate the gender effect on 14 foot dimensions of Taiwanese adults and to evaluate the correlations of the dimensions associated with stature and weight. The results of this study can be very useful for footwear design.

2 METHOD

2.1 *Participant*

Three hundred subjects including 150 males and 150 females, aged from 18 to 37 years (mean = 21.8) were recruited for this study. They were undergraduate and graduate students in a university and without any history of feet abnormalities. The basic demographic data and feet anthropometric data of the participants are shown in Table 1.

2.2 *Experimental apparatus*

The feet dimensions were collected by using a 3D foot scanner (INFOOT USB scanning system, IFU-S-01, I-Ware Laboratory Co., Ltd., Japan).

Table 1. The basic demographic data of the participants.

Terms	Males (N = 150)	Females (N = 150)
Age (year)	21.8 (2.7)*	21.8 (2.8)
Stature (cm)	174.0 (5.4)	161.1 (5.1)
Weight (kg)	67.7 (9.1)	51.9 (6.0)
BMI	23.4 (3.0)	20.0 (2.2)

* Mean (Standard Deviation).

The INFOOT scanning system has 8 CCD cameras and 4 laser projectors to construct foot shapes and collect dimensions. The accuracy of the foot scanner was within 1.0 mm (Kouchi & Mochimaru 2001). Fourteen foot dimensions of Right Foot (RF) and Left Foot (LF) were collected in this study. The definitions of the dimensions are presented in Table 2.

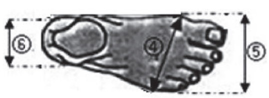
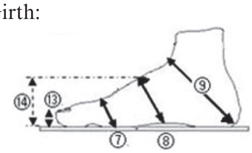
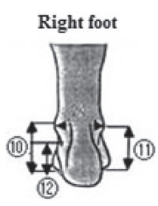
2.3 *Experimental design*

The independent variable in this study was gender. Fourteen foot dimensions were selected as dependent variables. A single-factor experiment design was used to evaluate the gender effect on 14 foot dimensions.

2.4 *Experimental procedure*

At the beginning of the experiment, each participant was requested to wash his/her feet and use tissue paper to dry their feet surface completely

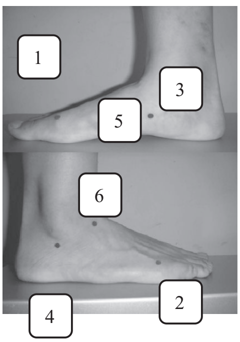
Table 2. The 14 foot dimensions.

	Dimensions	Definition
Length: 	Foot length	The distance along the X-direction from the end of heel to the tip of longest toe.
	Ball of foot length	The distance from the end of heel to the 1st Metatarsal point protrusion.
	Outside ball of foot length	The distance from the end of heel to the 5th Metatarsal point protrusion.
Breadth: 	Ball of foot breadth diagonal	The straight distance from the most medially placed point on the head of 1st metatarsal to the most laterally placed point located on the head of 5th metatarsal.
	Ball of foot breadth horizontal	The horizontal distance between 1st metatarsal to 5th metatarsal.
	Heel breadth	The widest distance of the heel.
Girth: 	Ball girth	Maximum girth over the first and fifth metatarsal joint.
	Instep girth	Maximum girth over middle cuneiform prominence (Clark, 1976).
	Short heel girth	Maximum girth around back heel point and tentative junction point of foot and the leg.
Height: 	Sphyrion height	The vertical distance (Z-direction) from the land to the most prominent point on the sphyrion.
	Sphyrion fibulare height	The vertical distance from the land to the most prominent point on the fibulare sphyrion.
	Navicular height	Measure vertical components of the vertical distance of a marker placed below the anatomic point of navicular.
	Instep height	The vertical distance form top height point of mid foot to land.
	Toe height	The vertical distance form top height point of toe to land.

before collecting data. This step was to avoid measuring errors due to the particles on feet surface. Applying the landmarks was helpful in obtaining the accurate measurement. The experimenter has been well trained to indentify the six anatomical landmarks, and to put markers on each participant's foot. The positions of the landmarks are shows in Fig. 1. After land marking, the participants were asked to keep their feet still to complete the scan. We scanned one foot at a time and then changed to the other. Each participant was scanned twice to ensure the image quality for subsequent analysis.

2.5 *Statistical analysis*

The t-test was conducted to evaluate the gender effect on 14 foot dimensions. Pair t-test was performed to evaluate the difference between feet. The correlation coefficients among the selected measurements were calculated. All the data were analyzed with SPSS 13.0 software package. The significance level was set at $\alpha = 0.05$.



1. Metatarsle tibiale
2. Metatarsle fibulare
3. Sphyrion
4. Sphyrion fibulare
5. Navicular
6. Tentative junction point of foot and leg

Figure 1. The position of six landmarks in present study.

3 RESULTS AND DISCUSSION

3.1 *Gender effect*

Table 3 shows that the means and standard deviations of the 14 dimensions and the gender difference on RF and LF. The result of independent t-tests in Table 3 indicates that there were significant differences in the foot dimensions between genders. Overall, males had greater foot dimensions than females and the difference between genders was about 24 mm and 10 mm in foot length and breadth, respectively. Agnihoyti et al. (2007) indicated that the gender difference in African was 30 mm in foot length and 10 mm in foot breadth.

3.2 *The difference between the left and right foot*

In addition, the differences between the left and right foot dimensions were evaluated and the results are shown in Table 4. For males, there are significant left and right foot differences in 14 foot dimensions except for foot breadth diagonal, ball girth and sphyrion height. For females, there are significant differ ences in all dimensions between left and right foot except for foot length, foot breadth

Table 3. The means and standard deviations of the 14 foot dimensions among genders.

Dimensions#	Males RF	Males LF	Females RF	Females LF	Gender effect
Foot length	259.2 (10.9)	260.0 (10.9)	235.5 (10.2)	235.6 (10.0)	RF*** LF***
Ball of foot length	188.5 (9.0)	190.6 (8.5)	171.9 (7.8)	173.4 (7.8)	RF*** LF***
Outside ball of foot length	170.9 (8.4)	169.9 (8.0)	154.4 (6.9)	153.5 (7.2)	RF*** LF***
Foot breadth diagonal	104.3 (4.9)	104.4 (4.9)	93.7 (4.9)	93.5 (4.6)	RF*** LF***
Foot breadth horizontal	102.4 (4.9)	101.7 (4.8)	91.5 (4.6)	90.8 (4.6)	RF*** LF***
Heel breadth	66.1 (3.7)	66.4 (3.8)	59.5 (3.1)	59.8 (3.2)	RF*** LF***
Ball girth	227.8 (11.2)	228.2 (11.8)	200.6 (10.4)	200.6 (9.6)	RF*** LF***
Instep girth	243.6 (11.7)	242.3 (11.5)	215.6 (10.1)	214.9 (9.9)	RF*** LF***
Short heel girth	328.9 (13.3)	327.7 (14.7)	296.1 (11.4)	294.9 (12.2)	RF*** LF***
Sphyrion height	67.6 (5.4)	67.3 (5.6)	61.3 (5.4)	60.5 (5.8)	RF*** LF***
Sphyrion fibulare height	52.4 (4.3)	54.2 (4.3)	47.5 (4.0)	48.4 (4.0)	RF*** LF***
Navicular height	37.0 (5.1)	35.7 (5.5)	33.1 (4.9)	31.8 (4.6)	RF*** LF***
Instep height	40.3 (2.9)	39.6 (4.8)	36.2 (2.4)	35.7 (2.5)	RF*** LF***
Top of ball girth height	68.0 (4.4)	66.9 (4.3)	59.1 (3.9)	58.8 (3.8)	RF*** LF***

#: All dimensions are in mm; RF: Right Foot; LF: Left Foot; * p < 0.05; ** p < 0.01; *** p < 0.001.

diagonal and instep girth. Moreover, Table 4 also shows that the right foot was smaller than left foot in foot length, ball of foot length, heel breadth and sphyrion fibulare height regardless of genders. It is interesting to note that Manna et al. (2001) indicated that significant difference between feet was observed in male's foot breath and female's foot length. Beside, Agnihotri et al. (2007) and Ozden et al. (2005) also found that there was no significant difference in foot length and breadth between left and right foot. Our result was not consistent with the previous studies. One of the reasons was due to the different ethnic populations and the habits developed (Anil et al. 1997).

3.3 *Correlations between the selected measurements*

The results of the correlations between stature, weight and 14 foot dimensions are presented in Table 5. The correlation was found to be significant at alpha level set 0.05. Stature and body weight had significant correlations with all dimensions except for navicular height. It seems that stature was closely related to length dimensions, including foot length, inside/outside ball of foot length, and body weight was correlated to breadth and girth measurements (correlation coefficient >0.5).

Table 4. Results of the paired T-tests between right foot and left foot.

Dimensions[#]	Male (RF-LF)	Female (RF-LF)
Foot length	–0.83**	–0.06
Ball of foot length	–2.09***	–1.45***
Outside ball of foot length	0.91**	0.73*
Foot breadth diagonal	–0.04	0.16
Foot breadth horizontal	0.64***	0.69***
Heel breadth	–0.32**	–0.31**
Ball girth	–0.04	0.84**
Instep girth	1.33**	0.78
Short heel girth	1.22**	1.12**
Sphyrion height	0.29	0.73*
Sphyrion fibulare height	–1.80***	–0.92***
Navicular height	1.30***	1.29***
Instep height	1.07***	0.57***
Top of ball girth height	0.70*	0.32*

[#]: All dimensions are in mm; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 5. Correlation between stature, weight and 14 foot dimensions.

	Stature		Weight	
Dimensions	Male	Female	Male	Female
Foot length	0.683***	0.734***	0.413***	0.352***
Ball of foot length	0.647***	0.673***	0.402***	0.319***
Outside ball of foot length	0.629***	0.691***	0.486***	0.384***
Foot breadth diagonal	0.247**	0.298***	0.510***	0.458***
Foot breadth horizontal	0.231**	0.298***	0.530***	0.499***
Heel breadth	0.319***	0.245**	0.564***	0.371***
Ball girth	0.137 NS	0.261**	0.677***	0.534***
Instep girth	0.283***	0.237**	0.637***	0.509***
Short heel girth	0.538***	0.539***	0.706***	0.595***
Sphyrion height	0.226**	0.343***	0.263**	0.288***
Sphyrion fibulare height	0.323***	0.349***	0.208*	0.187*
Navicular height	0.252**	0.095 NS	0.109 NS	0.101 NS
Instep height	0.030 NS	0.236*	0.418***	0.438***
Top of ball girth height	0.211*	0.251*	0.464***	0.387***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; NS: Non-Significant.

4 CONCLUSION

This study collected 14 foot anthropometric data by using 3D foot scanner. The average difference between males and females was about 24 mm and 10 mm in foot length and breadth, respectively. For the difference between feet, the right foot was smaller than left foot in foot length, ball of foot length, heel breadth and sphyrion fibulare height regardless of genders. Moreover, all the foot dimensions were significantly correlated to body weight and stature. The findings can be useful for designing footwear of different genders.

REFERENCES

Agnihotri, M.D., Shukla, S. & Purwar, B. 2007. Determination of sex from the foot measurements. *The Internet Journal of Forensic Science* 2(1).

Agrawal, K.N., Singh, R.K.P. & Satapathy, K.K. 2010. Anthropometric considerations for farm tools/machinery design for tribal workers of north eastern India. *Agricultural Engineering International* 7: 1406–1466.

Anil, A., Peker, T., Turgut, H.B. & Ulukent, S.C. 1997. An examination of the relationship between foot length, foot breath, ball girth, height and weight of Turkish university students aged between 17 and 25. *Anthropol Anz* 55: 79–87.

Clark, Ltd. Training Dept. 1976. *Manual of shoemaking*, 2nd edition, Training Department Clarks.

Killian, R.B., Nishimoto, G.S. & Page, J.C. 1998. Foot and ankle injuries related to rock climbing: The role of footwear. *JAM Podiart Med Assoc* 88: 365–374.

Kouchi, M. & Mochimaru, M. 2001. Development of a low cost foot-scanner for a custom shoe making system. In *5th ISB Footwear Biomechanics*: 58–59.

Krauss, I., Grau, S., Mauch, M., Maiwald, C. & Horstmann, T. 2008. Sex-related differences in foot shap. *Ergonomics* 51(11): 1693–1709.

Manna, I., Pradhan, D., Ghosh, S., Kar, S.K. & Dhara, P. 2001. A comparative study of foot dimension between adult male and female and evaluation of foot hazards due to using of footwear. *Journal of physiological anthropology and applied human science* 20(4): 241–246.

Ozden, H., Balci, Y., Demirüstü, C., Turgut, A. & Ertugrul, M. 2005. Stature and sex estimate using foot and shoe dimensions. *Forensic Science International* 147(2): 181–184.

Salami, O. 2009. Anthropometric data of hand, foot and ear or university students in Nigeria. *Leonardo Journal of Sciences* 15: 15–20.

Voracek, M., Fisher, M.L., Rupp, B., Lucas, D. & Fessler, D.M. 2007. Sex differences in relative foot length and perceived attractiveness of female feet: relationships among anthropometry, physique, and preference ratings. *Percept Motor Skills* 104(3): 1123–38.

Wunderlich, R.E. & Cavanagh, P.R. 2001. Gender differences in adult foot shape: implication for shoe design. *Medical Science Sports Exercise* 33(4): 605–611.

Xiong, S., Goonetilleke, R.S., Zhao, J., Li, W. & Witana, C.P. 2009. Foot deformations under different olad-bearing conditions and their relationships to stature and body weight. *Anthropological Science* 117(2): 77–88.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Determination of the hip joint center based on 3D surface anthropometry

Wei-Ying Chou & Wen-Ko Chiou
Chang Gung University, Tao-Yuan, Taiwan

ABSTRACT: The predictive methods are commonly used to estimate the Hip Joint Center (HJC) clinically. This study aimed to determine whether the predictive methods are suitable when applies to 3D scan data. We also sought to determine the errors between different predictive methods. 14 subjects were taken the whole body three-dimensional laser scan. The inter-anterior superior iliac spine distances (IAD) were acquired by 3D surface anthropometry, and two predictive methods that only relied on IAD were used to calculate HJC location. The HJC location was also determined using motion analysis. The distances between right and left HJCs were calculated, and the differences between those methods were studied. Results show the predictive methods were accurate estimated the HJC on the 3D scan data, and the most accurately method was by taking the point 14% IAD medial, 30% IAD distal and 19% IAD posterior to the Antero-Superior Iliac Spine (ASIS).

Keywords: Whole body scanner, 3D surface anthropometry, Hip joint center, ASIS

1 INTRODUCTION

Present paper describes a technique to locate the Hip Joint Center (HJC) on the 3D human scanned data, for deform these data more accurately.

Skeleton-driven deformation is the most universal method to drive the 3D scan data (Allen, Curless, & Popović, 2003). Generally, an articulated linkage representation of the human skeletal system remains to be the most widely adopted framework for defining body segment. Pervious paper described the procedure to define a skeleton for controlling the pose of the 3D scan data: by creating an articulated skeleton, then bound to the mesh model by the user, typically by manual indication of a correspondence between elements of each structure, or by computing the "center of the model" (Allen, Curless, & Popović, 2002). However, Mullaji et al. (2010) investigated the Computer Tomography (CT) images of 200 adults and found the Body Mass Index (BMI) had no influence on the HJC location. Extracting skeleton only form the shape of the mesh has been inevitably be affected by the soft tissue, the errors will give huge effect on the low extremity kinematics and kinetics.

Allen et al. (2003) described a technique rely on the anatomy base: creating a skeleton by manually choosing 2–3 points on the surface to act as markers for each joint in the skeleton, and then calculate the local position of these markers in the joint's coordinate frame. However, for this to be an accurate method, the equation of joint calculation should be clearly defined.

Previous study descried the optimal linkage representation should be take the central axes of the linkage need to be most closely approximate the real joint center (Zhang, Lee, & Braido, 2004). The accurate identification of joint centers is of great importance for the correction of human body kinematics and kinetics. Many studies reported the inaccuracies in joint center will result in joint translations (Piazza, Okita, & Cavanagh, 2001; Stagni, Leardini, Cappozzo, Grazia Benedetti, & Cappello, 2000). Locate the joint center more accuracy when human skeletal system modeling is the key to deform the 3D scan data more realistically.

The hip, knee and ankle joint are used to define anatomical frame of the lower extremity (Cappozzo, Catani, Della Croce, & Leardini, 1995). While the joint centers of knee and ankle are easier to locate, the deeply located hip joint is not easily identified. The accurate identification of HJC is of great importance for the correction of human body kinematics and kinetics (Harrington, Zavatsky, Lawson, Yuan, & Theologis, 2007; Hicks & Richards, 2005; Kirkwood, Culham, & Costigan, 1999; Piazza, et al., 2001; Stagni, et al., 2000).

Previous studies demonstrate many predictive methods to estimated the HJC location, such as predicted by percentages of bony landmarks (Bell, Pedersen, & Brand, 1990; Harrington, et al., 2007;

Kirkwood, et al., 1999; Seidel, Marchinda, Dijkers, & Soutas-Little, 1995), functional approach and sphere fit methods (Begon, Monnet, & Lacouture, 2007; Camomilla, Cereatti, Vannozzi, & Cappozzo, 2006; Ehrig, Taylor, Duda, & Heller, 2006; Hicks & Richards, 2005; Piazza, et al., 2001). Functional approach and sphere fit methods have limitation in their application to 3D scan database since the necessity of subjects to perform specific range of motion trials, nerveless work suggesting these methods may improve the accuracy in prediction. Seidel et al. stated that HJC was located relative to the ASIS: 14% of IADs medial, 34% of pelvic depth (distance from ASIS to posterior superior iliac spine) posterior, 79% of pelvic height (perpendicular form pubic center to the intra-ASISs line) inferior. Although the method is proved on accuracy, these are inappropriate to 3D whole body scan dataset since they rely on the location of pubic center, a place have been shelter by cloth and inappropriate to palpate or to mark in the process of scanning. Harrington et al. taken the MRI data to regression analysis and proposed an equation relies on IAD and pelvic depth (the distance between the midpoints of the line segments connecting the two ASIS and the two posterior superior iliac spine). However, the posterior superior iliac spine is difficult to observer from the 3D scan data. Bell et al. (1990) and OrthoTrak software estimated the location of the HJC in all three planes using a fixed percentage of the IAD. These methods located HJC only rely on the ASIS coordinates are general applied in clinical test, also are superior methods for applied on the 3D scan data. However, none of previous research has applied those methods on the 3D scan data. On the other hand, the 3D scan data have high degree of fidelity on the bony landmarks, making the applying of predictive methods to the scan data more feasible.

This study aimed to determine whether the predictive methods are suitable when applies to 3D scan data. We also sought to determine the errors between different predictive methods. Our hypothesis was that the predictive methods might show good accuracy when apply on the 3D scan data.

2 METHODS

14 subjects (8 male, 6 female), with a mean (SD) age of 21.4 (0.79) years old, body weight 61.6 (10.0) kg, body height of 1.68 (0.08) m were recruited in this study. All subjects with no history of lower limb problems gave informed consent and participated in this study, which was approved by the local ethics committee (no: 97-2538B). The whole body laser scan was taken of the subjects, all the procedures were in accordance with previous study (Lin, Chiou, Weng, Fang, & Liu, 2004). The subjects were asked to stand with shoulder width, maintain the posture before the scan have finish. Relevant anatomical landmarks of right and left Antero-Superior Iliac Spine (ASIS) were marked before scanned.

The Anthro3D software (Logistic Technology, Taiwan) were used to locate the coordinates of the ASIS (x, y, z) and measure the inter-ASIS distance (IAD) within the scanned dataset (Fig. 1). The x axis was defined as the anterior-posterior direction, y axis was defined as the medial-lateral direction, and z axis was defined as the superior-inferior direction.

The coordinates of HJC (x, y, z) were been gathered from two predictive methods that calculated using the surface anthropometry data. Those equations predicted HJC in every plane relied on IAD. When the coordinate system origin is located halfway between the right and left ASISs, one equation estimated HJC at 22% posterior, 30% inferior, and 36% lateral to the origin point (Bell, et al., 1990), and the other equation (software recommendations for OrthoTrak, Motion Analysis Corp., CA, USA) estimated HJC at 22% posterior, 34% inferior, and 32% lateral to the origin point (Table 1).

The HJC location was also determined using motion capture system, which is considered the gold standard. The subjects were barefoot, asked to stand with shoulder width for 5 seconds. 17 reflective markers were place in accordance with

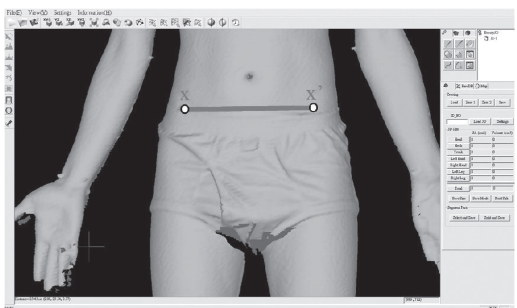


Figure 1. 3D surface anthropometry to measure the inter-ASIS distance (IAD, *line XX'*) on the 3D scanned example.

Table 1. Prediction methods from the literature for the right HJC coordinates in the pelvis.

	Posterior (x)	Lateral (y)	Inferior (z)
Bell, 1990	–0.19IAD	0.36IAD	–0.30IAD
OrthoTrak	–0.22IAD	0.32IAD	–0.34IAD

IAD = inter-ASIS distance (mm).

the Newington Children's hospital (Davis Iii, Õunpuu, Tyburski, & Gage, 1991). Five cameras were used to perform the video capture at a sampling frequency of 60 Hz. These data was synchronized to obtain the three-dimensional coordinates using the APAS motion analysis software (Ariel Dynamics, USA).

The locations of HJC derived from motion analysis were subtracted from the results estimated using each of the two equations. The resulting linear difference scores were analyzed using nonparametric 2-related Wilcoxon (SPSS v.10.0, SPSS Inc. Chicago, IL. USA). The differences of three anatomical axes are also calculated. The absolute distance along each axis was compared to determine the accuracy of each equation in the three anatomical planes and the signed distance was averaged and compared to determine the general direction of the error. The significance level was set at $P < 0.05$.

3 RESULTS

Subject coordinates for the left and right HJCs derived from motion analysis and the range of coordinate of male as well female are listed in Table 2. The mean HJC asymmetry was 1.90 ± 1.13 cm. The overall mean linear distances were 16.80 ± 1.91 cm.

Table 2. HJC location derived from motion analysis, the coordinate system origin is located halfway between the right and left ASISs.

	Posterior (x)	Lateral (y)	Inferior (z)
M1-R	–48.06	75.93	–54.60
M2-R	–52.80	68.43	–58.58
M3-R	–48.91	90.12	–51.18
M4-R	–50.86	70.87	–54.65
M5-R	–51.95	66.46	–48.99
M6-R	–45.62	94.46	–30.65
M7-R	–60.31	91.96	–51.88
M8-R	–65.17	81.11	–47.63
F1-R	–57.73	75.63	–49.73
F2-R	–60.58	69.34	–45.57
F3-R	–41.97	78.73	–58.07
F4-R	–58.10	93.17	–48.07
F5-R	–55.13	68.25	–53.65
F6-R	–48.93	82.41	–47.33
Minimum	–65.2	66.5	–58.6
Maximum	–42.0	94.5	–30.7
Harrington, 2007			
Minimum	–51.8	85.6	–73.9
Maximum	–33.4	96.1	–95.2

Table 3. Compare the difference (cm) of IAD and various errors (cm) of three anatomical axes with two predictive methods of locating the HJC.

	Bell, 1990	OrthoTrak	P value
Difference of IAD $N = 14$	0.95	1.23	0.363
Absolute errors of three anatomical axes			
Inferior-superior	2.45	3.43	0.000
Anterior-posterior	0.83	0.79	0.855
Medial-lateral	0.70	0.77	0.585
Signed errors of three anatomical axes			
Inferior-superior	2.45	3.43	
Anterior-posterior	0.44	0.44	
Medial-lateral	–0.45	–0.45	
$N = 28$			

*Significantly different from another method ($P < 0.05$).

The two predictive methods of finding HJC were first compared by calculating the linear distance of IAD derived by each predictive method and motion analysis. Comparing the linear difference of IAD scores indicated that there was a significant different between the predictive methods and the motion analysis, nonetheless this difference in Bell equation shows little clinical meaning (difference < 1 cm). No significant difference found between the two predictive methods.

The components of the HJC estimated errors along the anatomical axes were also calculated for each algorithm. In the inferior–superior direction, there was significant difference ($P < 0.001$) between the two predictive methods, and both of the equations tended to locate the HJC superior to the motion analysis. (Table 3).

Table 3 also shows the results of the signed distances different for the three anatomical axes. In the inferior–superior direction, the Bell equation located HJC superior to the OrthoTrak equation. Both of the equations predicted HJC inferior to the motion analysis. In the anterior-posterior direction, Bell equation trend to locate the HJC anterior to the motion analysis results, and OrthoTrak equation trend to locate the HJC more posterior. In the medial-lateral direction, both of the equations located HJC lateral to the motion analysis.

4 DISCUSSION

This study shows the Bell equation have good accuracy when apply on the 3D scan data (linear difference <1 cm), and the Bell equation is more accurate than the OrthoTrak equation in the inferior-superior direction.

Skeleton-driven deformation is the most universal method to drive the 3D scan data (Allen, et al., 2003). However, for this to be an accurate method, the procedure of joint calculation should be clearly defined. The predictive methods based on investigate of anatomy and biomechanics provides more accurate than creating manually. Moreover, the 3D scan data have high degree of fidelity on the bony landmarks, making the applying of predictive methods to these data more feasible.

The inaccurate of joint centers will huge inference the quality of the deformation and motion. Previous studies investigating the effect of hip joint center mislocation on kinematics and kinetics found that displacement of the HJC in the anterio–posterior and medial–lateral directions had the greatest effect at the hip and knee (Stagni, et al., 2000). This study shows the little errors (<1 cm) in these directions when using the predictive methods to estimate the HJC coordinates.

Present research used the motion analysis data to being the golden standard, since the difficult of acquire the MRI or CT images in health subjects. The motion analysis method and APAS motion analysis software was widely applied in lower extremity motion and gait analysis (Amiridis, Arabatzi, Violaris, Stavropoulos, & Hatzitaki, 2005; Lee, Kim, Kim, & Choi, 2007; Maulder, Bradshaw, & Keogh, 2008; Stuelcken, Portus, & Mason, 2005). Compare to the study of Harrington et al. (2007) which used the MRI to define the location of HJC from 7 health adults, our study shows large difference in axis *z* (this study: –30.7~–58.6; Harrington: –73.9~–95.2). This result was in accordance with previous study, the motion analysis have a trendy to locate HJC superior to the MRI or ultrasound (Hicks & Richards, 2005). The mean HJC asymmetry in this study was 1.90 ± 1.13 cm. These values are less than those reported by Leardini et al. (1999) of up to 2.59 cm. The overall mean linear distances were 16.80 ± 1.91 cm, these values are similar to that reported by Mullaji et al. (2010) of 16.0 ± 0.8 cm.

Nowadays, number of countries used 3D whole body scanners to collect anthropometric data and construct a database of 3D scanned human models (Daanen & Van De Water, 1998; Jones, West, Harris, & Read, 1989; Wang, 2005; Wang, Chang, & Yuen, 2003). The 3D scanned datasets are commonly used in anthropometric data collection, driving these human models will provide great facilitation to the designers and web shopping customers. For designers and manufacturers, the scanned examples can help them to produce well-fitted products, such as clothing and shoes (Kim & Kang, 2003; Lu & Wang, 2008; Lu, Wang, Chen, & Wu, 2010; Luximon & Luximon, 2009). For consumers, the scanned examples can help them to evaluate whether these products are suitable or not when put on their body, which is especially important in online shopping. The need for realistic integration of the costumers in virtual products try-on and operation calls for efficient and reliable techniques of to drive the scanned examples.

In conclusion, present study described a handy and accurate technique to determine the HJC of the 3D scan data using 3D surface anthropometry and the predictive methods. By measuring the IAD on the 3D scan data and applying the Bell predictive method to estimate the coordinates, this technique located the HJC on the 3D scan data accurately. This logic can also apply to other joint centers estimated. Future researches need to define the estimated methods on other joints for the whole human skeletal system defining, and the reliability of these methods also needs to investigate.

REFERENCES

Allen, B., Curless, B. & Popović, Z. 2002. *Articulated body deformation from range scan data.*

Allen, B., Curless, B. & Popović, Z. 2003. The space of human body shapes: Reconstruction and parameterization from range scans. *ACM Transactions on Graphics* 22(3): 587–594.

Amiridis, I.G., Arabatzi, F., Violaris, P., Stavropoulos, E. & Hatzitaki, V. 2005. Static balance improvement in elderly after dorsiflexors electrostimulation training. *European Journal of Applied Physiology* 94(4): 424–433.

Begon, M., Monnet, T. & Lacouture, P. 2007. Effects of movement for estimating the hip joint centre. *Gait and Posture* 25(3): 353–359.

Bell, A.L., Pedersen, D.R. & Brand, R.A. 1990. A comparison of the accuracy of several hip center location prediction methods. *Journal of Biomechanics* 23(6): 617–621.

Camomilla, V., Cereatti, A., Vannozzi, G., & Cappozzo, A. (2006). An optimized protocol for hip joint centre determination using the functional method. *Journal of Biomechanics* 39(6): 1096–1106.

Cappozzo, A., Catani, F., Della Croce, U. & Leardini, A. 1995. Position and orientation in space of bones during movement: Anatomical frame definition and determination. *Clinical Biomechanics* 10(4): 171–178.

Daanen, H.A.M. & Van De Water, G.J. 1998. Whole body scanners. *Displays* 19(3): 111–120.

Davis Iii, R.B., Õunpuu, S., Tyburski, D. & Gage, J.R. 1991. A gait analysis data collection and reduction technique. *Human Movement Science* 10(5): 575–587.

Ehrig, R.M., Taylor, W.R., Duda, G.N. & Heller, M.O. 2006. A survey of formal methods for determining the centre of rotation of ball joints. *Journal of Biomechanics* 39(15): 2798–2809.

Harrington, M.E., Zavatsky, A.B., Lawson, S.E.M., Yuan, Z. & Theologis, T.N. 2007. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. *Journal of Biomechanics* 40(3): 595–602.

Hicks, J.L. & Richards, J.G. 2005. Clinical applicability of using spherical fitting to find hip joint centers. *Gait and Posture* 22(2): 138–145.

Jones, P.R.M., West, G.M., Harris, D.H. & Read, J.B. 1989. The loughborough anthropometric shadow scanner (LASS). *Endeavour* 13(4): 162–168.

Kim, S.M. & Kang, T.J. 2003. Garment pattern generation from body scan data. *CAD Computer Aided Design* 35(7): 611–618.

Kirkwood, R.N., Culham, E.G. & Costigan, P. 1999. Radiographic and non-invasive determination of the hip joint center location: Effect on hip joint moments. *Clinical Biomechanics* 14(4): 227–235.

Leardini, A., Cappozzo, A., Catani, F., Toksvig-Larsen, S., Petitto, A., Sforza, V., et al. 1999. Validation of a functional method for the estimation of hip joint centre location. *Journal of Biomechanics* 32(1): 99–103.

Lee, J.Y., Kim, G., Kim, J.H. & Choi, S.H. 2007. Kinematic gait analysis of the hind limb after tibial plateau levelling osteotomy and cranial tibial wedge osteotomy in ten dogs. *Journal of Veterinary Medicine Series A: Physiology Pathology Clinical Medicine* 54(10): 579–584.

Lin, J.D., Chiou, W.K., Weng, H.F., Fang, J.T. & Liu, T.H. 2004. Application of three-dimensional body scanner: observation of prevalence of metabolic syndrome. *Clinical Nutrition* 23(6): 1313–1323.

Lu, J.M. & Wang, M.J.J. 2008. Automated anthropometric data collection using 3D whole body scanners. *Expert Systems with Applications* 35(1–2): 407–414.

Lu, J.M., Wang, M.J.J., Chen, C.W. & Wu, J.H. 2010. The development of an intelligent system for customized clothing making. *Expert Systems with Applications* 37(1): 799–803.

Luximon, A. & Luximon, Y. 2009. Shoe-last design innovation for better shoe fitting. *Computers in Industry* 60(8): 621–628.

Maulder, P.S., Bradshaw, E.J. & Keogh, J.W.L. 2008. Kinematic alterations due to different loading schemes in early acceleration sprint performance from starting blocks. *Journal of Strength and Conditioning Research* 22(6): 1992–2002.

Mullaji, A., Shetty, G.M., Kanna, R. & Sharma, A. 2010. Variability in the range of inter-anterior superior iliac spine distance and its correlation with femoral head centre. A prospective computed tomography study of 200 adults. *Skeletal Radiology* 39(4): 363–368.

Piazza, S.J., Okita, N. & Cavanagh, P.R. (2001). Accuracy of the functional method of hip joint center location: Effects of limited motion and varied implementation. *Journal of Biomechanics,* 34(7): 967–973.

Seidel, G.K., Marchinda, D.M., Dijkers, M. & Soutas-Little, R.W. 1995. Hip joint center location from palpable bony landmarks - A cadaver study. *Journal of Biomechanics* 28(8): 995–998.

Stagni, R., Leardini, A., Cappozzo, A., Grazia Benedetti, M. & Cappello, A. 2000. Effects of hip joint centre mislocation on gait analysis results. *Journal of Biomechanics* 33(11): 1479–1487.

Stuelcken, M.C., Portus, M.R. & Mason, B.R. 2005. Off-side front foot drives in men's high performance cricket. *Sports biomechanics / International Society of Biomechanics in Sports* 4(1): 17–35.

Wang, C.C.L. 2005. Parameterization and parametric design of mannequins. *CAD Computer Aided Design* 37(1): 83–98.

Wang, C.C.L., Chang, T.K.K. & Yuen, M.M.F. 2003. From laser-scanned data to feature human model: A system based on fuzzy logic concept. *CAD Computer Aided Design* 35(3): 241–253.

Zhang, X., Lee, S.W. & Braido, P. 2004. Towards an integrated high-fidelity linkage representation of the human skeletal system based on surface measurement. *International Journal of Industrial Ergonomics* 33(3): 215–227.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Experimental study on the relationship between hand-grip strength and stature

Liao K. Hsi
Department and Institute of Management, Diwan University, Taiwan

ABSTRACT: Hand-grip strength has received increasing attention from physicians and ergonomic researchers. Most studies try to associate hand-grip strength and anthropometric variables. Stature is an important indicator of human growth. However, the relationship between grip strength and stature was obviously unclear. The purpose of this study was to explore the statistically relationship between grip strength and stature. 200 healthy subjects (117 males and 83 females, age 20.10 ± 1.00 years) participated this study and were divided into three groups by stature: low, medium, and high. Maximum volitional contraction of grip (MVC) was measured using handgrip dynamometer. Five consecutive measurements were made in both the left and the right hands. The results showed that MVC was significantly and positively correlated with stature (r = .722, $p < .01$), and the three groups showed significant differences on MVC with ANOVA method (F = 10.238, $p = .000$). The study conducted ANCOVA based on the covariance of gender and body weight, the three groups also showed significant differences on MVC (F = 47.763, $p = .000$). The findings would serve as a reference for hospitalized patients as well as industrial staffs.

Keywords: hand-grip strength, maximum volitional contraction of grip, gender, stature, body weight

1 INTRODUCTION

1.1 *Hand-grip strength*

Hand-grip strength is an important skill for the human body to control objects. It assesses skeletal muscle function and has received increasing attention from physicians and ergonomic researchers in the last years (Schlüssel, Anjos & Kac, 2008). Recently, it has deserved particular attention as an indicator of nutritional status, particularly in hospitalized patients. Bohannon (2008) summarized the literature addressing the value of grip strength as a predictor of important outcomes. For example, hand-grip strength of cirrhotic patients has proven to be very low when compared to the predicted value by age (Coporrino, Faloppa, Réssio, Soares, Nakachina & Segre, 1988). Decreased strength, most often grip strength, has been prescribed as an important sign of frailty (Hirsch, Anderson, Newman, Kop, Jackson, Gottdiener, Tracy & Fried, 2006). Grip strength has been shown to be a legitimate indicator of nutritional status that may in some populations predict outcomes better than traditional nutritional markers such as weight to height ratios, weight loss, limb circumference, or serum albumin (Klidjian, Foster, Kammerling, Cooper, & Karran, 1980). Ergonomic field researchers were also much consideration to MVC. For example, Mahmoud, Purswell, & Schlegel (1996) demonstrated that carpal tunnel syndrome (CTS) and other cumulative trauma disorders (CTD's) were affected by MVC. Liao (2008) demonstrated that MVC was relevant with eye-hand coordination task performance. From above discussion, Hand-grip strength expressed important function in the physician and ergonomic field.

1.2 *Grip strength and anthropometric variables*

Most studies try to associate hand-grip strength and anthropometric variables such as gender, height, weight, etc., in order to predict the outcomes. There were many factors effect the hand-grip strength measurement, for example, gender, age, height, body weight and the dominant hand of the subjects (Chang, 2002; Liao, 2009). Stature (body height) is an important indicator of human growth, particularly in the younker. The peak of the curve of stature occurred at 15 years old for both males and females (Chang, 2002). Shih, Fu, & Wang (1996) have also indicated that the grip strengths are related with the heights of hand elbows. As a result, researcher infers that the height of a person may be possibly one of the influential factors on grips. Luna-Heredia, Martín-Peña, & Ruiz-Galiana (2005) found that grip strength in healthy individuals was correlated positively with height. Liao (2009) reported that MVC was in positive correlation with height and body weight. Although there were much research conducted the correlation between grip strength and height, however, the covariance effect among gender, stature and weight were seldom discussed statistically. In other words, the relationship between grip strength and stature was interesting and unclear. It is an important key point to understand the actually relation based on ANCOVA statistic method. The purpose of this study was to explore the statistic relationship between grip strength and stature. In other words, the aim was to understand whether the grip strength was actually affected by the body height or not, when expulsion the gender and body weight as co-variables. The result would serve as a reference for hospitalized patients as well as ergonomic industrial staffs.

2 RESEARCH METHOD

2.1 *Subjects*

There were a total of 200 subjects in the participation of related experiments, in which males are 117 and females are 83, with ages between 18 and 27 years old. The related anthropometric data of the subjects being tested are as Table 1 shown. The subjects being tested do not have any injuries related to the parts of muscles and joints.

2.2 *Experiment design*

The study includes one experiment making evaluations on the value of MVC. On the usage aspect of instrument equipments for MVC adopted hand grip dynamometer in minute style (Japanese style, Tkk 5001), as the examination tool. While doing tests, the handle diameter of hand grip dynamometer is adjusted to 50.8 mm (Grant, Habes, & Steward, 1992; Shih, Fu, & Wang, 1996). The reason to choose this handle scale is due it can produce the MVC, as well as the muscles are most unlikely to get tired (Ayoub & LoPresti, 1971; Liao, 2008). The related equipments and experiment designs of this study are illustrated in details in the Table 2. On the aspect of experiment design, the dependent variable of the experiment is MVC. The independent variable is stature (low, medium, and high). The moderator variables are gender (male/female) and body weight. To conduct the relationship between MVC and stature, researcher was used ANCOVA method and took gender and body weight as co-variable.

Table 1. Anthropometric data of the subjects being tested (117 Male, 83 Female).

Item	Gender	Average (S.D.)	Range
Age	Male	20.20 ± 0.98	18–27
	Female	19.96 ± 1.01	18–25
Stature (cm)	Male	171.31 ± 5.29	162–185
	Female	159.11 ± 6.16	151–175
Weight (kg)	Male	73.53 ± 9.21	51–85
	Female	47.67 ± 7.81	40–65

Table 2. Related equipments and experiment designs in the testing.

	Hand-grip strength
Method of measurement	After securing hand grip dynamometer with palm, to then have thumb in opposite grip with the other four fingers
Instrument	Hand grip dynamometer (Tkk 5001)
Experiment designs	Correlation, ANOVA, ANCOVA (Gender & weight are co-variables), Regression
Dependent variables	MVC
Independent variables	Stature

2.3 *Measurement procedures*

The experiment steps and procedures are refined and implemented according to the methods of pre-researchers (Caldwell, Chaffin, Dukes-Dobos, Kroemer, Laubach, Snook, & Wasserman, 1974; Liao, 2009). The testing position for MVC was done with sitting position while the subject's arms are naturally straight out in downward inclination. Before doing tests, to firstly let the subjects have sufficient rest, then to get familiar with the experiment procedures, afterwards, doing tests of the individual MVC. While making force applications, to hold tightly on the hand grip dynamometer around 3 seconds, then to get relax, moreover, to continue with such operations for 5 times, and with each operation to take a break for 10 seconds. After consecutively doing experiment for 30 minutes, to take a break of 5 minutes at least.

3 RESULTS AND DISCUSSIONS

The results and discussions of this study are aimed on the four aspects, correlation analysis, ANOVA, ANCOVA, and regression analysis of MVC, and to work on the illustrations respectively.

3.1 *Pearson correlation analysis test*

The mean and standard deviation of MVC of five consecutive measurements in both the left and the right hands are listed in Table 3 in below. The Average of five consecutive measurements of Left-hand MVC is 33.53 kg, and the Right-hand MVC is 37.28 kg. And right-and-left hand MVC is 35.41 kg.

The summary table of the Pearson Correlation is listed in the following (See Table 4). The Correlation Matrix shows (see Table 4) that the dependent variable (MVC) is positively correlated with stature and weight, and is negatively correlated with gender. The independent variables stature was positively correlated with weight and negatively correlated with gender.

3.2 *ANOVA analysis*

This study makes investigation of differences on MVC among the three groups by way of ANOVA. Subjects were divided into three groups by stature. Subject stature below 25% was assigned to "low stature group", above 25% was assigned to "high stature group", and the rest 50% subjects was assigned "medium stature group". The basic descriptive statistical facts of three groups are listed in Table 5.

When doing Analysis of Variance (ANOVA), it is required to take into consideration if the sampling is meeting with the basic assumption of ANOVA, including homogeneity. The verification of homogeneity was used the Levene Test. The results of Levene Test for homogeneity of variances, $F = 1.580$, $p = .209 > .05$. Thus, the dependent samples were not in violation of basic assumption

Table 3. Mean and standard deviation of MVC statistics.

MVC	Mean (kg) (5 times)	S.D.	N
Left-hand	33.53	9.19	200
Right-hand	37.28	9.15	200
Average	35.41	9.03	200

Table 4. The pearson correlation matrix.

	MVC[a]	Gender	Stature
MVC			
Gender	−.867**		
Stature	.722**	−.729**	
Weight	.782**	−.828**	.793**

a = Average of Left-hand MVC and Right-hand MVC.
** Correlation is significant at the 0.01 level (2-tailed).

Table 5. Basic descriptive statistical facts of three different stature groups.

Group	Ratio	N	Stature (cm)	Age
Low stature	Below 25%	50	160.5 ± 6.14	19.9 ± 1.00
Medium stature	Medial 50%	100	166.2 ± 8.56	20.1 ± 1.00
High stature	Above 25%	50	172.1 ± 4.78	20.3 ± 0.97
Total	100%	200	166.2 ± 8.26	20.1 ± 1.00

of Analysis of Variance (Girden, 1992), and it is okay to do further ANOVA analysis. The results of three different stature groups' ANOVA in the Maximum volitional contraction of grip (MVC) is listed in Table 6 in below.

By looking at Table 6, three groups are significant differences on MVC ($F_{.95\ (2,\ 197)} = 10.238$, $p = .000$). However, the Post-Hoc Multiple Comparisons showed that the high stature group was stronger than medium stature group (H > M) and high stature group was stronger than the low stature group (H > L).

3.3 *ANCOVA analysis*

For ANCOVA analysis the study has to pick gender and weight as the moderator variables to work on the Analysis of Covariance. The selecting principles of moderator variables are worked according to the recommendations of Bryman and Cramer (1997). On the aspect of Homogeneity test, it is mainly adopted with homogeneity of with-in regression. For the test results, the group of "stature * gender", "stature * weight", the homogeneity of with-in regression of two groups are with p value greater than .05 ($p = .910$ and $p = .260$, respectively). Thus, null hypothesis is accepted. That means the slope of regression line for each group being the same. In other words, after eliminating the interference of gender and weight, the MVC of each group will not change due of differences on each handling level of each independent variable, as a result, it is okay to work on ANCOVA. Thus, the basic assumption test on hand-grip strength is suitable to do ANCOVA (Joseph, Rolph, Ronald, & William, 1995).

The results of ANCOVA in the Maximum volitional contraction of grip (MVC) is listed in Table 7 in below. In Table 7, as the results done with ANCOVA, after excluding the interference of "gender", and "weight", there were significant differences on three MVCs' values, left-hand, right-hand, and two-hand from stature.

By looking at Table 7, for the results of ANCOVA are described as following. First, the

Table 6. Three different stature groups' ANOVA results in the MVC (two hands' value).

Source	SS	df	MS	F	Post-Hoc
Between Groups	1526.656	2	763.328	10.238***	H > M, H > L
Within Groups	14688.656	197	74.562		
Total	16215.311	199			

H = High stature, M = Medium stature, L = Low stature.

*** The significant at the .001 level (2-tailed).

Table 7. ANCOVA analysis with three different hands' MVCs.

Source	SS	df	MS	F	Sig.
Left-hand	156.069	2	782.534	56.314	.000
Right-hand	978.072	2	489.036	27.835	.000
Two-hand	1253.659	2	626.830	47.763	.000

test of between-subjects effects of Left-hand MVC is reaching significant level, $F_{.95\ (2,\ 197)} = 56.314$, $p = .000 < .001$. Second, the test of between-subjects effects of Right-hand MVC is reaching significant level, $F_{.95\ (2,\ 197)} = 27.835$, $p = .000 < .001$. Third, the test of between-subjects effects of Two-hand MVC is reaching significant level, $F_{.95\ (2,\ 197)} = 47.763$, $p = .000 < .001$. This indicates that the Low-stature group, Medium-stature group, and High-stature are reaching significant levels ($p < .001$). The result is the same trend as the ANOVA analysis has found. As an increasing stature, it would be increasing the hind-grip strength. The finding of the present study may get support from many researches (Luna-Heredia, Martín-Peña, & Ruiz-Galiana, 2005; Liao, 2009).

3.4 *Regression analysis*

Regression analysis is used gender, stature, and weight as independent variables to predict the dependent variable MVC. Based on the findings illustrated in Table 8, in regard to the study's determinants and two-hand MVC, the coefficient of the determinant (R^2) was .772 and the adjusted R^2 was .769. As it regarded two-hand MVC, the F-ratio value was 221.432 ($p < .001$). This was a clear sign of significance, based on the assessment of the determinants against MVC.

In order to assess the relative importance, beta coefficients are used: the higher the beta coefficients, the more important each factor. It is important to note that the gender and stature factors turned out to be statistically significant ($p < .05$) antecedents influencing MVC. The most important factor affecting MVC was found to be

Table 8. Regression of Two-hand MVC[a].

Independent Variable	b	Beta	t	Sig.
Constant	22.167		2.183	.030
Gender	–12.139	–.664	–10.666	.000
Stature	.157	.144	2.507	.013
Weight	6.931E-02	.118	1.694	.092

a: $R = .879$, $R^2 = .772$, adj. $R^2 = .769$, $F = 221.432$, $p < .001$.

"gender" with a beta value of –.664 followed by "stature" with .144. Weight value with a beta value of .118 was found to be the least important key factor affecting MVC.

It can be concluded that, in descending order, gender, stature, and weight value were found to be important determinants of MVC.

The standardized regression equation could be written as following: Equation (1):

$$\text{Two-hand MVC (kg)} = 22.167 - .664 \times \text{gender (male = 1, female = 2)} + .144 \times \text{stature (cm)} + .118 \times \text{weight (kg)} \quad (1)$$

4 CONCLUSIONS AND RECOMMENDATIONS

From above results and discussions, this study has come to the following conclusions. There is a positive correlation between MVC and stature ($p < .01$) or weight ($p < .01$), and there is a negative correlation between MVC and gender ($p < .01$). Thus, when evaluating human Maximum volitional contraction of grip (MVC), and when assigning works in the job operations, it is required to take into consideration of factors such as gender, stature and body weight.

ANOVA analysis indicated that three different stature groups, low, medium, and high were significant differences on MVC ($F_{.95\ (2,\ 197)} = 10.238$, $p = .000$). However, the Post-Hoc Multiple Comparisons showed that the high stature group was stronger than the medium stature group (H > M) and the high stature group was stronger than the low stature group (H > L).

For ANCOVA analysis the study has to pick gender and weight as the moderator variables to work on the Analysis of Covariance. ANCOVA analysis results indicated that three different stature groups, low, medium, and high were all significant differences (p = .000) on Left-hand MVC, Right-hand MVC, and Two-hand MVC. This result is meeting with the study of ANOVA analysis.

Regression analysis results indicated that the three independent variables, gender, stature, and weight could explain the variances of dependent variable MVC reaching to 77.2%. The most important factor affecting MVC was found to be "gender", followed by "stature". It can be concluded that gender and stature were found to be important determinants of MVC. The standardized regression equation, Equation (1), that this study had suggested could be used as reference for MVC measurement.

In summary, the study used several methods, such as correlation analysis, ANOVA, ANCOVA, and regression analysis, to explore the statistically relationship between hand-grip strength and stature. The study finally found that when expulsion the gender and body weight as co-variables, the grip strength was actually affected by the stature and the relationship between MVC and stature was obviously positively correlated (p < .01). That means that high stature people would be shown having strong Maximum volitional contraction of grip (MVC).

The research data obtained in this study can be offered as reference for the industries and relevant personnel on the selection/training and hand tools/equipment designs. And for hospital patients, when usage grip strength as outcomes to test their sign of frailty, it is required to considerate the stature and gender affecting.

It would be of great benefit to re-conduct this study in the future to measure and compare the findings over two or more occurrences. Although findings from this study do not illustrate concerns to consider as regards age, and dominant hand of the subjects, it is of great assurance to re-conduct this study before implementing the model constructed from this study. For further research, an additional study could also be conducted to measure MVC with age and other variables. That would provide insight as to its contribution to this assessment.

ACKNOWLEDGMENT

I would like to give my special thanks to National Science Council, Republic of China for funding the implementation of the research plan. The plan code: NSC-96-2221-E-434-006.

REFERENCES

Ayoub, M.M. & LoPresti, P. 1971. The determination of optimal size cylindrical handle by use of EMG. *Ergonomics* 14(4): 509–518.

Bohannon, R.W. 2008. Hand-grip dynamometry predicts future outcomes in aging adults. *Journal of Geriatric Physical Therapy* 31(1): 3–10.

Bryman, A. & Cramer, D. 1997. *Quantitative Data Analysis with SPSS for Window.* London: Rout ledge.

Caldwell, L.S., Chaffin, D.B., Dukes-Dobos, F.N., Kroemer, K.H.E., Laubach, L.L., Snook, S.H. & Wasserman, D.E. 1974. A proposed standard procedure for static muscle strength testing. *American Industrial Hygiene Association Journal* 35(4): 201–206.

Chang, S.Y. 2002. Grip and key pinch strength: norms for 7 to 22 years-old students in Taiwan. *Tzu Chi Medical Journal* 14(4): 241–252.

Coporrino, F.A., Faloppa, B.G.S., Réssio, C., Soares, F.H.C., Nakachina, L.R. & Segre, N.G. 1988. Estudo populacional de força de preensão palmar com dinamómetro Jamar®. *Rev Bras de Ortop,* 33(2): 150.

Girden, E.R. 1992. *ANOVA Repeated measures*, Newbury Park, CA: Sage.

Grant, K.A., Habes, D.J., & Steward, L.L. 1992. An analysis of handle designs for reducing manual effort: the influence of grip diameter. *International Journal of Industrial Ergonomics* 10: 199–206.

Hirsch, C., Anderson, M.L., Newman, A., Kop, W., Jackson, S., Gottdiener, J., Tracy, R. & Fried, L.P. 2006. The association of race with frailty: The cardiovascular health study. *Ann Epidemiol* 16(7): 545–553.

Joseph, F.H., Rolph, E.A., Ronald, L.T. & William, C.B. 1995. *Multivariate data analysis with readings,* Englewood Cliffs, New Jersey.

Klidjian, A.M., Foster, K.J., Kammerling, R.M., Cooper, A. & Karran, S.J. 1980. Relation of anthropometric and dynamometric variables to serious postoperative complications. *British Medical Journal* 281: 899–901.

Liao, K.H. 2008. *The correlative study of the spatial ability and hand-gripping control on eye-hand coordination task.* Chinese Institute of Industrial Engineers Conference, Taiwan, Dec. 13, 2008.

Liao, K.H. 2009. *A Study Concerning How Gender, Hands and the Sequence of Force Application Affected Grip and Hand-gripping Control*. The 10th Asia Pacific Industrial Engineering & Management Systems Conference (APIEMS) in Kitakyushu, Japan, Dec. 14–16, 2009.

Luna-Heredia, E., Martín-Peña, G. & Ruiz-Galiana, J. 2005. Handgrip dynamometry in healthy adults. *Clin Nutr*. 24(2): 250–258.

Mahmoud, A.A., Purswell, J.L. & Schlegel, R.E. 1996. Psychophysically determined work-cycle parameters for repetitive hand gripping. *International Journal of Industrial Ergonomics* 17(1): 35–42.

Schlüssel, M.M., Anjos, L.A. & Kac, G. 2008. Hand grip strength test and its use in nutritional assessment. *Rev. Nutr.* 21(2): 233–235.

Shih, Y.C., Fu, S.L. & Wang, M.J. 1996. The effect of hand posture on grip, pinch and torque exertions capabilities of hand. *Journal of Occupational Safety and Health*, 4(1): 43–58.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Investigation of trunk muscular activities during manual lifting with two elastic back belt types

Nopporn Kurustien, Keerin Mekhora & Wattana Jalayondeja
Faculty of Physical Therapy, Mahidol University, Thailand

Masaharu Kumashiro & Hiroyuki Izumi
Department of Ergonomics, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Japan

Suebsak Nanthavanij
Engineering Management Program, Sirindhorn International Institute of Technology, Thammasat University, Thailand

ABSTRACT: Nowadays, back belts are commonly used among industrial manual handling workers to prevent low back injury. This study aims to investigate how the lumbo-sacral belt and the pelvic belt affect trunk muscular activities during manual lifting. Both belts are elastic type but differ in the belt width and wearing position. Nine Japanese healthy male subjects whose ages range between 21 and 27 years participated in this study. The results showed that there were significant decreases ($p < 0.01$) of Transversus Abdominis (TrA)/Internal Oblique (IO) and Erector Spinae (ES) activities when wearing the pelvic belt and lifting 20-kg loads using the squat posture, and lifting 20-kg loads using the standing posture, respectively. All other conditions showed similar trend of decreases in trunk muscular activities.

Keywords: back belt, manual lifting, lumbo-sacral belt, pelvic belt, electromyography

1 INTRODUCTION

In Thailand, many industries commonly provide back belts to their materials handling workers to prevent back injury. Previous studies have reported many benefits of wearing a back belt such as reminding to lift properly, supporting the spine from shear loads, reducing compressive loading of the lumbar spine by increasing Intra-Abdominal Pressure (IAP), acting as a splint, reducing the range of motion, providing warmth to the lumbar region, enhancing proprioception via the pressure, reducing muscular fatigue, and providing stiffening to the torso (McGill 2002 & McGill 2005).

An ability to increase Intra-Abdominal pressure (IAP) was one of the most cited mechanisms of the back belt that helps to reduce compression force on the spinal column and assist back extensor muscles in producing extension torque. The IAP had been proved by many studies that it can provide lumbar stabilization, decrease compressive load on lumbar spine by stiffening the trunk, and contribute, to some extents, to generate spinal extensor moment by exerting a hydraulic force upon the diaphragm which acts anterior to the spine (Cholewicki 1999, Miyamoto 1999, Poppel 2000).

Transversus Abdominis (TrA) was found to play an important role in spinal stabilization since nearly two decades ago (Richardson 2004). Electromyographic studies during manual lifting showed parallel activity of the Transversus Abdominis (TrA) and internal Oblique Muscles (IO) when IAP increased, whereas the Rectus Abdominis (RA) muscle was not seen to work during lifting (Sullivan 1989). The TrA has feed-forward mechanism to stiffen the spinal column without being influenced by the direction of arm movement (Hodges 1997). However, this feed-forward could limit or delay an onset of activation time in people with back problems, prior to rapid arm movement (McGill 2002). At present, there have been no studies which investigate the TrA muscle activity in back belt usage.

Previous study indicated that elastic back belt is more effective than other types of back belt in decreasing back muscle activities (Granata 1997). In general, there are two types of elastic back belt: (1) lumbo-sacral belt, and (2) pelvic belt. Both belts differ in the belt width and wearing position.

The former is wider and is used for supporting the whole lumbar spine and the upper part of the sacrum, whereas the latter is narrower and is used for supporting only the pelvic and the lower part of the lumbar spine. Thus far, the effects of these two back belt types on trunk muscular activities have never been investigated before.

This study aims to investigate how the lumbo-sacral belt and pelvic belt affect the trunk muscular activities especially the feed-forward mechanism of TrA during manual lifting.

2 METHODS

2.1 *Experimental design*

A 2 × 3 × 2 (lifting position × back belt type × load level) within-subject design was used. Each lifting task involved 3 lifts with a 20-second rest period in between. Between successive lifting tasks, the resting of at least 3 minutes was provided to prevent fatigue. The EMG with the lift-off switch signal was recorded in all lifting tasks.

2.2 *Participants*

Nine healthy male volunteers were recruited from Japanese college students, with their ages ranged from 21 to 27 years. The mean height and weight of the participants were 1.71 ± 0.06 m and 61.34 ± 6.77 kg, respectively; the mean BMI was 21.07 ± 2.17. The participants were screened for past history of back pain. Only those with no back pain during the past 6 months were permitted to participate. All volunteers signed an informed consent in which its protocol has been approved by the ethic committee of Mahidol University.

2.3 *Instruments*

- Electromyography (EMG): TEAC LX-10 series recorder (TEAC Corporation, Tokyo, JAPAN) and EMG transmitter (BA1104 + TU-4, DIGITEX LAB. Co., Ltd., Tokyo, JAPAN) with active electrode.
- Material for lifting: 40 × 27 × 23 cm^3 plastic basket with 10 kg and 20 kg weights from barbell.
- Metronome (MP-150, Yamaha, Kyohan Co., Ltd.): Rhythm of 48 beats/minute was set to control the speed of lifting as 1 lift per 1.25 seconds.
- Back belts used in this experiment include: (1) lumbo-sacral belt (Multi Macro Co., Ltd., THAILAND), and (2) pelvic belt (MIDORI ANZEN Co., LTD., JAPAN). The fastening mechanism consists of 2 sets of Velcro straps; the main one is on the front and the more lateral enhancing one is on each side. Its back comes with four semi-rigid bars aligned on the back, parallel to the spine.

Lumbo-sacral belt. Lumbo-sacral belt is commonly used among Thai industrial workers. It has a posterior height of 20 cm and an anterior height of 12 cm. The belt is positioned by aligning the middle point between the upper and lower rings of the belt with the iliac crest to cover most of the span from the thorax to the pelvis (Fig.1, left).

Pelvic belt. Pelvic belt is commonly used among Japanese industrial workers. It has a posterior height of 14.5 cm and an anterior height of 9 cm height at the middle region of abdomen. The belt is positioned by aligning the upper ring of the belt at the iliac crest in order to stabilize the lower part of lumbar and pelvis (Fig.1, right).

2.4 *Procedure*

Initially, the examiner informed each participant about the purpose, procedure, benefit, and risks that may occur from the research. After being informed, the participant then signed the consent form.

The participant's skin was carefully prepared to reduce its impedance to 5 kΩ or lower by shaving excess hair with new razor blade, rubbing the skin with fine sand paper, and cleaning it with alcohol. Surface electrodes were applied with inter-electrode spacing of 2 cm in the direction of the muscle fibers.

For each participant, 5 pairs of electrodes were placed in the following positions by an experienced physical therapist: (1) External abdominal Oblique (EO), in direct line with the umbilicus (approximately 12–15 cm from umbilicus) (Marshall 2003); (2) Transversus Abdominis/Internal Abdominal oblique (TrA/IO), 2 cm medial and inferior to the right ASIS (Marshall 2003); (3) Rectus Abdominis (RA), 3 cm superior to the umbilicus and 2 cm lateral from midline (Marshall 2003); (4) Erector Spinae (ES), 2 cm laterally from the midline through the L3 spinal process (Arokoski 2001); (5) Multifidus (MF), the same level as the L5 spinous process, 2 cm laterally from the midline

Figure 1. Lumbo-sacral belt (left), and pelvic belt (right).

(Arokoski 2001, Arokoski 2004, Hodges 1997). The reference electrode was placed on the iliac crest (for RA, IO, and EO) and posterior superior iliac spine (PSIS) (for ES and MF). Each pair of electrodes were carefully covered with kinesiology tape to prevent any effect of belt pressure and slippery. All electrode wires were also kept in place with adhesive tape.

MVC Testing. After skin preparation, the participant's muscles were tested as follows: (1) RA, by raising the upper body straight upward against resistance from strap at the chest, in supine lying with knee fixing by strap; (2) IO, by raising the left shoulder upward to the right knee against resistance in supine lying with knee fixing by strap; (3) EO, by raising the right shoulder upward to the left knee against resistance in supine lying with knee fixing by strap; (4) ES and MF, by raising the upper back up against resistance in prone lying position with strap fixing at mid calf. For each testing position, the participant was instructed to perform three isometric contractions, each at his maximum effort for 3 seconds, with a 3-minute rest period in between.

Lifting Task. All lifting tasks were administered in a random order throughout the experiment. The back belt conditions were: no belt, pelvic belt, and lumbo-sacral belt. The load levels were: 10 kg and 20 kg. The lifting postures were: (1) standing and lifting the basket from the table (with the table height at the level of the tip of the middle digit for each participant) to the elbow level (with the elbows flexed at 90° (Fig. 2)), and (2) squatting and lifting the basket from the floor to the knuckle height in standing. In addition, lifting in standing posture was employed in this study to minimize the effects of muscle length change from trunk movement during lifting. The number of lift was 3 times with a 20 second rest period between two successive lifts. For the standing posture, the participant was instructed to hold the basket with his elbows flexed at 90° for 3 seconds without any trunk movement, lower the basket back on the table, and then rest in the standing posture for 20 seconds. For the squat posture, the speed of lift was control by a metronome set at 48 beats/minute or 1 lift per 1.25 seconds. From the standing posture, the participant squatted down and lifted the basket from the floor. He squatted down again to put the basket back on the floor. Then, he would stand up and rest for 20 seconds. After finishing the 3 lifts for each lifting condition, a rest period of at least 3 minutes was provided to the participant in comfortable sitting position.

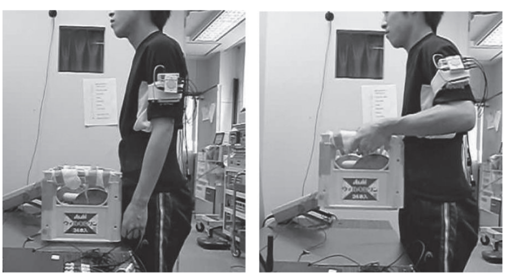


Figure 2. Lifting task in the standing posture: starting position (left) and finishing position with the elbows flexed at 90° (right).

A video camera was used to record all movements in a lateral view during lifting both in the standing and squat postures. The participant was instructed and trained to lift only in a sagittal plane and to keep his back straight.

Belt Position. Both types of back belt were put on the participant by the same physical therapist to assure that the position and pressure would be the same. For the pelvic belt, the upper border was place at the iliac crest level. For the lumbo-sacral belt, the midline between the upper and lower borders was placed at the iliac crest. The belt pressure was also checked every time by hand-checking and by asking the participant to assure that the belt tightness was comfortable.

2.5 *Data processing and analysis*

Raw EMG signals were processed using DADiSP SE-Pro2002. They were high-pass filtered above 5 Hz.

The signals of the switch installed at the base of the plastic basket to mark the timing of lift-off (an instance that the basket was lifted from the floor) and contact (an instance that the basket was put back on the floor). Based on the lift-off and contact signals, the EMG signals were collected. The *rms* data from the 3 lifts was averaged and normalized by the % MVC. The MVC was determined using the *rms* of the 1-second highest amplitudes for each isometric contraction test for each muscle. Then, the mean *rms* EMG of the two highest MVCs was used to normalize the *rms* EMG of the lifting data.

The onset of EMG was defined as the point where the mean of 75 consecutive samples (50 ms) exceeded the baseline activity by three Standard Deviations (SD). The baseline activity was defined as a period of 1000 ms before a lift-off position. Then, the latency onset was determined using the following formula:

Latency onset = Lift-off time–the point where the mean of 75 consecutive samples exceeds the baseline activity by 3 SDs.

Since this study is only concerned with the feed-forward activity of TrA, only the EMG activity that occurs before the lift-off time was considered.

2.6 *Statistical analysis*

A Statistical analysis was conducted using SPSS for Windows (Version 16). A paired student t-test was used to compare the average *rms* and latency onset among three back belt conditions. The data is presented as the mean ± SD throughout the paper.

3 RESULTS

Table 1 presents the means and standard deviations of the normalized EMG (NEMG) (in %MVC) of all trunk muscles. It shows that when the load was increased from 10 to 20 kg, the NEMG also increased in all trunk muscles. The paired t-test comparison between each pair of back belt conditions (no belt *vs*. pelvic belt, no belt *vs*. lumbo-sacral belt, and pelvic belt *vs*. lumbo-sacral belt) indicated a significant effect of back belt condition.

There was a significant decrease (p = 0.001, p = 0.005) of TrA/IO and ES muscle activation in the standing and squatting lifting postures, respectively, with a 20 kg load. The same trend of TrA/IO and ES decreased activation when wearing the pelvic belt was also found in all other postures and load conditions (Fig. 3 and Fig. 4).

Table 2 presents the latency onset of TrA/IO relative to the lift-off time. The result showed a significant effect of back belt conditions. Wearing a lumbo-sacral belt significantly decrease the onset of TrA/IO in all lifting task conditions ($p < 0.05$). The same trend direction was also found when wearing a pelvic belt (Fig. 5).

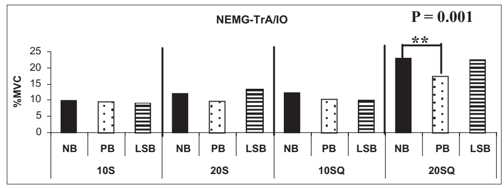


Figure 3. Comparison of %MVC of TrA/IO (Transversus abdominis/internal abdominal oblique muscle) among the 3 back belt conditions (NB: No Belt, PB: Pelvic Belt, LSB: Lumbo-Sacral Belt), and 2 load – 2 lifting posture conditions (10S: 10 kg in standing, 20S: 20 kg in standing, 10SQ: 10 kg in squatting, 20SQ: 20 kg in squatting) (n = 9).

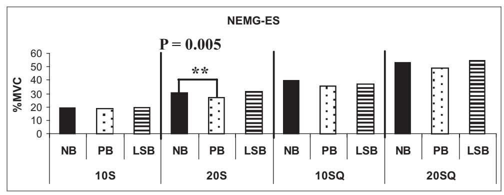


Figure 4. Comparison of %MVC of ES (external abdominal oblique muscle) among the 3 back belt conditions (NB: No Belt, PB: Pelvic Belt, LSB: Lumbo-Sacral Belt) and 2 load – 2 lifting posture conditions (10S: 10 kg in standing, 20S: 20 kg in standing, 10SQ: 10 kg in squatting, 20SQ: 20 kg in squatting) (n = 9).

Table 1. Root mean square (rms) EMG (%MVC) in 2 lifting postures, 2 loads (10 and 20 kg), and 3 back belt conditions (NB: No Belt, PB: Pelvic Belt, and LS: Lumbo-Sacral Belt) (n = 9).

Lifting posture	Load (kg)	Belt	RMS EMG of Trunk Muscles (%MVC)									
			RA		TrA/IO		EO		ES		MF	
Standing	10	NB	2.26	(3.54)	9.75	(7.98)	3.81	(2.62)	19.17	(7.76)	18.26	(6.94)
		PB	2.43	(3.93)	9.25	(10.58)	3.10	(1.82)	18.21	(7.71)	17.31	(8.43)
		LSB	3.16	(4.86)	9.05	(8.60)	3.15	(1.83)	19.20	(7.69)	17.56	(6.86)
	20	NB	3.64	(5.86)	11.90	(7.68)	8.83	(6.89)	30.24	(10.75)**	28.10	(7.84)
		PB	3.28	(5.36)	9.67	(8.15)	8.53	(6.67)	27.14	(10.79)**	27.95	(13.09)
		LSB	3.49	(5.27)	13.22	(14.17)	5.45	(3.36)	30.79	(11.76)	29.46	(10.83)
Squatting	10	NB	2.36	(4.22)	12.22	(8.93)	6.13	(3.74)	39.32	(12.96)	34.71	(10.24)
		PB	1.72	(2.58)	10.19	(8.30)	5.08	(2.82)	35.50	(11.27)	34.42	(10.38)
		LSB	2.46	(3.61)	9.76	(9.65)	6.69	(4.06)	36.45	(13.37)	39.19	(13.72)
	20	NB	2.79	(4.43)	22.97	(19.29)**	8.75	(4.91)	52.74	(16.19)	46.64	(12.09)
		PB	2.15	(3.18)	17.31	(16.74)**	9.21	(7.84)	48.92	(12.78)	49.08	(10.47)
		LSB	2.80	(4.58)	22.47	(23.05)	12.50	(9.60)	54.01	(15.51)	54.11	(15.22)

The numbers shown are the average and standard deviation (in the parentheses) values of 5 trunk muscles (RA: Rectus Abdominis, IO: Internal Oblique, EO: External Oblique, ES: Erector Spinae, MF: Multifidus). *p < 0.05, **p < 0.01, ***p < 0.001.

Table 2. Latency onset of TrA/IO muscle relative to the lift-off time (n = 9).

Lifting posture	Load (kg)	Latency onset of TrA/IO (m/sec)					
		No belt		Pelvic belt		Lumbo-sacral belt	
Standing	10	296.68	(159.30)	169.86	(100.49)	159.89	(71.38)
	20	337.91	(72.57)	255.56	(68.69)	216.92	(67.12)
Squatting	10	215.79	(35.38)	154.63	(33.16)	145.16	(64.34)
	20	226.24	(84.91)	183.80	(55.20)	166.55	(60.42)

The numbers shown are the average and standard deviation (in the parentheses) values of the onset time relative to the lift-off time.

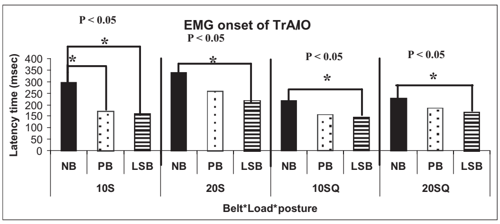


Figure 5. Comparison of latency of onset time of TrA/IO muscles among the 3 belt conditions (NB: No Belt, PB: Pelvic Belt, LSB: Lumbo-Sacral Belt), and 2 load – 2 lifting posture conditions (10S: 10 kg in standing, 20S: 20 kg in standing, 10SQ: 10 kg in squatting, 20SQ: 20 kg in squatting) (n = 9).

4 DISCUSSION

This study examined the immediate effect of 2 back belt types on trunk muscle activation, especially TrA muscle, among inexperienced back belt users during manual lifting. The result showed that only the pelvic belt significantly decreased the TrA/IO muscle activitiy in the 20 kg squat lifting task. The squat lifting task from the floor to knuckle height (with 20 kg load) in this study required more trunk stabilization effort than the other 3 lifting tasks because of the highest load and trunk moment required in changing the position from squatting to standing. Obviously, it was found that lifting tasks with lower load level or lesser trunk movement in the standing posture seemed to have decreased TrA muscle activity. This result indicated that the pelvic belt was effective in stabilizing the lumbar spine during lifting as seen from the decreased TrA muscle activity.

Lifting in the standing posture in this study was used to control confounding factors of back belt usage on trunk moment. This is because changes in the trunk muscle EMG due to back belt usage, at least part of them, must be attributed to the kinematic changes with back belt application. The EMG is highly correlated and sensitive to changes in trunk moment (Lee 2002). The lumbo-sacral belt is wider than the pelvic belt (20 cm and 12 cm, respectively). Thus, it could limit the trunk moment during squat lifting more than the pelvic belt. However, the result did not show any significant difference. Therefore, a decrease in TrA activity in the squat lifting posture in this study was influenced by the pelvic belt since the result was similar to that from lifting in the standing posture (Fig. 3).

When wearing the back belt, no changes in trunk muscle EMG were found except the ES in which its activity was significantly decreased when wearing the pelvic belt during the lifting of 20 kg load in the standing posture (Fig. 4). This result also indicated the effective of pelvic belt in reducing back muscle activity during lifting.

The pelvic belt is considerably narrower than the lumbo-sacral belt, but it focuses on the point that directly stabilizes the lower spine which is the most movement part. The lumbo-sacral belt, on the other hand, is wider and generally stabilizes the whole lumbar spine and pelvis. Thus, unlike the pelvic belt, it does not have any specific focus point on any part of spine. When considering the belt width, the results of this study support that both TrA and ES muscle activities were decreased during lifting when wearing the pelvic belt.

This study was the first study to determine the onset EMG of TrA during back belt usage. The results show that the delay of TrA onset is mostly seen when wearing the lumbo-sacral belt (Fig. 5). This result may be induced by the belt pressure that covers all lumbar spine and sacrum during lifting; whereas, the pelvic belt only covers the lower part of the lumbar spine and pelvis.

To determine the onset of TrA, the participant needs to relax all trunk muscles during pre-lifting period. Each participant was instructed to breathe naturally and relax all muscles before starting the lifting task. Therefore, this study could not control the breathing pattern during lifting.

However, this study was a preliminary study to report the effect of back belt usage on TrA in a small group of inexperienced back belt users. Future studies should consider a larger number of participants. A long-term effect of back belt usage among experienced users should also be investigated.

In addition, the decreased EMG activity of TrA during lifting with back belt found in this study could possibly cause the related muscle to become weaker in the future.

5 CONCLUSION

The pelvic belt resulted in a significant decrease in TrA and ES muscle activities in squat lifting and standing postures, respectively, with the 20 kg load. As for the onset time of TrA, the lumbosacral belt had a significant delay onset in all lifting task condition. The pelvic belt only showed a significant delay onset in the standing posture with 10 kg load. The long-term effect on experienced back belt users still needs further investigation.

ACKNOWLEDGMENTS

The authors would like to show much appreciation, impression, and gratitude to the members of the Department of Ergonomics, University of Occupational and Environmental Health, Japan, for their contribution to this study. Especially, special thanks to Ms. Chalermsiri Theppitak, Ms. Praphatson Klangsin and Ms. Mehrnoosh Movahed for their helps throughout the data collection period. The authors also thank Asst. Prof. Jemma Coleman for giving helpful suggestions.

REFERENCES

Arokoski, J.P., Valta, T., Airaksinen, O. & Kankaanpa, ¨a¨. M. 2001. Back and abdominal muscle function during stabilization exercises. *Arch Phys Med Rehabil* 82: 1089–1098.

Arokoski, J.P., Valta, T., Kankaanpa, M. & Airaksinen, O. 2004. Activation of lumbar paraspinal and abdominal muscles during therapeutic exercises in chronic low back pain patients. *Arch Phys Med Rehabil* 85: 823–832.

Cholewicki, J., Juluru, K. & McGill, S.M. 1999. Intra-abdominal pressure mechanism for stabilizing the lumbar spine. *J Biomech* 32(1): 13–17.

Granata, K.P., Marras, W.S. & Davis, K.G. 1997. Biomechanical assessment of lifting dynamics, muscle activity and spinal loads while using three different styles of lifting belt. *Clin Biomech* 12(2): 107–115.

Hodges, P.W. & Richardson, C.A. 1997. Feedforward contraction of transversus abdominis is not influenced by the direction of arm movement. *Exp Brain Res* 114: 362–370.

Lee, Y.H. & Kang, S.M. 2002. Effect of belt pressure and breath held on trunk electromyography. *Spine* 27(3): 282–290.

Marshall, P. & Murphy, B. 2003. The validity and reliability of surface EMG to assess the neuromuscular response of the abdominal muscles to rapid limb movement. *Journal of Electromyography and Kinesiology* 13: 477–489.

McGill, S. 2002. The question of back belts. In: *Low back disorders: Evidence based prevention and rehabilitation.* USA: Human Kinetics: 193–201.

McGill, S. 2002. Injury prevention. In: *Low back disorders: Evidence-based prevention and rehabilitation*. Human Kinetics: 147–158.

McGill, S.M. 2005. *On the use weight belts*. A review invited by the National Strength and Conditioning Association.

Miyamoto, K., Iinuma, N., Maeda, M., Wada, E. & Shimizu, K. 1999. Effects of abdominal belts on intra-abdominal pressure, intramuscular pressure in the erector spinae muscles and myoelectrical activities of trunk muscles. *Clin Biomech* 14: 79–87.

Richardson, C., Hodges, P. & Hides, J. 2004. *Therapeutic exercise for lumbopelvic stabilization: A motor control approach for the treatment and prevention of low back pain*, 2nd ed. Edinburgh: Churchill Livingstone.

Sullivan, M.S. 1989. Back Support Mechanisms during Manual Lifting. *Phys Ther* 69(1): 38–45.

van Poppel, M.N.M, de Looze, M.P, Koes, B.W., Smid, T. & Bouter, L.M. 2002. Mechanisms of action of lumbar supports: a systematic review. *Spine* 25(16): 2103–2113.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Effects of holding umbrellas and wind blown on forearm muscle activities and grip pressure

Bor-Shong Liu, Wen-Chang Chou, Rui-Zhe Wu & Chien-Sung Wang
Department of Industrial Engineering and Management, St. John's University, Taipei, Taiwan

ABSTRACT: The purpose of this study was to examine the effects of holding umbrella and wind blown on forearm muscle activities and grip pressure. Ten various umbrellas have been evaluated. The angles of wrist and elbow joint and EMG of extensor carpi radialis and biceps were recorded by Biopac system. In addition, the tactile grip pressure also measured by Grip System. Each participant performed thirty bouts at one of ten types of umbrellas and three holding conditions (upraised without wind blown, upraised with wind blown, and raised forward with wind blown). Results of ANOVA showed that the angles of wrist and elbow were significant differences between umbrellas. In addition, EMG activities on both muscles are higher while raised forward with wind blown. The mean grip pressures were higher while holding smaller diameter of umbrellas handle. Of course, the grip pressure increased with wind blown. Results of present study could provide the information for designing of umbrellas.

Keywords: umbrella, holding postures, hand tools

1 INTRODUCTION

An umbrella is a canopy designed to protect against rain or sunlight. Although very common in our daily lives, the basic design of umbrellas has not changed significantly over the past years. Of cause, umbrella is one kind of hand tools. In the past research, hand tool design was focused on tool function in order to improve task efficiency. Furthermore, design of hand tools was to pay special attention to the ergonomical design of tools and try to adapt the individual requirements for each user, then it would be possible to make the user both more effect and prevent many injuries through overloading (Berglund and Eriksson, 1995). Some studies reported that well-designed handle of hand tools can facilitate improved performance and reduce user musculoskeletal injuries (Chi and Drury, 1988; Okunribibo and Haslegrave, 1999). There is a need to study how the types of umbrella affect user's responses of forearm muscles. The purpose of this study was to examine the effects of holding umbrella and wind blown on forearm muscle activities and grip pressure. Results of present study could provide the information for designing of umbrellas.

2 METHODS

2.1 *Participants*

Twelve undergraduate students at the St. John's University were recruited in present study. Their mean age, mean stature and mean body mass were 21.7 years old, 165.6 cm and 64 kg, respectively. All subjects were healthy and reported no musculoskeletal problems that might influence performance detrimentally. All subjects claimed to 'dominantly right-handed' and used the right hand during the study.

2.2 *Apparatus and materials*

Total of ten umbrellas have been evaluated in present study (Figure 1). Specifications of ten umbrellas were showed in Table 1. The tube length was from 50 cm to 78.4 cm and weight of umbrella was 164 g to 502 g. Area of umbrella was 5024 cm^2 to 14733 cm^2. In addition, handle length was from 1.5 cm to 19 cm.

Figure 1. Ten umbrellas have been evaluated in present study.

Table 1. Specifications of ten umbrellas.

No	Tube length (cm)	Radius (cm)	Area (cm^2)	Handle circle (cm)	Handle length (cm)	Weight (g)
1	78.4	68.5	14733	7.1	6.5	502
2	66.5	59.5	11116	7.5	6.5	502
3	68	58.5	10745	6	16	336
4	64	53.5	8987	10	19	414
5	50	45	6358	10	1.5	224
6	57	48	7234	12.5	2.5	280
7	53	47	6936	11.5	3	164
8	50	48	7234	8	3	168
9	52	48	7234	20.5	8	302
10	53	40	5024	7.5	6.5	336

2.3 *Experimental procedures*

Upon arrival, the participants were given general instructions about the experiment. Subjects were trained and known experimental task until they were able to make smooth manipulation. Further, the miniature joint angle sensor (S720, Measurand Inc., Canada) and shape sensor (S700) were mounted on the neutral axis of wrist and elbow respectively. Thus, angles of wrist and elbow were measured (Figure 2). In addition, the EMG activity was recorded over the right side of Extensor Carpi Radialis and biceps by AcqKnowledge on Biopac systems (MP150, Biopac Inc., USA). The skin area of current interest was shaved and cleaned with 70% alcohol solution. surface electrodes were attached over the muscle of concern, parallel to the muscle fibers, with an inter-electrode center distance of 30 mm, impedance <20 kΩ. Reference electrodes were placed over bone. The signals were amplified, band-pass filtered 20–500 Hz, A/D converted and sampled at 1000 Hz.

Reference EMG values for normalization (RVC) were obtained during Maximum isometric Voluntary Contraction (MVC) of the specific muscle group for two trials of about 5 sec. In addition, the tactile grip pressure also measured by Grip System (Figure 3). A total of 30 trials with one of ten types umbrellas and three holding condition involving upraised without wind blown (UP), upraised with wind blown (UW) and raised forward with wind blown (FW). The order of these trials was randomly assigned for each subject.

2.4 *Data analysis*

A randomized complete block design (blocks as individual subjects) with two within-subject factors (types of umbrella and holding postures) was used for this study. Twelve participants performed three holding postures with ten types of umbrellas. The dependent variables were angles of wrist and elbow, RVC (%) for extensor carpi radialis and biceps. Further, analysis of variance (ANOVA) was utilized to identify significant differences between conditions for dependent variables. Statistical significance was set at a probability level of 0.05.

Figure 2. The angles of wrist and elbow were measured.

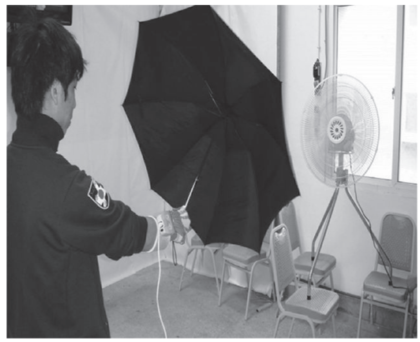


Figure 3. Grip pressure measured by Grip System.

3 RESULTS

3.1 *Types of umbrella*

Results of ANOVA revealed that these were significant difference in RVC of extensor carpi radialis between types of umbrella ($p < 0.05$). The mean RVC of extensor carpi radialis was higher while holding No. 1 and No. 3 of umbrella (Figure 4). In addition, the mean RVC of biceps was also higher in holding No. 1 of umbrella.

For analysis of grip pressure, these were significant difference between types of umbrella ($p < 0.001$). The mean grip pressure was higher while holding No. 1, No. 3 and No. 10 of umbrella. These are 55.6, 47.3 and 48.7 kgf, respectively.

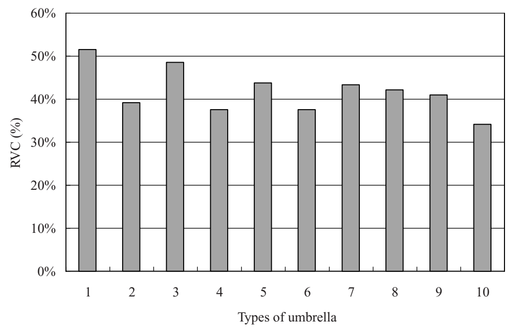


Figure 4. Mean RVC of extensor carpi radialis between umbrellas.

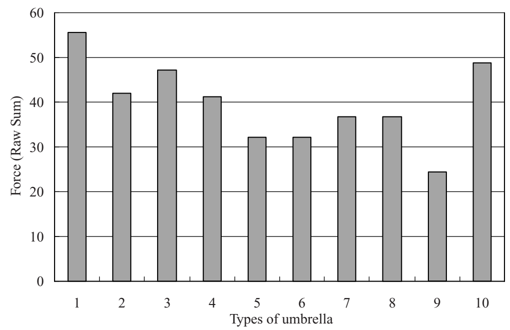


Figure 5. Mean grip pressure between umbrellas.

3.2 *Effects of wind blown*

The wind blown affects strongly on muscle activities of forearm ($p < 0.05$). Figure 6 showed the mean RVC of extensor carpi radialis and biceps between holding postures. EMG activities were higher in extensor carpi radialis while subject holding umbrella on upraise with wind blown, particularly (Figure 6).

Figure 7 revealed the distribution of grip pressure while holding upraised with wind blown (left) and raised forward with wind blown (right). The contact areas of hand were significant difference. The main contact areas of hand were on thenar muscles and midpalmar muscles while subject holding umbrella on upraised with wind blown. By contrast, the main contact areas of hand were on hypothenar muscles and tip muscles of ring finger. In addition, Figure 8 showed the comparison of grip pressure distribution by holding upraised with wind blown (green line) and raised forward with wind blown (red line). The grip pressure maintained on 30 kgf while holding upraised with wind blown. By contrast, the distribution of grip pressure was unstable form about 10 to 40 kgf.

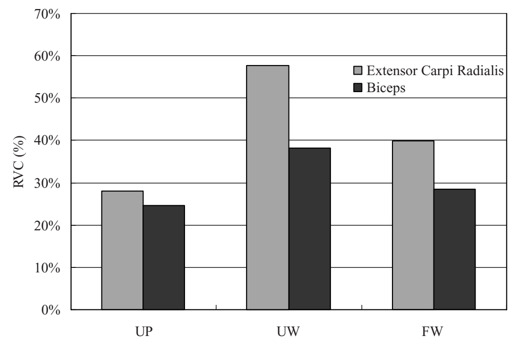


Figure 6. Mean RVC of extensor carpi radialis between umbrellas.

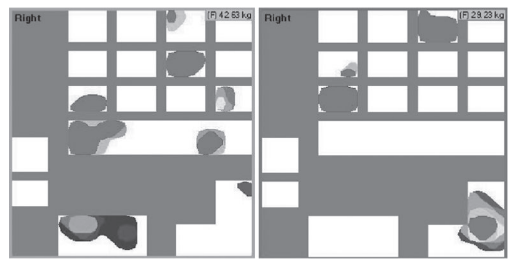


Figure 7. The distribution of grip pressure while holding upraise with wind blown (left) and raised forward with wind blown (right).

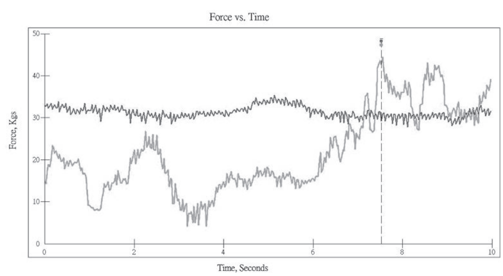


Figure 8. Comparison of grip pressure distribution by holding upraised with wind blown (green line) and raised forward with wind blown (red line).

3.3 *Interaction between umbrellas and wind blown*

The interesting result from the viewpoint of RVC of extensor carpi radialis was the significant interaction between types of umbrella and wind blown ($p < 0.05$). Surely, the EMG activities of forearm were lower while holding umbrellas without wind blown (UP line). In average, while holding umbrella on raised forward posture with wind blown was also lower levels of EMG activities on forearm muscles. For worse cases of higher EMG activities, subjects were holding the No. 3, No. 4 and No. 7 of umbrella on upraised posture with wind blown. On the other hand, No. 6 umbrella had larger contact area of hand handle. Thus, EMG activities of forearm muscles and grip pressure were lower level.

4 DISCUSSION AND CONCLUSION

Hand tools are still the primary interface for operators at daily life, in spite of all the automation efforts made by modern world. In the past, hand tool design was focused on tool function in order to improve task efficiency and allow for standardization. The tool should perform the task which it was designed for and respond to the needs of the greatest possible number of users (Aptel & Claudon, 2002; Marsot & Claudon, 2004). Among the various tool handle design characteristics, handle diameter has been an important factor to maximize grip strength, minimize on the digit flexor tendons, first metacarpal ulnar collateral and carpometacarpal ligaments (Meagher, 1987) and it can also influence force exertion in manual work (Grant et al., 1992; Blackwell et al., 1999). The evaluation of subjective handle comfort ratings indicated that participants preferred midsized handles over the smallest or largest handles which were associated with highest or lowest total finger forces, respectively. Results of present study also showed that handle characteristics (shape, length) were significant effects on EMG activities of forearm and distribution of grip force. Kong & Lowe (2005) tested maximum grip force on cylindrical aluminum handles maximum grip force on cylindrical aluminum handles to evaluate the relationships between handle diameter (25–50 mm diameter handles), perceived comfort, finger and phalange force distribution, and electromyographic efficiency of finger flexor and extensor muscle activity. Their results also reported that there was a significant negative correlation between the total finger force and handle diameter. Participants exhibited greater total finger force with the same amount of EMG activity with the small diameter handles than with the large diameter handles. For effects of wind blown, subjects need more effort in forearm and grip force while holding umbrellas in upraise with wind blown.

Participants preferred circular and hexagonal longitudinal-shaped and double frustum and cone lateral-shaped handles over the triangular longitudinal-shaped handles, and cylindrical and reversed double frustum lateral-shaped handles (Kong et al., 2008). Chang et al. (1999) evaluated the effects of handle types of garden tools on ergonomic effectiveness, and subjective judgment of tactile feel and control.

Their results showed that the grip force and EMG analyses were less effort with hollow fiberglass handle.

The optimal surface material and cross-section of power tool handles are not sufficiently described in ergonomic literature (Bjoring et al., 1999). Results of previous study showed that foam rubber on the handle is a more preferable covering material compared to harder rubber and it does not increase the muscular activity. Thus, material and shapes of umbrella handle should be examined in further research.

ACKNOWLEDGEMENTS

This study was supported by a grant from the National Science Council of Taiwan (98-2815-C-129-003-E). Authors would like to thank all participants during collecting the data.

REFERENCES

Aptel, M. & Claudon, L. 2002. Integration of ergonomics into hand tool design: principle and presentation of an example. International *Journal of Occupational Safety and Ergonomics* 8(1): 107–115.

Bjoring, G., Johansson, L. & Hagg, G. 1999. Choice of handle characteristics for pistol grip power tools. *International Journal of Industrial Ergonomics* 24: 647–656.

Blackwell, J.R., Kornatz, K.W. & Heath, E.M. 1999. Effect of grip span on maximal grip force and fatigue of flexor digitorum superficialis. *Applied Ergonomics* 30: 401–405.

Chang, S.R. Park, S. & Freivalds, A. 1999. Ergonomic evaluation of the effects of handle types on garden tools. *International Journal of Industrial Ergonomics* 24: 99–105.

Chi, C.F. & Drury, C.G. 1988. Cross validation of measures of handle/human fit Applied Ergonomics 19(4): 309–314.

Grant, K.A., Habes, D.J. & Steward, L.L. 1992. An analysis of handle designs for reducing manual effort: the influence of grip diameter. *International Journal of Industrial Ergonomics* 10: 199–206.

Kong, Y.K. & Lowe, B.D. 2005. Optimal cylindrical handle diameter for grip force tasks. *International Journal of Industrial Ergonomics* 35: 495–507.

Kong, Y.K., Lowe, B.D., Lee, S.J. & Krieg, E.F. 2008. Evaluation of handle shapes for screwdriving. *Applied Ergonomics* 39: 191–198.

Marsot, J. & Claudon, L. 2004. Design and ergonomics. Methods for integrating ergonomics at hand tool design stage *International Journal of Occupational Safety and Ergonomics* 10(1): 13–23.

Meagher, S.W. 1987. Tool design for prevention of hand and wrist injuries *Journal of Hand Surgery* 12(A): 855–857.

Okunribibo, O.O. & Haslegrave, C.M. 1999. Effect of handle design for cylinder trolleys *Applied Ergonomics* 30(4): 407–419.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Differences in lifting postures at various exertion heights between workers and novices

Yi-Lang Chen & Chia-Jung Chen
Department of Industrial Engineering and Management, Mingchi University of Technology, Taipei, Taiwan

Yu-Chi Lee
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

ABSTRACT: The purpose of this study was to compare participant's lifting postures at various exertion heights between experienced workers and novices. Forty-two experienced workers and novices (21 of each) were recruited and required to determine their static-lifting strengths under various height levels (10~150 cm in increments of 10 cm) using two exertion methods (vertically upward lifting, VUL, and toward body lifting, TBL). Results showed that the VUL forces were much higher than the TBL at 15 height positions ($p < 0.001$). Forces in all 30 task combinations showed no difference between the two groups except for the VUL at heights of 100~120 cm. Novices' VUL forces were on the average 4.6 ~7.6 kg lower than workers', while exerting forces at 100–120 cm (all $p < 0.05$). Workers in VUL postures were highly differentiated from novices while performing near-floor positions, but the forces were equivalent to each other. That is, workers tended to adopt a safer (i.e., more flexed knees) and more skillful technique to generate forces than did novices. The results suggested that lifting strength data generally collected on students should be carefully used in the task design.

Keywords: lifting postures, experienced worker, novice, exertion heights

1 INTRODUCTION

Even though automation technology is extensively used today, Manual Materials-Handling (MMH) still exists in many different task situations, especially in newly developed service trades such as hypermarkets, shopping malls, and wholesale stores. The concerns for these handling tasks by ergonomists focus on whether human capabilities match the job demands. Human strengths are a primary measure of an individual's physical capabilities, particularly those that permit a person to exert force or sustain external loading without inflicting personal injury (Mital et al. 1993).

Many studies have established the human strength data under various task variables, which were usually derived from student participants (Yates et al. 1980, Lee 2004). These studies conducted on student participants are on the basis of: ease in recruitment, greater reliability, lower cost, and availability for a long-term study (Mital 1987). However, it is essential to identify the suitability of an individual's capabilities before directly applying the results, conducted on students, to the entire industrial population.

Much of the literature has reported that the lifting techniques of highly skilled workers substantially differ from those of novices (Authier et al. 1996, Gagnon et al. 1996). Field studies have also shown that workers are using methods other than the recommended ones (Kuorinka et al. 1994, Baril-Gingras and Lortie 1995), and lack a consensus on the best methods (Authier et al. 1996). For a given task, in general, the experienced workers employ lower biomechanical spinal loads (Marras et al. 2006), less back-muscle activities (Keir and MacDonell 2004), more psychophysically accepted lifting-weights (Mital 1987), and higher subjective discomfort thresholds (Parakkat et al. 2007) than the novices. On the contrary, some investigations suggest that the response patterns of industrial and non-industrial workers to task variables in manual-lifting activities are similar (Mital and Manivasagan 1983). The controversies may attribute to the diverse experimental settings among these studies.

To our knowledge, no comparative studies have been found in the literature that systemically assesses the simulated lifting strengths of workers and novices under full-range heights. Since the large-scale warehouse stores (e.g., the hypermarket) with high employee turnover rates almost always recruit novice workers or part-timers every week in Taiwan (Chen 2008), it is becoming more and more important to develop training programs and more suitable shelf designs for handling tasks, by understanding the differences in lifting strengths between workers and novices. This study, therefore, consists of exertion heights and exertion methods to differentiate the lifting strengths of experienced workers from novices. Furthermore, the participant's testing postures in all task combinations were also recorded for interpreting the collected strength data.

2 METHODS

2.1 *Participants*

Forty-two experienced male worker and novice participants (21 of each), with no prior history of any musculoskeletal disorder, volunteered for this study and received an hourly wage for finishing all test conditions. Twenty-one workers, with at least two-years experience levels, were chosen as experienced participants. They were selected from a large hypermarket, and their duties consisted mainly in varied replenishing tasks onto the shelves. Another 21 novice participants (with no manual material handling experience) were recruited from university students. The mean (SD) age was 23.1 (3.0) years for the experienced group, and 23.3 (2.4) years for the novice group. Their anthropometric data and isometric strengths, employed with the methods of Chaffin (1975), were measured. There were no significant differences by t-test between the two groups with respect to their physique and strength measurements (all $p > 0.05$).

2.2 *Experimental apparatus*

The static-lifting strengths were measured using a Static-Lifting Strength Tester (SLST), as illustrated in Figure 1. The SLST, with an incremental height setting for measuring strengths, consisted of a standing platform, a steel frame with 20 positioning holes (ranging from 10 cm to 200 cm in increments of 5 cm), a sliding height stopper along the frame, and a 55 cm handle bar (diameter: 3.5 cm) attached to the stopper. The force applied onto the handle bar was measured by a load cell that connected with the bar and was rigidly placed on the stopper, and then the strength signal (60 Hz) was transferred into an A-D converter and a digital readout unit (JSES Model 32628, U.S.A.). The A-D converter was calibrated eight times prior to the testing against known static loads. The accuracy of the measurement was within ±1% of the lifting strength.

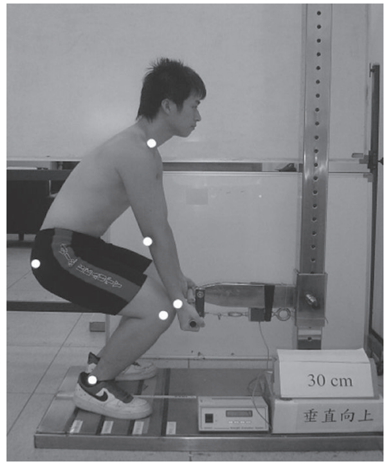


Figure 1. The device of static-lifting strength and schematic testing posture in the study.

2.3 *Body posture recording*

The testing postures of randomly sampled participants (one third of participants) were videotaped in this study. Six adhesive reflective markers were attached to the right side of the participant's wrist, elbow, shoulder, hip, knee, and ankle joints, as is also shown in Figure 1. In the study, the joint angles were all defined as the inter-joint angles between the two adjacent segments (Figure 2). The digital video camera, which was set up at a distance of 5 m from participants, was used to record the positions of each joint marker when participants performed the strength test. The camera heights were always adjusted to orthogonally align the participant's hip positions to eliminate as much as possible the distortion errors resulting from the varied testing heights. Image processing software was then used to calculate the relevant joint angles.

2.4 *Experimental design*

This study examined the differences in static-lifting strengths between experienced workers and novices under 15 lifting heights as well as under different exertion methods. Fifteen lifting heights, ranging from 10 to 150 cm and spaced 10 cm apart, were set by the SLST. There were two exertion methods performed by each participant at all 15 height levels. One exertion method was Vertically Upward Lifting (VUL) and the other was Toward Body Lifting (TBL). The VUL has been generally accepted as the standard method in previous strength measurement studies, whereas

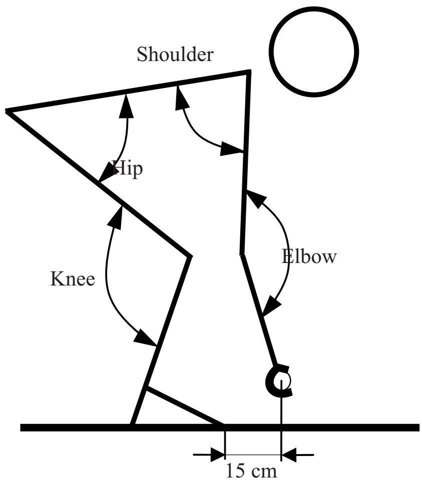


Figure 2. The definition of the body joint angles while testing in this study.

the TBL refers to the finding of a study by Garg et al. (1983). Garg et al. observed that the subjects pulled the load toward the body when lifting psychophysically determined maximum weights. Lifting at an angle resulted in a decrease in moment at the upper arms and spinal L5/S1 joints and an increase in moment at the knees and ankles. In this study, we would like to identify the strength differences caused by different exertion directions. All participants were requested to perform the maximum lifting strengths for each task combination for at least three repetitions. Each strength measurement was repeated until three readings were obtained, which were consistent within a range of 10%. As a result, a total of 3,780 strength data (42 participants × 2 exertion methods × 15 heights × 3 repetitions) was determined.

2.5 *First line of text or heading*

All participants were familiarized with the experimental procedures, and stretched themselves at least 10 min before the data was collected. During the experiment, the participants wore light clothing and rubber gym shoes. They were randomly requested to perform all maximum strengths at a specific exertion height and direction with free style but symmetrical lifting postures. The strength-testing procedure was performed according to the methodology used by Chaffin (1975) and Ayoub et al. (1978). A period of 5 s was chosen as sufficient time to build up and maintain a constant force, while being brief enough to be endured without perceiving any muscular fatigue. A minimum rest period of 2 min was required between successive trials. The horizontal-projection distance between the tip of the shoes of the participants and the handle bar was set at 15 cm. As was observed in a recent field study, this distance took the most advantageous testing posture to apply force by participants for full-range heights. Moreover, the distance was maintained at about half the width of the box that was used extensively in the Taiwanese hypermarkets (Chen 2008).

While determining the lifting strength, we used a nested design for analysis. Each participant was considered a block. We analyzed the strength data using a two-way analysis of variance (ANOVA) and used Duncan's Multiple-range test (Duncan's MRT) for post-hoc comparisons. Differences in participants' strengths of VUL and TBL were checked by a t-test. An alpha level of 0.05 was selected as the minimum level of significance.

3 RESULTS

3.1 *Difference in strengths between the groups*

Table 1 shows the summary ANOVA result of lifting strengths in different participant groups and independent task variables. As shown in the table, the exertion height did influence all of the TBL forces ($p < 0.001$), as well as the novices' VUL forces ($p < 0.001$). It was noteworthy that the worker's VUL forces were not affected by heights. When averaged across the exertion heights, Duncan's MRT result also showed that the VUL forces among workers were much higher than the TBL (31.67 kg: 11.67 kg); as were the forces among novices (30.45 kg: 10.33 kg).

Different strength profiles under various exertion heights are illustrated in Figure 3. As aforementioned, the VUL forces were always higher than the TBL, regardless of working experience. The result of the t-test revealed that there were no differences between workers' and novices' TBL forces at all 15 height levels. That is, the worker and the novice participants demonstrated a similar strength profile while performing the TBL. However, the VUL force profiles exhibited an inconsistent trend for the two groups. The t-tests were performed and significant differences in VUL strengths were found

Table 1. Significance of effects of independent variables to the lifting strengths by ANOVA.

	Workers		Novices	
	VUL	TBL	VUL	TBL
Exertion heights	N.S.	$p < 0.001$	$p < 0.001$	$p < 0.001$

N.S.: non-significant.

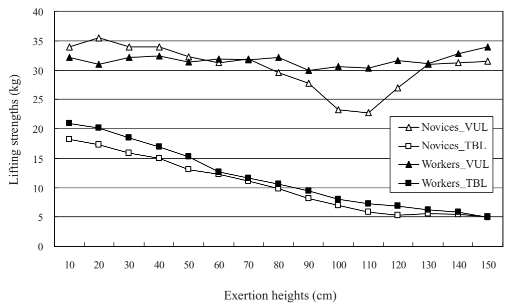


Figure 3. Lifting strength profiles under various exertion heights and methods.

between workers and novices at heights of 100, 110, and 120 cm. Higher strengths were determined by workers and with differences ranging from 4.57 kg to 7.61 kg (all $p < 0.05$).

3.2 *Posture analyses*

Comparisons of lifting postures while exerting maximum strength indicated that there were higher knee, shoulder, and elbow angles at lower height levels (e.g., 10~30 cm) in the novices' VUL tests, as shown in Figure 4. This means that a more stooped technique was chosen by novices while lifting upwardly from near-floor levels, as opposed to a more flexed knee strategy by workers. From the figure, it can also be seen that workers employed relatively consistent patterns of joint angles than did the novices, for whichever VUL or TBL that was performed. Figure 5 further shows the postures of the VUL at a height of 10 cm by two representative experienced and novice participants, respectively. In contrast to the TBL, quite different postures were adopted between the two groups for the VUL, and nevertheless the strengths remained unchanged (Figure 3).

4 DISCUSSION

The results of this study provide an indication that there exists a discrepancy in vertically-upward lifting strengths at the heights from 100 cm to 120 cm between the two participant groups. Besides that, similar trends in strengths, along with increased heights, were exhibited by different exertion methods, regardless of working experience. This implied that the lifting-strength data collected on novices (as those students in our study) should be still cautioned to the task design, especially when the task was to be conceivably done by novice workers.

The novice's VUL strength profile was in agreement with the result of Lee's study (2004). Lee examined the lifting strengths of twelve university student participants at 10, 45, 75, 105 and 140 cm, and found the lowest strength to be at the exertion height of 105 cm. However, he did not further explain why the strength at 140 cm was higher than that at 105 cm. As illustrated by the VUL postures in Figure 4, the shoulder was placed in a minimum extension, whereas the hip reached its maximum extension at about a 110 cm height of exertion. This can partially explain why at these height levels the lower strengths were found, since the novices may apply force only by their upper arms' own efforts. It is well-recognized that the arm and shoulder strengths are dominant while lifting from higher positions, and the leg and back muscles are stronger than the upper extremities. We can therefore attribute this to a relatively disadvantageous position in the upper extremities at about the participants' elbow height (data of 109.0 cm). When the height exceeded that, the shoulders of novices began to extend, then gradually changed to a more beneficial posture for generating forces.

Posture analysis revealed an interesting result in the study. Working experience influenced the adopted posture while exerting, especially in near-floor VUL strengths. Figure 5 illustrates the comparison of the testing postures for two exertion methods between worker and novice participants at a height of 10 cm. For the VUL test, workers tended to choose a more knee-flexion strategy, whereas the novices tended to choose a more stooped one. It is generally accepted that a stooped lifting posture, which can provide higher lifting strengths, would impose more compressive forces on the spinal L5/S1 disc. From the figure, we also can observe a larger horizontal distance (i.e., moment arm) from the handle to the participants' lower back in novices than that of in workers. This would lead the workers to be safer for testing than would the novices. The result did not match with Gagnon et al's study (1996). They found that the expert handler reduced knee flexion more than the novice did, to save energy more advantageously. However, only the static-lifting strengths examined in our study were not similar to Gagnon et al's dynamic lifting act. While testing, no instruction or hint about testing posture was given to participants in this study. Any difference in postures between the two groups in this study therefore can be reasonably attributed to the influence of working experience. As shown in Figure 4, the workers revealed a more consistent body posture for both VUL and TBL tests. Contrarily, the novices chose different postures depending on what exertion method or height was requested. Even so, note that there was no difference in strengths between workers and novices while performing VUL at lower

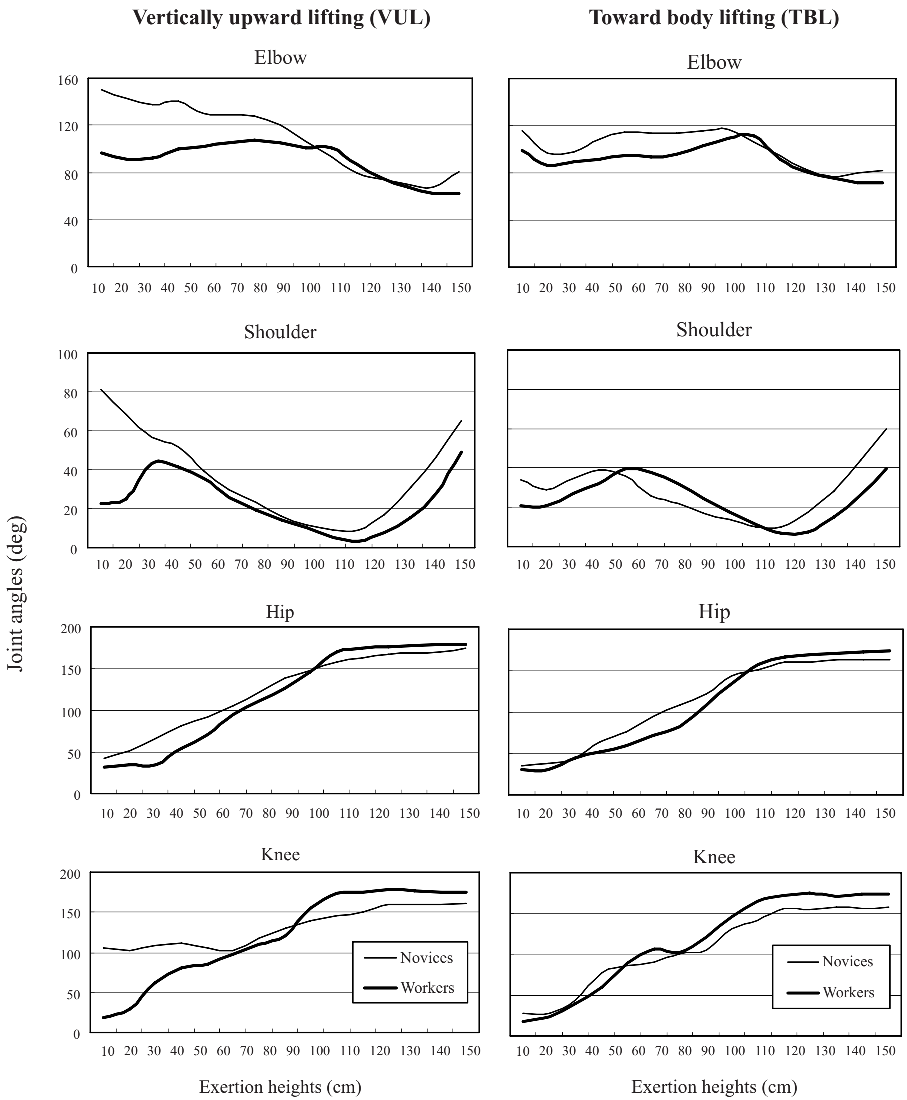


Figure 4. Comparison of the testing postures while participants performed the VUL and TBL.

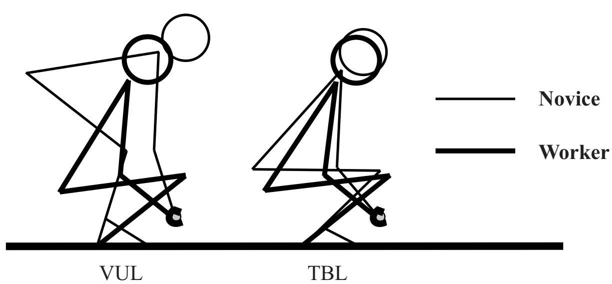


Figure 5. Comparison of the schematic postures between the experienced workers and novices while performing static lifting at a height of 10 cm.

height levels. The workers, experienced in realistic lifting tasks, would choose a safer and more skillful technique than the novices did to generate lifting strengths.

5 CONCLUSION

This study shows that the static-lifting strengths of novices were significantly lower than those of experienced workers while upward lifting near the participant's elbow height. A safer technique was chosen by workers at near-floor upward lifting; even so, the corresponding strengths did not decrease in comparison to the novice's strengths. Strengths that were exerted toward the body were exceptionally lower than the vertically lift-up for both groups. The results of this study provide the inference that experienced workers may have learned a protective strategy and a more effective technique to satisfy their daily job demand. Furthermore, the results also indicated that the lifting strength data collected on novices (e.g., student) should be carefully used for whichever task variable is to be taken into consideration.

REFERENCES

Authier, M., Lortie, M. & Gagnon, M. 1996. Manual handling techniques: Comparing novices and experts. *International Journal of Industrial Ergonomics* 17: 419–429.

Ayoub, M.M., Bethea, N.J., Deivanayagam, S., Asfour, S.S., Bakken, G.M., Liles, D., Mital, A. & Sherif, M. 1978. *Determination and modeling of lifting capacity.* Final Report, HEW [NIOSH], Cincinnati, OH.

Baril-Gingras, G. & Lortie, M. 1995. The handling of objects other than boxes: univariate analysis of handling techniques in a large transport company. *Ergonomics* 38: 905–925.

Chaffin, D.B. 1975. Ergonomics guide for the assessment of human static strength. *American Industrial Hygiene Association Journal* 36: 505–511.

Chen, C.J. 2008. *Investigation of musculoskeletal disorders for the warehouse workers in Taiwan*. Master Dissertation, Mingchi University of Technology, Taipei, Taiwan.

Gagnon, M., Plamondon, A., Gravel, D. & Lortie, M. 1996. Knee movement strategies differentiate expert from novice workers in asymmetrical manual materials handling. *Journal of Biomechanics* 29: 1445–1453.

Garg, A., Sharma, D., Chaffin, D.B. & Schmidler, J.M. 1983. Biomechanical stresses as related to motion trajectory of lifting. *Human Factors* 25: 527–539.

Keir, P.J. & MacDonell, C.W. 2004. Muscle activity during patient transfers: A preliminary study on the influence of lift assists and experience. *Ergonomics* 47: 296–306.

Kuorinka, I., Lortie, M. & Gautreau, M. 1994. Manual handling in warehouses: the illusion of correct working postures. *Ergonomics* 37: 655–661.

Lee, T.H. 2004. Static lifting strengths at different exertion heights. *International Journal of Industrial Ergonomics* 34: 263–269.

Marras, W.S., Parakkat, J., Chany, A.M., Yang, G., Burr, D. & Lavender, S.A. 2006. Spine loading as a function of lifting frequency, exposure duration, and work experience. *Clinical Biomechanics* 21: 345–352.

Mital, A. 1987. Patterns of differences between the maximum weights of lift acceptable to experienced and inexperienced materials handlers. *Ergonomics* 30: 1137–1147.

Mital, A., Garg, A., Karwowski, W., Kumar, S., Smith, J.L. & Ayoub, M.M. 1993. Status in human strength research and application, *IIE Transactions* 25: 57–69.

Mital, A. & Manivasagan, I. 1983. Maximum acceptable weight of lift as a function of material density, center of gravity location, hand preference, and frequency. *Human Factors* 25: 33–42.

Parakkat, J., Yang, G., Chany, A.M., Burr, D. & Marras, W.S. 2007. The influence of lift frequency, lift duration and work experience on discomfort reporting. *Ergonomics* 50: 396–409.

Yates, J.W., Kamon, E., Rodgers, S.H. & Champney, P.C. 1980. Static lifting strength and maximal isometric voluntary contractions of back, arm and shoulder muscles. *Ergonomics* 23: 37–47.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Regression analysis of fatigue at various shoulder postures in sitting

Ji-Soo Park & Jung-Yong Kim
Department of Industrial and Management Engineering, Hanyang University, Ansan, Korea

ABSTRACT: The purpose of this study is to investigate the level of fatigue of shoulder muscle in various working postures in sitting. An experiment was designed to evaluate the six shoulder muscles at nine shoulder postures including 30 degrees of adduction/abduction, and neutral postures as well as 60, 90, and 120 degrees of shoulder flexion. Thirteen subjects participated in the experiment to collect MVC (Maximum Voluntary Contraction) and MPF (Mean Power Frequency) of SEMG (Surface Electromyography). The participants maintained the shoulder posture in each of nine pre-determined flexion and abduction/adduction shoulder angle for 60 seconds. A regression was calculated out of 60 data points of MPF values, and the slope value was used as fatiguing rate of individual muscle. It was found that a certain muscle was fatigued fast at particular posture compared to other muscles, which would mean that a certain shoulder muscle at particular posture could be more exposed to the risk of musculoskeletal disorders than other muscles.

Keywords: local muscle fatigue, fatiguing rate, shoulder abduction, electromyography

1 INTRODUCTION

The musculoskeletal problems of the shoulder are caused by the highly repetitive work, forceful exertions, high level of static contractions, prolonged static loads, and extreme postures, as well as combinations of these factors (Larsson et al., 2007). 30.4% of workers in manufacturing factories experienced the shoulder pain and injury in Korea. The symptom of MusculoSkeletal Disorders (MSDs) among nurses were reported such as 27.2% of light pain, 25% of moderate pain and 17.3% of severe pain (Kee and Seo, 2007). The localized musculoskeletal disorders were prevailing among workers of light industries in the neck-shoulder part (Åstrand and Rodahl, 1986). 30 to 40% of workers of timber factories were exposed to the musculoskeletal disorder on the neck or upper limb during 12 month survey in Netherlands and Belgium (Buckle and Devereux, 2002).

The shoulder problem is difficult to assess due to the complicated shoulder structure and difficulties in measuring the activity of shoulder muscle. The GlenoHumeral (GH) joint and rotator cuff muscles are affected by the shoulder muscles including the trapezius, serratus anterior, deltoid, supra-spinatus, infra-spinatus, teres minor, subscapularis and so on (Norkin and Levangie, 1992; Basmajian and De Luca, 1985).

Kim et al. (2003) evaluated the shoulder by measuring Maximum Voluntary Contraction (MVC), Root Mean Square (RMS) of ElectroMyoGraphy (EMG) and subjective workload at various shoulder postures. They measured shoulder workload at twenty five working shoulder angles. The results showed that workload increased when abduction and adduction angle increased. Also, the shoulder workload increased as the angle of shoulder flexion increased. Ekstrom et al. (2005) measured the muscle activities of the serratus anterior and upper, middle and lower parts of the trapezius at nine postures among 30 subjects, and the serratus anterior was found to have the high level of EMG activity of the scapula at upwardly rotated position compared with other muscles. Garg et al. (2006) measured the fatigue of the shoulder girdle by using different combination of weight of workpieces, weight of hand-tools, shoulder postures, arm upward and arm downward at automotive assembly operation. They measured SEMG (Surface ElectroMyoGraphy), RPE (Rating of Perceived Exertion), fatigue and pain at 60, 90 and 120 degrees of shoulder flexion. And, no fatigue was found either in the trapezius or the deltoid for most of the tasks when using the Median Power Frequency (MPF). Minning et al. (2007) measured the shoulder muscle fatigue during isometric shoulder elevation at 90 degrees. The middle deltoid was found to be fatigued than other shoulder muscles. But, those studies did not examine the combination of flexion and abduction/adduction.

Thus, in this study, the level of fatigue of shoulder muscles were investigated in various working postures including flexion and abduction/adduction during sitting in order to find which the muscles would be easily fatigued and at particular shoulder posture.

2 METHOD

2.1 *Participants*

Thirteen participants volunteered in this study. They did not experience shoulder problems for six months. Their average age was 26.23 (±5.89) years. Average height and weight were 169.35 (±7.79) cm, 63.50 (±13.32) kg, respectively.

2.2 *Apparatus*

A force indicator (maximum capacity of 50 kg·force (kgf), Bongshin Co., in South Korea) was used in order to measure the sustained shoulder strengths during arm lifting. It was installed on a platform with a chair that was safely changed on it. The position of chair was adjusted for comfortable sitting postures and arm reach to the handle of each participant (Figure 1).

Six channels of bipolar surface electrodes (Coulbourn Instrument Inc., USA) were used with a gain of x2500, bandpass filter of 1 to 1000 Hz, and a sampling rate of 1024 Hz for muscle activity measurement. An analog to digital (A/D) converter and a personnel computer were used for data acquisition.

2.3 *Experimental design*

The independent variables were the shoulder flexion angle (three levels of 60, 90, and 120 degree in the sagittal plane) and shoulder ad/abduction angle (−30 and 30 degrees in transverse plane). Six muscles (Figure 2) were also independent variables.

The dependent variables were the normalized MPF value of six shoulder muscles at workload

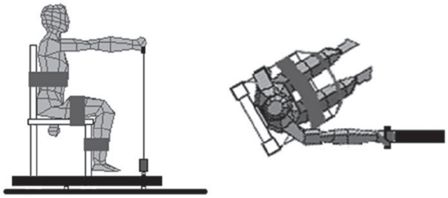


Figure 1. The 90 degrees of flexion and 30 degrees of abducttion of shoulder posture and on the chair platform.

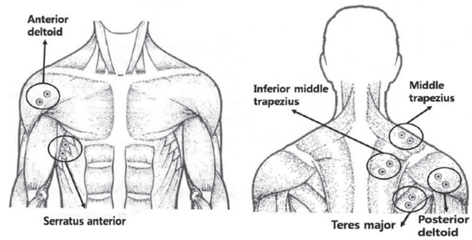


Figure 2. Muscle location for shoulder movement.

of 30% MVC. The nine postures of shoulder were measured with a digital goniometer placing its center at the acromion.

2.4 *Experimental procedures*

First of all, the instruction of experiment was informed when the participant arrived at the testing site. Participant was seated on the chair with the feet off the ground and then the torso, lower legs, and upper legs were strapped on the seat to isolate them from any involvement of force exertions during the experiment. Surface electrodes were attached to the middle trapezius, inferior middle trapezius, anterior and posterior deltoid, serratus anterior and teres major in an appropriate location. The scapulae and rotated the upper arm are abducted the middle trapezius and inferior middle trapezius. The deltoid is responsible to abduct of the shoulder and elevated of the upper arm. The serratus anterior is responsible to adduct of the upper arm and elevated of the scapulae. The teres major is responsible to abduct of the upper arm (Åstrand and Rodahl, 1986; Norkin and Levangie, 1992; Shin et al., 1993) (Figure 2). And, EMG signal maintained the resistance of 10 Ω and electrode distance of 15 mm.

The participant was fitted to the arm reach toward the handle comfortably to pull up the handle. MVC was measured twice with given shoulder flexion, adduction, and abduction angles during five seconds. The average value of MVCs was used to calculate the 30% MVC for EMG measurements. The EMG signals were measured at each of nine shoulder postures during 60 seconds. The MPF value was calculated the signals from the first and last five seconds. Five-minute resting intervals were provided to the participant between each trial.

2.5 *Data analysis*

The Mean Power Frequency (MPF) was calculated froms 1 second data of EMG signal during 50 seconds by using FFT (Fast Fourier Transformation). The MPF value was normalized based on the initial value of the experiment measured during a resting state. And, the MPF values collected during experiment were used to compute regression equation and finally calculated the slope of equation and used as the rate of fatigue of individual muscle in this study.

3 RESULTS

3.1 *Maximum voluntary contraction*

The MVC result was set as 100% in the neutral posture (zero flexion and zero ab/adduction), and then MVCs at 60, 90, and 120 degrees of flexion were compared. The result showed that 30% at 60 degrees, 20% at 90 degrees, and 20% at 120 degrees of the flexion postures, respectively. By the same method, the MVCs at 30 degrees of adduction and 60, 90, 120 degrees of flexion postures were recorded as 20%, 16%, and 15%, respectively. And, the MVCs at 30 degrees of abduction posture were recorded as 21%, 15%, and 15% (Figure 3).

3.2 *Muscle fatigue*

The ANOVA results of normalized MPFs at the start and finish of the experiment showed the difference depending on the muscles and changes of flexion and abduction/adduction as shown in Table 1.

The middle trapezius and inferior middle trapezius were not significantly fatigued at all postures. The decline of MPF slope was less than 10% at all postures. The anterior deltoid was significantly fatigued all postures. The decline of MPF slope was more than 10% at all postures (Table 2).

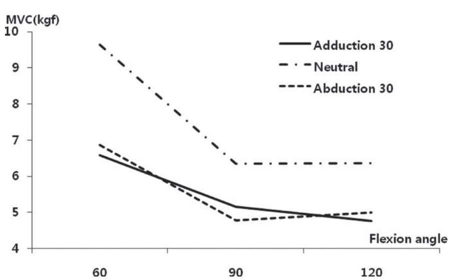


Figure 3. Variation of MVC at each flexion angle.

Table 1. Difference of MPF at each posture, muscle and gender (*:α = 0.05).

Source	Type III sum of squares	d.f.	Mean square	F	Sig
Ab/adduction	1636.59	2	818.29	0.89	0.410
Flexion	13265.96	3	4421.99	4.82	*0.002
Muscle	62407.15	6	10401.19	11.35	*0.001

Table 2. The slope of linear regression that is the ratio of MPF shift at various shoulder postures.

Abduction angles	Flexion angles	Middle trapezius (%)	Inferior middle trapezius (%)	Antrior deltoid (%)
–30	60	0.01	1.99	15.06
	90	1.52	2.12	15.59
	120	3.63	1.13	15.29
0	60	0.45	8.46	12.04
	90	0.01	1.73	14.86
	120	5.16	0.01	14.87
30	60	1.56	5.58	10.40
	90	0.82	7.04	15.84
	120	6.12	6.30	16.98
Abduction angles	**Flexion-angles**	**Posteri-ordeltoid (%)**	**Serratu-santerior (%)**	**Teres-major (%)**
–30	60	5.33	5.92	14.72
	90	13.87	13.01	17.73
	120	7.78	10.51	13.51
0	60	7.15	13.41	5.93
	90	10.47	15.19	11.94
	120	10.90	14.05	14.95
30	60	7.19	12.55	8.87
	90	14.22	16.62	13.65
	120	13.75	20.65	10.33

The posterior deltoid was significantly fatigued 30 degrees of adduction and 90 degrees of flexion. And, neutral and 30 degrees of abduction postures were fatigued more than 90 degrees of flexion. The serratus anterior was significantly fatigued except for 30 degrees of adduction and 60 degrees of flexion posture. The teres major was significantly fatigued except for neutral and 60 degrees of flexion, 30 degrees of abduction and 60 degrees of flexion postures (Table 2).

4 DISCUSSION

Regarding the level of workload causing the shoulder muscle fatigue, Öberg (1995) reported that the 11% of slope change caused by MPF slope with heavier than 2 kg load and 3 minutes endurance time was great enough to make subjects experience the fatigue. In the current study, 10% of slope change by average MPF slope was observed under 30% MVC and 60 seconds endurance time, which was considered as a substantial level of fatigue that participants could perceive during the experiment.

In terms of the results of MVC (Figure 3), there was no significant difference between 90 and 120 degrees of flexion of shoulder. The amount of MPF slope showed no difference between those shoulder angles either. This could be due to the uncomfortable posture at 120 degrees flexion in spite of short moment arm compared with the 90 degrees flexion.

Many researchers have reported that the local muscle fatigue was due to shoulder angles which were greater than 60 degrees of flexion with/without abduction (Larsson et al., 2007; Garg et al., 2006; NIOSH, 1997). Minning (1991) measured the EMG signal of the upper trapezius, lower trapezius, serratus anterior and middle deltoid at 60 degrees of abduction and 90 degrees of flexion. The results indicated that the MPF of the middle deltoid was greater than other muscles which were similar with the current result such that the MPF slope of the deltoid muscle was greater than 10% at 30 degrees of abduction and 90 degrees of flexion.

In this study, the serratus anterior showed a great MPF slope at 30 degrees of abduction and 120 degrees of flexion. The teres major also showed a great MPF slope at 30 degrees of adduction and 90 degrees of flexion. The great MPF slope of the anterior deltoid was shown at 30 degrees of abduction and 120 degrees of flexion. The great MPF slope of the posterior deltoid was found 30 degrees of abduction and 90 degrees of flexion. According to these results, it can be said that any shoulder muscle can be easily exposed to the risk of MSDs under a given working posture.

Thus, it is very important to understand the cause and effect relationship between the shoulder posture and fatigue. It is expected that the information disclosed in this study would assist ergonomist in predicting the relative workload at various shoulder postures. Since the study only covered static muscle exertion, applying the result to workplace with mostly dynamic tasks would require additional consideration depending on the nature of individual task at workplace. Therefore, a further and continuous investigation of the shoulder fatigue needs to be conducted to better understand the mechanism of shoulder MSDs and prevent them as much as possible.

5 CONCLUSIONS

This study measured the difference of muscle fatigue at various shoulder postures. The easily fatigued or possibly vulnerable muscles were found at specific working postures. It is expected that the results of this study can be used as the applicable information for workplace design in spite of experimental shortcomings.

REFERENCES

Åstrand, P.O. & Rodahl, K. 1986. *Textbook of work physiology: Physiological bases of exercise.* 3rd ed. McGraw-Hill, New York.

Basmajian John, V. & Carlo, J. De Luca. 1985. *Muscle alive: Their functions revealed by electromyography.* 5th ed. Williams & Wilkins: London.

Buckle, P.W. & Devereux, J.J. 2002. The nature of work-related neck and upper limb musculoskeletal disorders. *Applied Ergonomics* 33: 207–217.

Ebaugh, D.D., McClure, P.W. & Karduna, A.R. 2006. Effects of shoulder muscle fatigue caused by repetitive overhead activities on scapulothoracic and glenohumeral kinematics. *Journal of Electromyography and Kinesiology* 16: 224–235.

Ekstrom, R.A., Soderberg, G.L. & Donatelli, R.A. 2005. Normalization procedures using maximum voluntary isometric contractions for the serratus anterior and trapezius muscles during surface EMG analysis. *Journal of Electromyography and Kinesiology* 15: 418–428.

Farina, D., Madeleine, P., Graven-Nielsen, T., Merletti, R. & Arendt-Nielsen, L. 2002. Standardising surface electromyogram recordings for assessment of activity and fatigue in the human upper trapezius muscle. *European Journal of Applied Physiology* 86: 469–478.

Garg, A., Hegmann, K.T., Schwoerer, B.J. & Kapellusch, J.M. 2002. The effect of maximum voluntary contraction on endurance times for the shoulder girdle. *International Journal of Industrial Ergonomics* 30: 103–113.

Garg, A., Hegmann, K. & Kapellusch, J. 2006. Short-cycle overhead work and shoulder girdle muscle fatigue. *International Journal of Industrial Ergonomics* 36: 581–597.

Kee, D. & Seo, S.R. 2007. Musculoskeletal disorders among nursing personnel in Korea. *International Journal of Industrial Ergonomics* 37: 207–212.

Kim, J.Y., Chung, M.K. & Park, J.S. 2003. Measurement of physical work capacity during arm and shoulder lifting at various shoulder flexion and Ad/Abduction angles. *Human Factors and Ergonomics in Manufacturing* 13(2): 153–163.

Minning, S., Eliot, C.A., Uhl, T.L. & Malone, T.R. 2007. EMG analysis of shoulder muscle fatigue during resisted isometric shoulder elevation. *Journal of Electromyography and Kinesiology* 17: 153–159.

NIOSH 1997. In Bernard, B.P. (Ed.) *Musculoskeletal disorders and workplace factors: A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back.* DHHS (NIOSH) Publication No. 97–141. US Department of Health and Human Services, NIOSH, Cincinnati, Ohio.

Norkin, C.C. & Levangie, P.K. 1992. *Joint structure and function: A comprehensive analysis.* 2nd ed. F.A. Davis Company.

Öberg, T. 1995. Muscle fatigue and calibration of EMG measurements. *Journal of Electromyography and Kinesiology* 5(4): 239–243.

Pedersen, J., Lönn, J., Hellström, F., Djupsjöbacka, M. & Johansson, H. 1999. Localized muscle fatigue decreases the acuity of the movement sense in the human shoulder. *Medicine & Science in Sports & Exercise*, 31(7): 1047–1052.

Shin, M.K., Koo, J.Y., Kim, B.W., Shim, J.W., Yoo, K.W. & Choi, K.P. 1993. *Human Anatomy*. Hyunmoon Co.: Seoul. Korea.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Influence of circadian rhythm on maintenance of body shape

Shinya Kume
Graduate School of Science and Engineering, Ritsumeikan University, Shiga, Japan

Hiroshi Hagiwara
College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan

Namiko Tokumitsu & Saori Sakamoto
Wacoal Corporation Human Science Research Center, Kyoto, Japan

ABSTRACT: We examined differences in daily life, focusing on circadian rhythms, among women whose body shapes changed in different ways over 20 years. We examined an evaluation of daily living using amount of activity and clarified the influence of lifestyle on the maintenance of body shape. A circadian rhythm experiment was conducted on weekdays for 1 week with 24 healthy women in their 40s and 50s. Amount of activity of the non-dominant wrist and trunk, subjective evaluation of sleep quality, and subjective state of activity were surveyed. In order to maintain a constant body shape throughout life, it is important to sustain a high level of trunk movement during the day, sleep well at night, have a varied pace of life, and maintain a regular sleep-wake cycle. These results suggest that using portable accelerometers as "life recorders" for self-discipline can be effective in maintaining lifestyle and body shape.

Keywords: circadian rhythm, body shape changes, amount of activity, sleep quality, evaluation of daily living

1 INTRODUCTION

A recently developed piece of equipment, called the "life recorder" (Otani 2008), allows us to easily and unobtrusively obtain valuable physiological information over a long period. The "life recorder" has gained attention in the area of self-discipline for the prevention of metabolic syndrome and other lifestyle-related diseases. However, the function of this device is primarily calculating calories and metabolic equivalents (METs) from measurements of physical activity during the day. In order to record important lifestyle data, however, it is also necessary to take measurements at night, including during sleep. In addition, evaluation based on physiological and behavioral phenomena in daily situations is important in evaluating daily life.

The present study focused on circadian rhythm, a biological rhythm based on an approximately 24-hour cycle that influences physiological and behavioral phenomena. Several studies have investigated circadian rhythm in various subjects, including individuals with sleep disorder, shift workers, people with jet lag, and Alzheimer's disorder patients, in a variety of environmental conditions. However, little is known about circadian rhythm in healthy people in normal daily situations. Huang et al. (2002), in a study on the sleep and rest-activity patterns of 65 young, middle-aged, old and elderly subjects in daily situations, found that older subjects showed weakened and fragmented circadian sleep and rest-activity rhythms, and differences were apparent between subjects within the same age groups. The question we ask here is whether changes in physical or health conditions over time are affected by circadian rhythm.

The purpose of this study was to examine an evaluation of daily life using amount of activity and to clarify the influence of lifestyle on the maintenance of body shape throughout life. We considered physical and health changes to represent body shape changes. The present paper employs an evaluation of daily life focusing on circadian rhythm in people with different body shape changes.

We previously investigated body shape changes in the same woman over approximately 20 years. In the present study, we assigned body shape changes to a numerical scale and classified 93 women in their 40s and 50s into 3 groups. We then evaluated the daily life of each subject, focusing on circadian rhythm.

2 EXPERIMENTAL METHOD

2.1 *Body shape changes*

2.1.1 *Body shape vector*

Body shape (body form) was defined as the shape of the subject's silhouette projected onto a vertical plane (Kurokawa et al. 1984). Body shape data were collected using a sequence of nodes on the projected silhouette image. The nodes were extracted from the silhouette image using a curve-dividing algorithm so that linear interpolating splines passing through the image could be used to reconstruct the original body shape within a certain error. In this sense, the obtained node sequences can be said to describe the body form. In addition, it was shown that an average body shape can be drawn by arithmetically averaging the coordinates of the corresponding nodes among the different silhouette images.

In this way, human body shape data (front view and side view) projected onto a vertical plane can be compressed into 173 nodes, the locations of which were optimized in order to most accurately describe the body shapes of at least 99% of women. In this research, body shape vector, as the horizontal coordinates of 173 node locations, was used to classify body shape into patterns.

2.1.2 *Classification of body shape pattern*

We compared the mean body shape at each age (20s, n = 418; 30s, n = 771; 40s, n = 378; 50s, n = 102) to the individual body shape of each subject over approximately 20 years, and then classified 93 women in their 40s and 50s into 3 groups (groups A, B, and C) as follows (Fig. 1).

Group A comprised women who, when they were in their 20s, had body shapes close to the mean body shape for women in their 20s and 30s, and whose body shape had changed little over the 20 years between their 20s and their 40s (i.e., women who had maintained the same body shape from their 20s through their 40s).

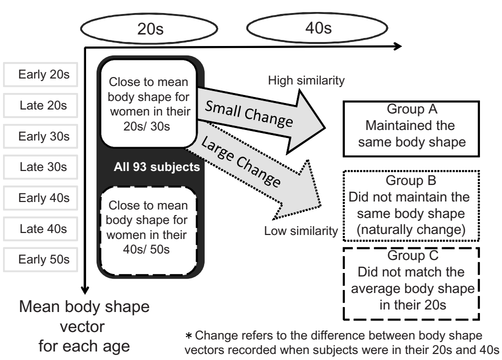


Figure 1. Body shape changes.

Group B comprised women who, when they were in their 20s, had body shapes close to the mean body shape for women in their 20s and 30s, and whose body shape changed markedly over the 20 years between their 20s and their 40s (i.e., women who had not maintained the same body shape from their 20s through their 40s due to natural changes in weight and fat distribution).

Group C comprised women who, when they were in their 20s, had body shapes close to the mean body shape for women in their 40s and 50s (i.e., women whose body shape did not match the average for their age group when they were in their 20s).

As mentioned above, "close to the mean body shape" means that the difference between the mean body shape vector for each age and the individual body shape vector was relatively small. "Small" and "large" changes in body shape were considered to have taken place when the difference between the body shape vector recorded when a subject was in her 20s and that recorded when she was in her 40s was among the 10 smallest values (high similarity) or the 10 largest values (low similarity), respectively.

2.1.3 *Characteristics of groups (A, B, and C)*

Although we will not discuss the numerical data in detail here, the fundamental indexes measured in advance in the three groups were as follows. Weight, body fat percentage, fat mass, Body Mass Index (BMI), and trunk circumference were lower in group A than in groups B and C. Muscle and bone mass were not different between groups A and B, and were lower in group C than in groups A and B. Indexes of fundamental physical strength (e.g., Kraus-Weber Test, an evaluation of trunk fitness) were higher in group A than in groups B and C. These results suggest that the difference in body shape change between groups A and B is the result of different rates of body fat acquisition during aging, and that motor ability is higher in group A than in group B.

2.2 *Subjects*

Twenty-four healthy, non-medicated women (group A, n = 9; group B, n = 8; group C, n = 7) in their 40s and 50s were selected from the general public to participate in the experiment.

All subjects wore two portable accelerometers, one on the non-dominant wrist and the other around the waist, at all times except when bathing and changing clothes. In addition, subjects completed Oguri-Shirakawa-Azumi sleep inventory for middle age and aged (OSA sleep inventory MA version) (Yamamoto et al. 1999) immediately after waking and kept daily log. Daily log was used to record time use, activities and sleep status. It is entered their activities during the investigation onto the diary, on which time scales were printed.

During the experimental period, subjects were instructed to go about their daily lives as usual. Informed, written consent was obtained from each subject prior to participation in the study.

2.3 *Materials*

Micro-mini Actigraph (Ambulatory Monitoring Inc., Ardsley, NY, USA) was used to measure the amount of physical activity of the non-dominant wrist. The Actigraph, an acceleration-sensitive device resembling a wristwatch, was set to proportional integrating mode (low-PIM), which measures the intensity of a movement by integrating the deviations from 0 V every 0.1 seconds and stores the intensity once every minute.

Life-corder Plus (Suzuken Co Ltd., Nagoya, Japan) was used to measure the amount of physical activity of the trunk. The Life-corder measures the intensity (signal of 0.5 or 1 to 9) of movement every 4 seconds and stores the highest frequency signal for each 2-minute period once every 2 minutes. The level of intensity is recorded on a scale of 1 (lowest intensity) to 9 (maximum intensity).

OSA sleep inventory (MA version) is a quantitative index that evaluates five factors: the first factor is sleepiness on waking, the second factor is initiation and maintenance of sleep, the third factor is frequency of dreaming, the fourth factor is recovery from fatigue, and the fifth factor is sleep length. OSA contains 16 questions to evaluate subjective sleepiness and is completed immediately after waking. The MA version is used for people who do not have enough time for writing in clinical scene to complete the full survey, and for middle-aged and elderly subjects who cannot properly respond to many questions.

Daily log is a record of time use, activities, and sleep status that filled onto scaled recording paper throughout a day, and kept recorded during investigation period. The log contains aspects of behavior as follows: sleep, nap, rest, work, housework, move, exercise, meals, bath, and periods when the accelerometer was not worn. In addition, it also contains questions about subjective sleep quality, condition of health, and amount of activity.

SPSS (SPSS Inc., Chicago, Illinois, USA) was used for statistical analyses. A two-tailed t-test was used for independent samples to assess the specific difference between groups.

2.4 *Experimental procedures*

The circadian rhythm experiment was conducted on weekdays for 1 week in July or August. The amount of activity of the non-dominant wrist and trunk, subjective evaluation of sleep quality, and subjective state of activity were surveyed on weekdays for 1 week while subjects carried out their daily activities as usual.

3 ANALYTICAL METHOD

3.1 *Amount of activity*

Amount of activity was classified on a time series for weekdays during active periods, sleeping periods, and non-measurement periods (e.g., bathing) using daily log.

The mean amount of activity during the Active Period (mean AP) gives an indication of the usual daily activity level during the daytime. Mean AP values are high for subjects with a high level of physical activity and low for those with a low level of physical activity during the daytime.

The mean amount of activity during the Sleep Period (mean SP) gives an indication of sleep quality. Mean SP values are high for subjects with poor sleep quality (i.e., light sleep) and low for those with high sleep quality (i.e., deep sleep). A recent investigation has demonstrated that most physical activity during sleep occurs during the shift from deep sleep to light sleep, or during light sleep (Jo & Hagiwara 2009).

3.2 *Indicator of variation in pace of life*

An indicator of variation in pace of life was used to measure the intensity of the sleep-wake cycle. Variation in pace of life was high for subjects with a high level of physical activity during the daytime and high sleep quality. Variation in pace of life was calculated as the ratio of mean AP to mean SP according to Equation 1 below:

$$\text{Variation in pace of life} = \frac{\text{Mean of AP}}{\text{Mean of SP}} \qquad (1)$$

3.3 *Indicator of regularity of life*

An indicator of regularity of life was used to measure the regularity of the sleep-wake cycle. Sleep-wake cycle was numerically expressed as 0 (sleep) or 1 (awake). For analysis, sleep-wake data (0 or 1) were recorded over a 1-minute interval at the same times on 3 consecutive days (Tuesday, Wednesday, and Thursday) using daily log. For example, 000 indicates that the subject was asleep (0) at the same time on all 3 days; 111 indicates that the subject was awake (1) at the same time on all 3 days; and 001, 010, 100, 011, 101, and 110 indicate irregular sleep-wake cycles for the indicated times (0.5).

Regularity of life was plotted on a line graph for a 24-hour period and the sleep-wake cycles of all 24 subjects were classified into distinctive patterns.

3.4 *Subjective sleepiness*

OSA sleep inventory was used to get subjective sleep score. It was compared with the standard score of 50 obtained from data from 670 subjects aged 26 to 75 (Yamamoto et al. 1999). Scores higher than 50 were considered to indicate high sleep quality.

3.5 *Sleep parameters*

Sleep parameters including bedtime, time at waking, total sleep time (hours from bedtime on Monday to waking on Friday, including naptime), and total sleep frequency (number of times from bedtime on Monday to waking on Friday, including naptime) were derived from information recorded in activity diaries.

4 RESULTS

4.1 *Amount of activity*

Amount of activity of the non-dominant wrist is shown in Figure 2; amount of activity of the trunk is shown in Figure 3. The mean and standard deviation of mean amount of activity of the non-dominant wrist for each group are shown in Figure 4; the mean and standard deviation of mean amount of activity of the trunk for each group are shown in Figure 5.

During active periods, the mean amount of wrist movement was higher in group B than in groups A and C (Fig. 4-a). The mean amount of trunk activity was higher in group A than in groups B and C (Fig. 5-a). During active periods, the recorded amount of activity suggests that group B exhibited more active wrist movement than groups A and C, but group A exhibited more active trunk movement than groups B and C. Thus, group A had a high level of general activity daily life, while groups B and C had a high level of wrist movement and a low level of general body movement in daily life.

During sleep periods, the mean amounts of wrist and trunk activity were lower in group A than in groups B and C (Fig. 4-b, Fig. 5-b). A two-tailed t-test for independent samples revealed a significant difference ($p < 0.05$) between groups A and C (Fig. 4-b). During sleep periods, the recorded amount of activity suggests that group A had higher sleep quality than groups B and C.

Variation in wrist and trunk movement was found to be higher in group A than in groups B and C (Fig. 4-c, Fig. 5-c). A two-tailed t-test for

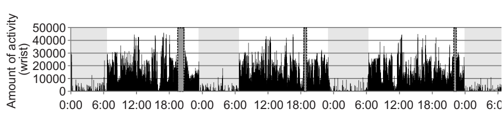


Figure 2. Amount of activity of the non-dominant wrist for 1 of the 24 subjects as recorded by Micro-mini Actigraph on consecutive weekdays beginning at bedtime on Monday to waking on Friday. White areas indicate active periods. Lightly shaded areas indicate sleep periods. Dark shaded areas between dashed lines indicate non-measurement periods.

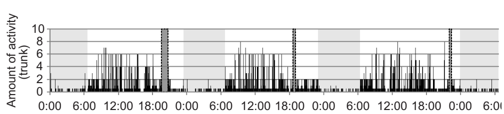


Figure 3. Amount of activity of the non-dominant wrist for 1 of the 24 subjects as recorded by Micro-mini Actigraph on consecutive weekdays beginning at bedtime on Monday to waking on Friday. White areas indicate active periods. Lightly shaded areas indicate sleep periods. Dark shaded areas between dashed lines indicate non-measurement periods.

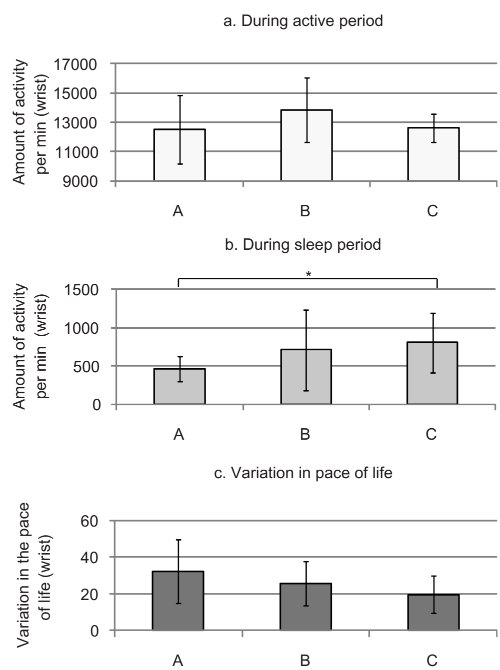


Figure 4. Mean and standard deviation of mean amount of activity of the non-dominant wrist for each group (group A, n = 9; group B, n = 8; group C, n = 6). A significant difference ($p < 0.05$) in sleep period was noted between groups A and C.

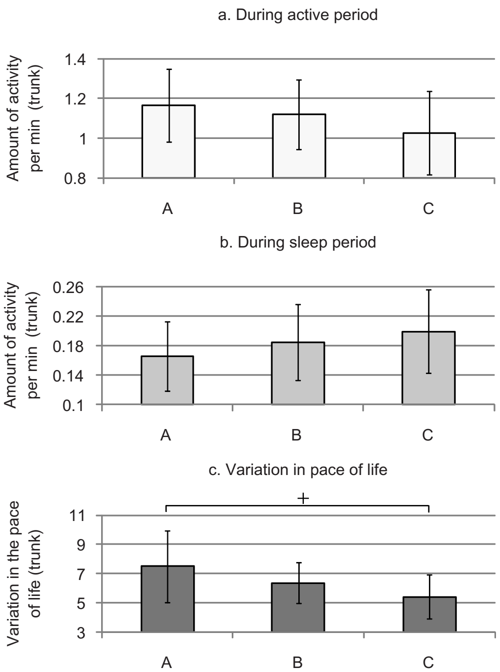


Figure 5. Mean and standard deviation of mean amount of activity of the trunk for each groups (group A, n = 8; group B, n = 8; group C, n = 7). A significant difference ($p < 0.10$) in variation in pace of life was noted between groups A and C.

independent samples revealed a significant difference ($p < 0.10$) between groups A and C (Fig. 5-c). Variation in pace of life suggests that group A was more active and rested more than groups B and C. Group A was very active in the daytime and slept well at night.

4.2 *Indicator of regularity of life*

The mean and standard deviation of rate of irregular sleep-wake cycle (0.5) for each group are shown in Figure 6. The rate of irregular sleep-wake cycle (0.5) was lower in group A than in groups B and C. Sleep-wake cycle was classified into 5 types: regular, irregular, arousal during sleep, daytime nap, and nighttime work.

Data for the regular type are shown in Figure 7-a.

In this type, instances of irregular sleep-wake cycle are rare and the time at waking and bedtime are relatively constant. The irregular type is shown in Figure 7-b. In this type, periods of irregular sleep-wake cycle are long and the time at waking and bedtime are not constant. The arousal during sleep type is shown in Figure 7-c. This type exhibits irregular sleep-wake cycle (0.5) during sleep.

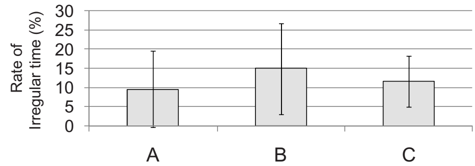


Figure 6. Rate of irregular sleep-wake cycle (0.5) was lower in group A than in groups B and C.

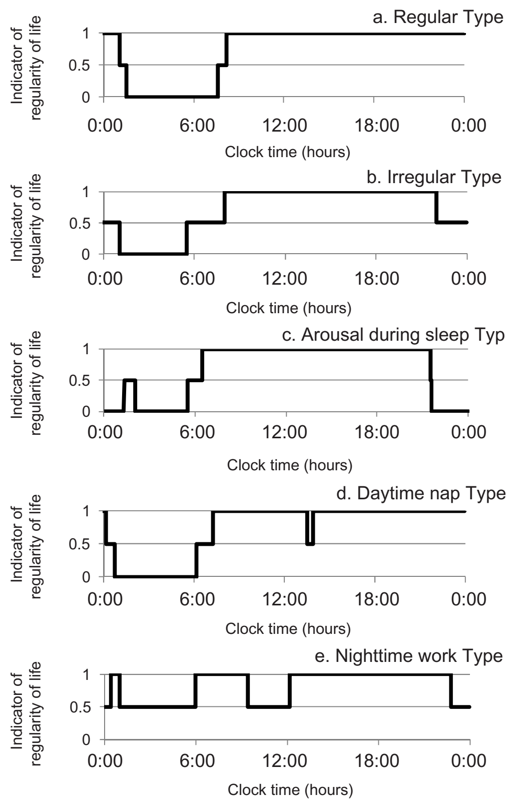


Figure 7. Sleep-wake cycle was classified into 5 types (regular, irregular, arousal during sleep, daytime nap, and nighttime work).

The daytime nap type is shown in Figure 7-d. This type exhibits irregular sleep-wake cycle (0.5) during the day. The nighttime work type is shown in Figure 7-e. This type has no regular sleep (0).

The number of patterns of sleep-wake cycle are shown in Table 1. Group A included more subjects with regular type, while groups B and C included more subjects with irregular and arousal during sleep types. There were no differences among the groups in the frequency of daytime nap type. One subject in group B was classified as nighttime work type.

4.3 *Subjective sleepiness*

The mean and the standard deviation of subjective sleep score were calculated (Table 2). Subjective sleep score in all groups was lower than the standard score of 50 except for the third factor of frequency of dreaming in groups B and C. Regarding the quality of sleep, scores for the second factor of initiation and maintenance of sleep and the fourth factor of recovery from fatigue were higher in group A than in groups B and C.

4.4 *Sleep parameters*

The mean and standard deviation of sleep parameters were calculated (Table 3). Total sleep time (from bedtime on Monday to waking on Friday, including naptime) was 25–26 hours in all groups. Total sleep frequency (from bedtime on Monday to waking on Friday, including naptime) was 5–6 times in all groups. Mean bedtime ranged from 23:51 to 0:13 in all groups. Mean waking time ranged from 6:40 to 6:50 in all groups. Therefore, all subjects slept for approximately 6.5 hours during each sleep period. There were no significant differences among the groups for these sleep parameters.

Table 1. Patterns of sleep-wake cycle.

Types	Group A Number	Group B Number	Group C Number
Regular	5	1	2
Irregular	1	1	3
Arousal during sleep	1	3	2
Daytime nap	2	2	0 (*2)
Nighttime work	0	1	0

* Includes other types with daytime nap.

Table 2. Subjective sleepiness.

Factor	Group A Score*	Group C Score	Group B Score
Sleepiness on rising	42.6	40.9	44.8
Initiation and maintenance	48.7	45.1	48.1
Frequency of dreaming	48.4	56.4	51.5
Recovery from fatigue	46.4	42.5	44.4
Sleep length	43.2	42.8	42.9

* Standardized scores.

Table 3. Sleep parameters.

Values	Group A mean (SD)	Group B mean (SD)	Group C mean (SD)
TST (hours)	25.78 (4.23)	26.11 (5.49)	24.90 (3.63)
TSF (number)	5.67 (1.94)	5.25 (1.04)	5.43 (2.70)
Mean BT	0:13 (1:01)	23:51 (1:23)	0:14 (1:06)
Mean WT	6:41 (0:37)	6:45 (0:52)	6:50 (1:12)

* Values are mean (standard deviation; TST, total sleep time; TSF, total sleep frequency; Mean BT, mean bedtime; Mean WT, mean waking time).

5 DISCUSSION

The purpose of this study was to examine an evaluation of daily life using amount of activity and to clarify the effect of lifestyle on the maintenance of body shape throughout life. We examined differences in daily life, focusing on circadian rhythms, in women whose body shapes had changed in different ways over 20 years.

The present results are consistent with the findings of a previous studies using a similar "life recorder" system, which found that evaluation of physical activity during the daytime is important for physical or health changes (Hara et al. 2006, Matsumura et al. 2008). Sleep quality, variation in pace of life, and regularity of life are closely correlated with physical and health changes. Therefore, in order to identify important lifestyle factors, it is necessary to evaluate activity levels throughout the day, including during sleep.

The present findings agree with the results of previous study that has noted differences in activity levels during active period and sleep period and activity pattern among people of the same ages (Huang et al. 2002). Additionally, people who experience different body shape changes (physical or health changes) have different circadian rhythms. Therefore, there may be a correlation between the maintenance of regular circadian rhythm and the maintenance of body shape (and other physical or health factors) during the course of aging.

Maintaining circadian rhythms, such as rest-activity and sleep-wake cycle, may be one factor behind the maintenance of body shape with aging. If body type is maintained, it is thought that an appropriate amount of body fat and motor ability can be maintained as well.

In the future, to maintain body shape, it will be necessary to investigate lifestyle during the daytime in detail and prescription for high quality of sleep. In the present study, the results were recorded over the course of only 1 week under usual daily environmental condition, and subjects' lifestyles during the experimental period were not necessarily the same as during the previous 20 years. In addition, in order to accurately investigate lifestyle, data must be continuously recorded. In the future, the effects of gender, age, and occupation should also be examined. To identify differences in lifestyle, it will be necessary to research daily life in detail, including amount of activity.

6 CONCLUSION

The present study investigated an evaluation of daily living focusing on circadian rhythm in subjects who had experienced different body shape changes over 20 years.

The results suggest that circadian rhythm—including a high level of trunk movement during the day, sleeping well at night, having a varied pace of life, and maintaining a regular sleep-wake cycle—is important in the maintenance of a constant body shape throughout life.

Portable accelerometers allowed us to easily and unobtrusively measure the amount of activity. From the measured data, we were able to estimate lifestyle-related body shape changes focusing on circadian rhythm. Thus, the present method is considered to represent the effective use of a "life recorder" for health care and self-discipline.

ACKNOWLEDGEMENTS

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 22500415, 2010.

REFERENCES

Hara, T. et al. 2006. The relationship between body weight reduction and intensity of daily physical activities assessed with 3-dimension accelerometer. *Japanese Journal of Physical Fitness and Sports Medicine* 55: 385–392.

Huang, Y-L. et al. 2002. Age-associated difference in circadian sleep-wake and rest-activity rhythms. *Physiology & Behavior* 76: 597–603.

Jo, K. & Hagiwara, H. 2009. Evaluation of body movement during sleep using an infrared motion sensor. *Japanese Society for Medical and Biological Engineering* 47(1): 7–14.

Kurokawa, T. et al. 1984. A method for describing human body forms. *Journal of Society for Instrumentation and Control Engineering* 20(9): 829–836.

Matsumura, Y. et al. 2008. Physical activity measurements based on 3-axis acceleration method. *Panasonic electric works technical report* 56 (2): 67–72.

Otani, T. (ed.) 2008. Life recorders monitor daily activity. *NIKKEI ELECTRONICS ASIA* July.

Yamamoto, Y. et al. 1999. Standardization of revised version of OSA sleep inventory for middle age and aged. *Brain Science and Mental Disorders* 10: 401–409.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Flexor tendon force ratios affect grip force and finger joint motion on power grip: A cadaver model

Shi-Hyun Park
University of Nebraska at Omaha, Omaha, USA

Jesun Hwang & Andris Freivalds
Pennsylvania State University, University Park, USA

ABSTRACT: We studied the effect of flexor tendon force ratio with regard to the grip force and the finger joint angle on power grip motion in a cadaver model using the hand motion simulator. The hand motion simulator was developed with a force delivery unit and data acquisition system. The simulator controlled both flexor tendon forces (FDP and FDS) ranged from 1:1 to 1:6 of FD:FDP. 40% of FDS force to total tendon force (i.e. 3:2 FDP to FDS) showed the highest grip force. Also, the internal tendon force was 5.3 times higher than externally applied grip force, and the efficiency of the forces was best at the 3:2 FDP to FDS force ratio. The finger joint angles were changed according to the tendon force ratios of FDP to FDS. These results demonstrate that awkward finger motion and posture may lose the force efficiency between internal and external forces and cause higher internal impact that may be a risk factor causing carpal tunnel syndrome.

Keywords: hand, FDP, FDS, flexor, handle

1 INTRODUCTION

1.1 *Forceful exertions*

The Hand Motion Forceful exertion of tendons while gripping hand tools may be one of the factors that lead to the development of work-related musculoskeletal disorders (WRMSDs) (NIOSH 2005). To reduce the impact of gripping a handle, understanding finger functionality associated with the specific muscle is crucial. Forceful exertion required to do the task plays an important role in the onset of WMSDs. More force equals more muscular effort, and consequently, a longer time is needed to recover between tasks. Since in repetitive work, as a rule, there is not sufficient time for recovery, the more forceful movements develop fatigue much faster (Chaffin 1973). Forceful exertions of the upper extremities may cause upper extremity musculoskeletal disorders such as joint inflammation, muscle spasms, sprains, tendinitis, or diseases of the peripheral nerves (Armstrong et al. 1999, Silverstein et al. 1987).

1.2 *Finger force and motion*

Forceful Finger forces have been studied (Dickson et al. 1972, Ohtsuki 1981). However, application of the result to hand tool design was very few. The distribution of finger force changes as the type of grip changes. Therefore, the design of grasping and control should provide both maintenance of a secure grip and comfort of the hand posture (Oh 1993). In addition, Gilbert et al. (1988) found out that there is fatigue in the fingers due to the range of motion of the fingers even though the minimal force is required in finger tasks. Thus, understanding the relationship of finger motions to flexor tendon force is required to reduce fatigue and risk factors in grasping a tool. Yet, the exact motion effect associated with an individual flexor muscle is still unknown. Especially, none of studies were performed on the multi-finger grasping (power grip) with handles in a cadaver model. Hence, the purpose of this study was to examine grip forces and finger joint motions generated by different flexor tendon force ratios.

Motivation of this study came from the hypothesis that finger joint angles will have different postures depending on flexor tendons force ratios. Therefore, it is hypothesized that there is an optimal flexor tendon force ratio to minimize internal tendon forces and maximize externally applied grip force. It is also hypothesized that the finger joint angles with the optimal tendon force ratio should be an optimal power grip posture reducing internal impact and causing less fatigue in the finger.

2 METHODS

2.1 *Apparatus*

The Hand Motion Simulator (HMS) was built to generate flexor tendon forces and simulate hand motions and postures with a cadaver hand. Muscle forces generated by two linear actuators with force feedback control were applied to the tendons of the extrinsic muscles of the hand (Fig. 1). The HMS was composed of five essential parts: frame supporting a specimen, motion delivery unit through stepper motor driven linear actuators applying forces to the muscle tendons, data acquisition unit for force transducers measuring internal and external forces, and operating program to control the HMS. In addition, finer force distributions were measured using FlexiForce force Sensitive Resistor (FSR) sensors (Tekscan, Boston, MA) on each phalange (distal, middle, proximal, and metacarpal) (Fig. 2). We previously validated the simulator with a force feedback control through the comparison between desired forces and actual tendon forces generated by the hand motion simulator (Park 2009).

2.2 *Specimen*

One female fresh-frozen human cadaveric left hand specimen was used in this study. The specimen was amputated at the middle humerus and was free from apparent musculo-skeletal disorders and anatomical abnormalities. After thawing overnight at room temperature, the specimen was minimally dissected to expose the musculotendinous junctions of the extrinsic muscles. The specimen was prepared with the entire forearm below the elbow joint and mounted into the Hand Motion Simulator after the preparation. Since we only focused on flexor muscles among the extrinsic muscles, flexors of the extrinsic were grouped into each tendon.

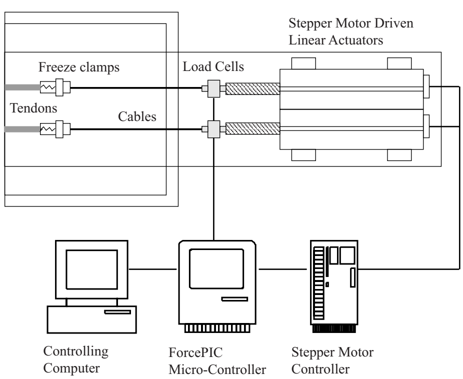


Figure 1. The components of Force Delivery Unit (DFU).

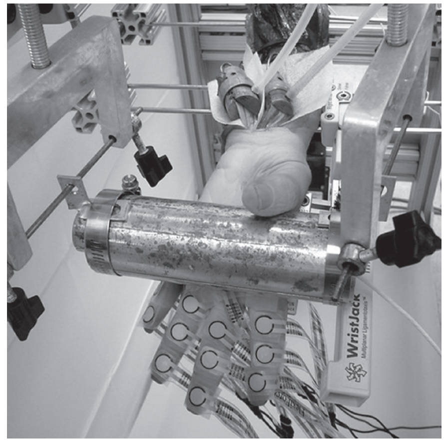


Figure 2. Specimen with FSRs attached on each phalange and mounted on a custom fixation device in the hand motion simulator with a cylindrical handle.

The Flexor Digitorum (FDP) and the Flexor Digitorum Superficialis (FDS) were the main finger flexor muscles and were selected to simulate the power grip motion. To maintain the structure of interior tissues including muscles, tendons and vascular system, only the FDP and FDS were extracted and isolated from the other muscles in the forearm and other tissues left undisturbed. The wrist fixator maintained a functionally neutral wrist angle of 20° extension allowing free motion for all fingers (Li, 2001). Each flexor tendon separated from flexor muscles was securely coupled with a freeze clamp (Sharkey, 1995) and coolant tubes for liquid nitrogen were connected to the freeze clamps.

2.3 *Tendon force control*

Total tendon force (sum of FDP and FDS force) was fixed at 200 N and both FDP and FDS forces were adjusted with regular ratios: FDP force decreased from 180 N to 100 N and FDS simultaneously increased from 100 N to 180 N. Consequently, the total force was always 200 N, but the tendon force ratio of FDP and FDS was different at the different combination of two tendon forces. Grip force generated by pulling internal flexor tendons was the dependent variable of each condition.

2.4 *Procedure and data analysis*

Three trials were run for each five different handles and five different tendon force ratios, with three minutes intervals between each trial. All of the 15 trials (5 tendon force ratios $\times$ 3 trials) were completely randomized for each handle. Two-way ANCOVA test was used to evaluate the effects of five tendon force ratios and five handle sizes in power grip motion. Significant effects were further explored using Tukey's pair-wise comparisons. The dependent variable was contact finger forces measured by force sensitive resisters on each phalange. The significance level was set as 0.05. Statistical analyses were performed in the statistics toolbox of Minitab 13.0 (Minitab Inc., State College, PA, USA).

3 RESULTS

Both the handle size and the relative FDS% showed significant influence on grip force ($p < 0.005$). Also, the interaction effect of the handle and the relative FDS% was also significant ($p < 0.05$).

3.1 *Tendon force ratio*

To explore the tendon force ratio of the FDP and FDS, the difference of finger forces were analyzed with the different tendon force ratios. The variation of each finger force summing phalange forces was significant with fingers and relative percentage of FDS force (%FDS). However, there was no interaction effect of finger and %FDS. As shown in the Figure 3, the mean finger force was significantly higher at the 40% FDS force ratio. In the pair-wise comparison, the mean finger force at 40% FDS increased significantly (48%) from the 10% and 20% FDS force level (Table 1). In terms of the efficiency between tendon force (internal force) vs. finger force (external force), the highest ratio of flexor tendon force to the applied finger grip force was 7.9 F generated by the 10% FDS force and the lowest ratio was 4.1 F with the 40% FDS force (3:2 PDF to FDS tendon force ratio). Thus, 4.1 F was most efficient ratio between internal force and external force, and the ratio of FDP and FDS to the applied finger grip force was 2.5 F and 1.6 F, respectively. The finger forces increased significantly on the index and middle finger as the%FDS force increased, but those dropped slightly at the 50% of FDS. Also, finger forces at the 40% of FDS force showed more contact forces than other tendon force ratios. Specifically, index and middle finger contact force increased significantly up to 40% FDS, but the dropped at 50% FDS.

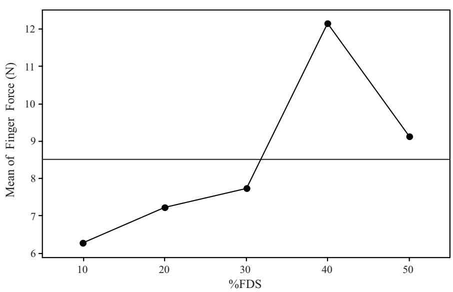


Figure 3. Main effects for the mean finger force.

3.2 *Finger joint angle*

In terms of the finger joint angle, the effect of handle, the relative percentage of FDS force (%FDS) and the interaction of handle and %FDS were all significant ($p < 0.005$). Figure 4 shows the captured grip postures representing finger joint angles in lateral view depending on different tendon force ratios and handle sizes. For the handle size effect, the joint angles at the MCP and PIP joint showed a descending pattern as the handle diameter increased, but DIP joint angle showed convex pattern according to handle size increasing. Essentially, flexions of MCP and PIP joint decreased as the handle size increased, but the DIP joint angle was most flexed for the middle handle size. The loading of the tendon forces with different ratio (FDP vs. FDS) produced different finger joint motions. According to increasing the percent FDS force from 10% to 50% of total tendon force, on average, the MCP joint angle gradually increased from 19.7° to 23.8°; in contrast, the PIP joint angle showed a rapid rise from 70.1° to 83.4° in the high proportion of FDS while the DIP joint angle displayed a rapid decline from 32.8° to 2.7°. These results clearly show that the tendon force ratio of FDP to FDS affects the

Table 1. The tota/individual finger forces and force distributions in responding to %FDS.

%FDS	Total finger force (N)	Tendon force	FDP force	FDS force	Mean contact forces of individual fingers (N) and distribution (%)			
					Index	Middle	Ring	Little
50%	36.6	5.5 F	2.7 F	2.7 F	11.1 (30.33%)	19.2 (52.46%)	4.3 (11.75%)	2.0 (5.46%)
40%	48.8	4.1 F	2.5 F	1.6 F	15.2 (31.15%)	24.4 (50.00%)	5.4 (11.07%)	3.8 (7.79%)
30%	31.0	6.5 F	4.5 F	1.9 F	8.6 (27.74%)	17.4 (56.13%)	3.1 (10.00%)	1.9 (6.13%)
20%	28.9	6.9 F	5.3 F	1.4 F	6.3 (21.80%)	18.2 (62.98%)	2.6 (9.00%)	1.8 (6.23%)
10%	25.1	7.9 F	7.2 F	0.8 F	3.7 (14.74%)	14.4 (57.37%)	5.2 (20.72%)	1.8 (7.17%)

(*F*: applied grip force unit).

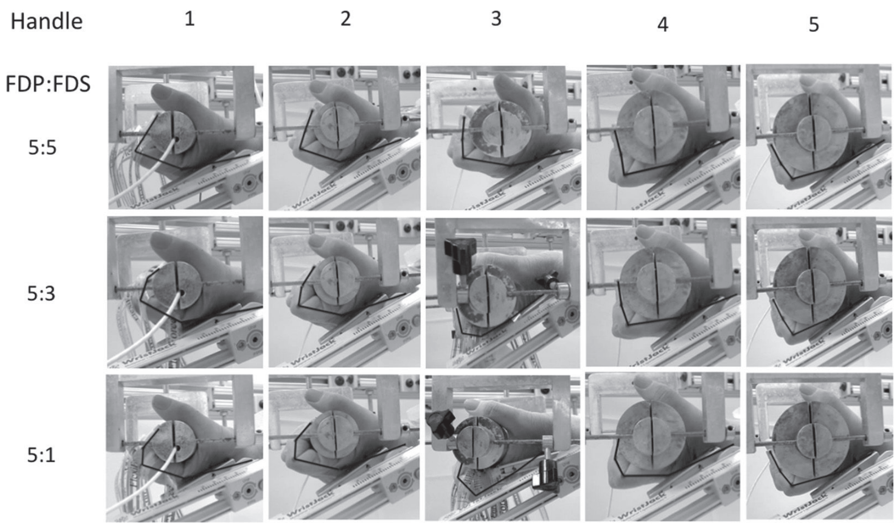


Figure 4. Finger joint flexion during grasping different sizes of handles with different flexor tendon force ratio.

finger joint motion while grasping a handle. Higher percent FDP force with less FDS force produced flat PIP joint angles and flexed DIP joint angles; In contrast, the same proportion of FDP to FDS force (i.e. PDF: PDS = 1: 1) produced flexed PIP joints and flattened DIP joints. Thus, the DIP would be too flexed with low portion of FDS force compared with FDP force, and would be too flat with same level of FDS to FDP force. Therefore, 3:2 FDP to FDS tendon ratio that already concluded before can be supported by this mechanism of finger joint angles by different tendon force ratios.

4 DISCUSSION

In this study, we demonstrated the finger joint motion induced by flexor tendons was closely related with a tendon force ratio of FDP to FDS. Also, the finger joint motion derived by different tendon force ratios affected the grip force generated by pulling flexor tendons in power grip motion. From our findings, we could design optimal handle size to minimize tendon forces and maximize grip forces. Furthermore, the results of this study provided novel insights into the kinematic role of individual flexor tendons.

4.1 *Tendon force ratio*

In terms of the tendon force ratio between FDP and FDS, the efficiency of external forces to internal forces was improved as the%FDS increased (10% to 40%), but it dropped at 50% FDS (i.e. FDP:FDS = 1:1). The efficiency at the 40% FDS force (i.e. FDP:FDS = 3:2) was significantly better (4.1 F) and it was worst for the 10% FDS force (7.9 F) in power grip. Consequently, the efficiency of the external to internal forces was optimal at 40% FDS force (3:2 FDP to FDS force ratio) with the smallest diameter handle. These findings can be compared with the results from previous studies. However, although researchers generally agreed on the proportion of muscle and tendon forces required for the externally applied forces, they presented various ranges of these ratios for the external force to internal tendon force. In the result of biomechanical model, Kong (2001) showed the average ratios were, for an external force of F, 9.05 F for FDP and 2.83 F for FDS in power grip motion. Brand et al. (1981), who analyzed relative tension and potential excursion of muscles in the forearm and hand by dividing the fiber length into the volume of each muscle (the cross-sectional area of the muscle) and relative tension capacities of forearm and hand muscle, found a 3:2 profundus-to-superficialis tendon-force ratio, which agreed with our results. Chao et al. (1989) also reported 3.17 and 1.51 for FDP and FDS in grasp motion, respectively. In terms of validation experiment, Schuind et al. (1992) reported that 7.92 for FDP and 1.73 for FDS in tip pinch motion. The variability of these results may be expected because all researchers did not use the same finger characteristics: moment arm, finger configurations, and angles of the applied forces to the fingertip regarding the function of intrinsic vs. extrinsic muscles during power grip motion.

4.2 *Finger joint angles*

The simulated contraction of extrinsic muscles generated concurrent flexion at the interphalangeal

(IP) joints, i.e. the joints were rotated in one direction only. High proportion of FDP force with less FDS force made PIP joint angle decrease and DIP joint angle increase. In contrast, the same proportion of FDP and FDS force, i.e. 1:1 FDP to FDS, produced flexed PIP joint and flatten DIP joint. Therefore, 3:2 FDP to FDS tendon force ratio showed most identical curvature motion. In terms of the metacarpophalangeal (MCP) joints, it was hard to control the MCP joint by pulling the extrinsic flexor muscles (FDP and FDS). The MCP joints of specimen were extended in rest condition due to high flexibility of the MCP joints. Thus, we had to hold up MCP joints until PIP and DIP joints were fully flexed during activating the grip motion. The role of the extrinsic finger flexor muscles (FDP and FDS) in initiating rotation of the MCP joint and in coordinating flexion at the MCP, PIP and DIP joints remains a matter of some debate (Kamper 2002). A previous study reported that the intrinsic muscles (Lumbrical and Ulnar/Radious interosseous) were seen as the primary MCP flexors, especially in regard to initiation of MCP flexion (Moore and Dalley, 1999). Also, a cadaver study utilizing static loading of the FDS tendon found that significant PIP flexion occurred before the loads become sufficient to initiate MCP flexion (Delattre et al. 1983). Thus, it was a limitation of the cadaver experiment that MCP joint could not be well controlled by pulling extrinsic flexor tendon forces. In addition, low contact forces of index, ring and little finger can be explained by that three fingers shorter than the middle finger may lose some of their mechanical advantage and thus unable to generate more pressure. Although the magnitude of applied contact force was quite lower than other study, the efficiency of internal force to external force was consistent with previous studies.

5 CONCLUSION

The study showed that the tendon force ratio affects grip force and finger joint angles on power grip. In other words, awkward finger motion and posture during grasping a handle may cause unbalanced tendon forces on flexors and degraded grip force. Grip design should be based on the maintenance of comfortable finger joint angles for optimum grasping and finger control.

REFERENCES

Armstrong, T.J. 1999. Cumulative trauma disorders of the hand and wrist: Ergonomics guide. *American Industrial Hygiene Association.*

Brand, P.W., Beach, R.B. & Thompson, D.E. 1981. Relative tension and potential excursion of muscles in the forearm and hand. *Journal of Hand Surgery*. 6: 209–218.

Chaffin, D.B. 1973. Localized muscle fatigue–definiton and measurement. *J. Occupational Med.* 15(4): 346–54.

Chao, E.Y., An, K.N., Cooney, W.P. & Linscheid, R.L. 1989. Biomechanics of the Hand: A Basic Research Study, *World Scientific, Singapore*.

Delattre, J.F., Ducasse, A., Flament, J.B. & Kenesi, C. 1983. The mechanical role of the digital fibrous sheath: application to reconstructive surgery of the flexor tendon. *Anatomia Chinica*. 5: 187–197.

Dickson, A., Petrie, A., Nicolle, F. & Calnan, J. 1972. A device for measuring the force of the digits of the hand. *Biomedical Engineering* July: 270–273.

Gilbert, B.C., Hahn, H.A., Gilmore, W.E. & Schurman, D.L. 1998. Thumbs up: Anthropometry of the first finger. *Human Factors* 20(6): 747–750.

Kamper, D.G., Hornby, T.H. & Rymer, W.Z. 2002. Extrinsic flexor muscles generate concurrent flexion of all three finger joints, *Journal of Biomechanics* 35: 1581–1589.

Kong, Y.K. 2001. Optimum design of handle shape through biomechanical modeling of hand tendon forces. *Ph.D. dissertation, University Park, PA: Pennsylvania State University.*

Li, Z.M., Zatsiorsky, V.M. & Latash, M.L. 2001. The effect of finger extensor mechanism on the flexor force during isometric tasks. *Journal of Biomechanics* 34: 1097–1102.

Moore, K.L. & Dalley II, A.F. 1999. Clinically Oriented Anatomy. *Lippincott Williams & Wilkins, Philadelphia, PA*.

NIOSH, 2001. National occupational research agenda for musculoskeletal disorders. *Cincinnati, OH, National Institute for Occupational Safety and Health*.

Oh, S. & Radwin, R.G. 1993. Pistol grip power tool handle and trigger size effects on grip exertions and operator preference. *Human Factors* 35(3): 551–569.

Ohtsuki, T. 1981. Inhibition of individual fingers during grip strength exertion. *Ergonomics* 24(1): 21–36.

Park, S., Freivalds, A., Sharkey, N. & Lowe, B. 2009. A new device for measuring flexor tendon forces and grip force: A cadaver model. *American Society of Biomechanics 2009 Annual Meeting, Penn State University, PA*.

Schuind, F., Garcia-Elias, M., Cooney III., W.P. & An, K.N. 1992. Flexor tendon forces: in vivo measurements. *Journal of Hand Surgery (American)* 17(2): 291–298.

Sharkey, N.A., Smith, T.S. & Lundmark, D.C. 1995. Freeze clamping musculo-tendinous junctions for in vitro simulation of joint mechanics. *Journal of biomechanics* 28(5): 631–635.

Silverstein, B.A., Fine, L.J. & Armstrong, T.J. 1987. Occupational factors and carpal tunnel syndrome. *American Journal of Industrial Medicine* 11: 343–358.

Part II: Usability & user experience

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A proposal of simple usability evaluation method and its application

Toshiki Yamaoka
Wakayama University, Wakayama, Japan

Satsuki Tukuda
Sharp Corporation, Japan

ABSTRACT: The purpose of paper is described how to evaluate designs using a Simple Usability evaluation Method (SUM). The SUM can evaluate GUI using the evaluation items of navigation, term and redundancy of information and other (operation time) which were extracted from 3P task analysis and ASQ. We compared the SUM with the protocol analysis. Hence the Sum got a lot of problems than the protocol analysis. So, the SUM is very convenient, easy to use and efficient evaluation tool for GUI mainly.

Keywords: evaluation method, usability, 3P task analysis, ASQ, SUM

1 INTRODUCTION

The purpose of paper is described how to evaluate designs using a simple usability evaluation method. Generally speaking, a lot of usability evaluation methods were proposed but engineers and designers were complicated to use them. They are designed for usability professionals.

2 THE OUTLINE OF THE SIMPLE USABILITY EVALUATION METHOD

Originally the Simple Usability evaluation Method (SUM) is an usability evaluation method for GUI mainly and constructed based on 3P (point) task analysis and ASQ (Lewis,1995).

3P task analysis and ASQ are as follows.

2.1 *3P (point) Task analysis method (Yamaoka, 2001)*

Consider a typical scenario where the product being surveyed will be used. Write down the tasks that will be performed in each scenario in the order that they come up. If a task is comprised of subtasks, then enter these subtasks in the task column as well. Write down real and anticipated user problems in the human information processing sequence consisting of effective acquirement of information → ease of understanding and judgment → comfortable operation. The cues when you are looking at problem areas are 1) poor layout, 2) difficulty of seeing, 3) no emphasis, 4) lack of information, 5) mapping for effective acquirement of information and then terminology, 1) indecipherable, 2) no affordance, 3) confusing, easy to mistake, 4) no feedback, 5) procedural problems, 6) inconsistency, 7) problems in mental model for ease of understanding and judgment and then

1. incongruity with humans' physical characteristics (posture, fitness and torque (force necessary for operation)),
2. cumbersomeness for comfortable operation. Finally consider proposals for resolving the real and anticipated problems that were extracted.

It is a good idea to refer to the resolution proposals when they come up because they present opportunities for working out solutions. A rough design proposal can be developed by condensing the resolution proposals.

2.2 *ASQ (the After-Scenario Questionnaire) (Lewis, 1995)*

ASQ (the After-Scenario Questionnaire) is a three-item questionnaire which is as follows:

1. Ease of task completion
2. Time to complete a task
3. Adequacy of support information (on-line help, messages and so on)

They ask participants user satisfaction with system usability. The questionnaire is very short, easy and effective for engineers and designers.

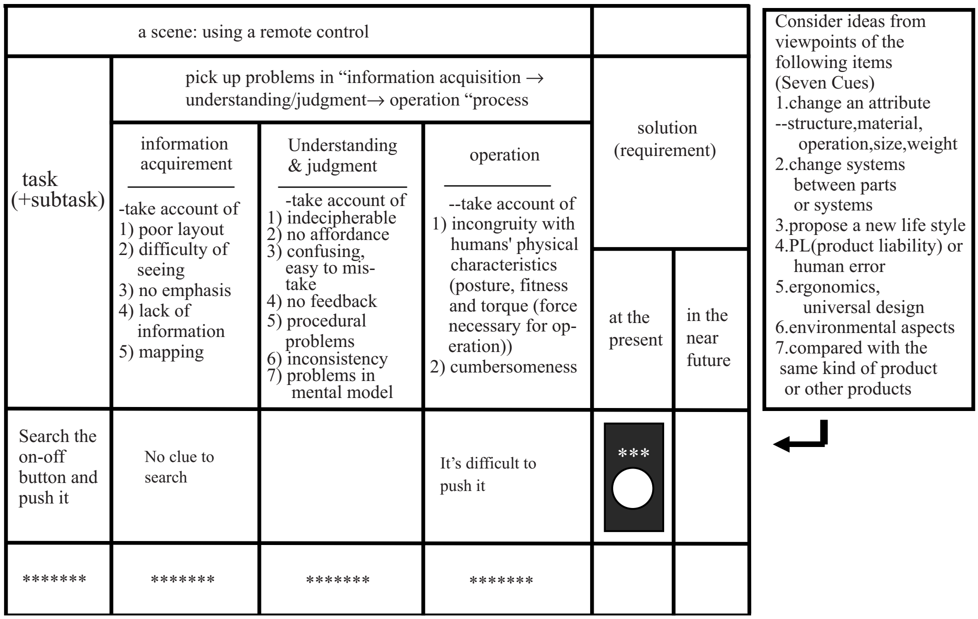


Figure 1. 3Point task Analysis (yamaoka, 2001).

2.3 *SUM (the simple usability evaluation method)*

Navigation (included function model and structured model in mental model), term and redundancy of information in the SUM is extracted from 3P task analysis and ASQ (Fig. 2). They are very important for usability evaluation. Most of problems are caused by navigation (relationship among operation parts), terms and mental model (functional model and structural model) mainly. Functional model means a kind of procedure, namely "How to use it". Structural model means "How it works" and shows a structure of system. Others (operation time) are evaluation item in order to extract other problems. If users take a lot of times to operate in a task and no problems are found in navigation, terms and redundancy of information, other problems seem to be included in the task. We can use the operation time as a clue in order to extract usability problems.

Tasks are evaluated from viewpoint of navigation, terms and redundancy of information mainly.

Criteria of the evaluation are as follows (Table1).

1. When users can go to next task in spite of one problem, the level of the problem is defined as bad–1. If a task has one problem (bad–1) based on three criteria (navigation, term and redundancy, others), the evaluation score is 0.
2. If a task has two problems (bad–1) based on three criteria (navigation, term and redundancy, others), the evaluation score is–1.
3. When users cannot go to next task by one problem, the level of the problem is defined as bad–2. If a task has one problem (bad–2) based on three criteria (navigation, term and redundancy, others), the evaluation score is–1.
4. If a task has no problem based on three criteria (navigation, term and redundancy, others), the evaluation score is +1.
5. If users take a lot of time to operate in spite of bad–1, the bad–1 becomes bad–2.

Although the length of the time depends on tasks, permitted time length is usually within 60 or 120 seconds.

2.4 *How to employ the SUM*

The procedure to employ it is as follows.

1. The scene and the tasks are identified.
2. The task is analyzed based on navigation, terms and redundancy of information. Problems are extracted.
3. The bad–1 or 2 of problems are decided by the judgment of whether user can go to next task or not.
4. The evaluation score is decided based on the number of bad–1 and bad–2.
5. The synthetic evaluation score is calculated by the total of tasks score. As the numbers of

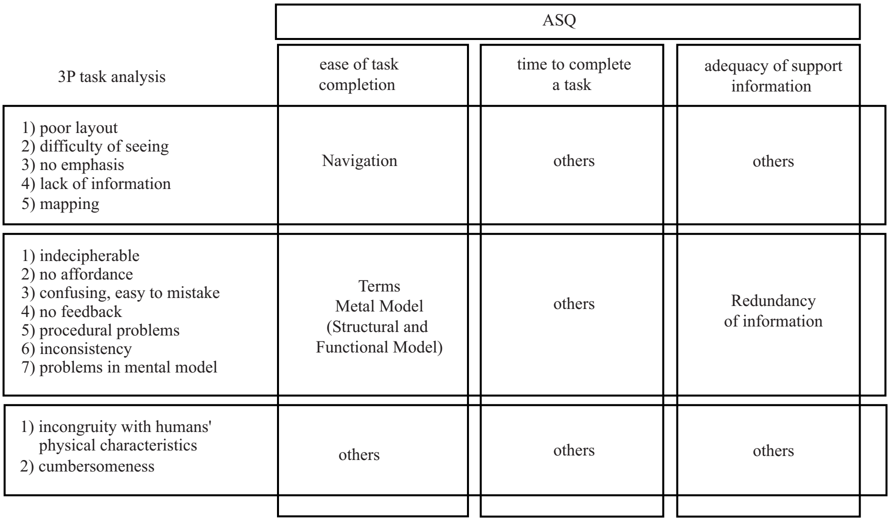


Figure 2. Relationship between 3P task analysis and ASQ.

Table 1. The Simple Usability evaluation Method (SUM).

Task (subtask)	Navigation / Functional and Structural Model	Terms and Redundancy of information	Others (operation time)	Evaluation (+1,0-1)
Task(1)	bad-1 (can go next)	good	good	0
Task (2)	bad -2 (cannot go next)	good	good	-1
Task (3)	good	bad-1	bad-1	-1
Task (4)	good	good	good	+1
Task (5)	good	good	bad-1	0
Task (n)	good	good	bad-1	0

bad–1or 2 regarding the three evaluation items (of navigation, terms and redundancy of information) can be counted, we can understand the bad evaluation item.

6. The problems as to bad–2 are examined. Hence the important requirements are extracted.

3 EVALUATED USING THE SUM AND COMPARED WITH PROTOCOL ANALYSIS

3.1 *An example evaluated using the SUM*

A WEB site in Japan was evaluated by the SUM (Table 2). As the evaluation items such as navigation, term and redundancy, others are clear, it was easy to extract problems on the screen.

3.2 *Comparison between the SUM and protocol analysis*

1. Participants 3 university students (male, 19–22 years old)
2. 3 tasks of retrieval was done
3. Results (Table 3)

The SUM can get a lot of problems than the protocol analysis. The results show total 37 problems as the SUM and 25 problems as the protocol analysis.

Table 2. A WEB site in Japan was analyzed by the SUM.

Task (subtask)	Navigation Functional and Structural Model	Terms and Redundancy of information	Others (operation time)	Evaluation (+1,0-1)
Task (1) Search the specified term to go to the next task.	bad-1 no navigation to guide on the screen	good	good	0
Task (2) Push the button	good	good	bad-1 It is hard to see characters because of low contrast	0
Task (3) Search the specified term to go to the next task.	bad-1 no navigation to guide on the screen. Disturb the searching by lots of information	good	good	0

Table 3. Comparison between the SUM and protocol analysis.

	The number of problems by the SUM	The number of problems by the protocol analysis
Participant 1	Task 1: 7 Task 2: 5 Task 3: 5	Task 1: 3 Task 2: 3 Task 3: 3
Participant 2	Task 1: 3 Task 2: 2 Task 3: 8	Task 1: 2 Task 2: 0 Task 3: 5
Participant 3	Task 1: 2 Task 2: 1 Task 3: 3	Task 1: 5 Task 2: 3 Task 3: 1
Total	Task 1: 12 Task 2: 8 Task 3: 17	Task 1: 10 Task 2: 6 Task 3: 9

REFERENCES

Lewis, J.R. 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human–Computer Interaction* 7 (1): 57–78.

Yamaoka, T. 2001. Human design technology as a new product design method. In *First international conference on planning and design.* Taipei, CD JP003-F 01–10.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Usability evaluation of an interactive service on mobile phone

Karl W. Sandberg & Håkan Andersson
Mid Sweden University, Sundsvall, Sweden

ABSTRACT: Social networking services have become increasingly widespread and, together with GPS equipped with mobile phones give citizens the opportunity to create new innovative types of social networks. The service is evolving, technology-centric to positioning using GPS and Google Maps Web service should also identify existing services interfaces are analysed, and be associated with basic functions of social networks to create a new innovative service. The purpose of this work is to illustrate the user-centred interaction design for mobile phones. Research design includes user requirement analysis, prototype development and usability evaluation. The results from usability evaluation in a simulated PC-based environment, support the user tasks in an efficient manner, one can work efficiently and feels that the system is sufficiently adapted to users' needs.

Keywords: mobile phone, prototype, interaction design, interactive service, user interface, usability

1 INTRODUCTION

The mobile phone has become a natural part of everyday life for over two billion people. Mobile devices are often exposed to statements made by selling the mobile devices to help us communicate better, make us become more creative and more efficient in our work. In many, cases of new development are met but not expectations, as complex and unusable interface makes the services and functionality in small mobile phones and PDAs unusable because of poor usability (Jones & Marsden 2006).

User interface and opportunities for interaction with mobile devices is more demanding for developers with limited screen size, small buttons and lack of standardized interaction tool makes it more difficult for a user to interact with this device than with an ordinary desktop or laptop computer. Rapid technological development and a host of new services have been made to mobile applications and systems are often oriented around technology and technology development as opposed to being user-centred and designed for high usability (Jones & Marsden, 2006; Sandberg & Kanter, 2009).

Some mobile phone services offered by telephone operators have become very popular, such as SMS and MMS, while other types of services that software-based Web applications and games also are becoming more common. The rapid evolution of technology and the large number of brands and different techniques have made it even today there is no accepted standard for how mobile devices must be designed and there is often overlap of functionality between mobile phones, PDAs, MP3 players and similar devices. This was not the case until the millennium when the purpose of the various mobile units was more distinct: mobile telephones to handle communications and handheld computers to manage information. Today the border erased something and it is a popular topic with many different views whether such mobile phones is designed to be simple devices for communication or evolve to become complex multifunctional units (Jones & Marsden 2006).

Mobile devices come in many sizes, from handheld devices to tiny sensors. Most mobile devices have buttons that you can press and many units are also touch-sensitive displays. There is usually some form of wheel or arrow keys to highlight and select information and interaction options on the display. The aesthetic design of mobile devices has also become increasingly important, and today is the mobile phone number for an accessory like a wallet or day planner (Jones & Marsden 2006).

A service that is very popular in the current situation is Facebook (2009), which is a social's network where users can submit personal information, make contacts, and send messages and emails between each other (Facebook 2009). A service that has become increasingly common in mobile phones is support for GPS and can play to their geographical location on a map using services like Google Maps (2009). The development of wireless sensors can also quicken up the ability to integrate the mobile phone world with wireless sensor networks and offer new types of service and technical solutions.

1.1 *Research design*

A prototype of a user interface for a social network will be developed, which has support for mapping using GPS, and the ability to use wireless sensor networks. The interface should be user-centred and focus on usability, and also aesthetically pleasing. The interface is developed based on recognized methods and standards, and then evaluated in a small test group with the usability testing. The project will serve as a prototype for an interface to a social network can facilitate communication with other people, and the opportunity to interact with their environment by linking a social network with new technologies such as GPS-based map services in mobile phone which allows for new innovative services.

1.2 *Method*

Usability testing and user interviews delimited to include only a small group of five people who fit the described user profile. Testing restricted further to testing in the PC environment where the user interacts with an illustrated mobile phone. Testing is also limited to a particular type of mobile phone with a given screen resolution (Zwick, Schmitz, & Kuhl, 2005).

2 INTERACTION OPPORTUNITIES ON MOBILE PHONE

Mobile phones usually have very limited interaction possibilities, but these vary somewhat between different models and brands. There is however an ISO standard for the keys: 12 keys, 10 for numbers 0–9 and the other two of the characters * and #. Eight of the numeric keys also have three letters associated with them. Writing more than a few words with such a keyboard can be frustrating where the standard input is pressing repeated times on the same button to select the correct character. Studies have shown that such an input method, limiting entry to 21 words per minute by experts in comparison with 60 words per minute for normal users on a regular keyboard standard for PCs. The input system for mobile devices has improved example, using T9 technology from Tegic, which uses a vocabulary that enables the mobile phone may guess the right word, a number of key-press and look up potential words in the dictionary to limit the selection (Kinnander 1997).

Another alternative is the so-called “Fastap” keyboard, which consists of a small keyboard with 50 different keys in the order of an ordinary credit card Fastap (2006). Since the keys are so small, the keyboard is used with the aid of an algorithm that determines which keys the user probably intended for multiple keys are pressed accidentally. Although a lot of other less common and successful keyboard solutions have been in the mobile industry.

Techniques which have become increasingly common in recent years are touch-sensitive screens that allow the user to interact directly by pointing and tapping the screen. Solutions with the gyrosensor, user can lean your mobile phone to interact with the information contents of the display. Voice control-solutions are also available, but there are large problems with those due to noise and interference, and that it is difficult to link voice to a certain context. Another problem with voice recognition is that they rely on word lists and tables which are causing problems on the word used are not recognized (Jones & Marsden 2006).

2.1 *Usability*

Successful mobile products and services are the products and services that are useful and provide the user with a consistent, rich experience when it is used. If a mobile service is not highly valued offers little functionality will not be used no matter how well designed user interface is. On the other hand, a bad user interfaces to do a service still in use. For example, use SMS service in a very large extent, although it is very cumbersome to write text messages in a mobile phone (Jones & Marsden 2006, Sandberg & Pan 2009).

Useful products and services. A mobile device and its utility services is affected by two factors: how easy it is to use, i.e. how it presents its function and how well it fits as a resource in the user’s everyday life. Low usefulness appears in two guises. First is the question of the device or service is useful in itself. Secondly, it is about the unit works in harmony with the things that surround it. Low utility can prove when the developer of such a mobile service has not thought of how well the service interacts with a user’s other resources and can enrich their lives (Jones & Marsden 2006, Sandberg & Pan 2009).

Guidelines and standards. Usability by itself is largely about how easily the user can translate an imaginary target for interaction with the device or service used. If such a person turns the sound on his mobile phone on during a meeting, the user must translate the objective: “to turn the sound off” commands to the system in their mobile phone. At the same time, the user based on the interaction he or she had with mobile phone is able to see what their actions achieved, i.e. how the system has changed.

Four standards for system designers to take into account are designed for high usability:

- Ensuring a high degree of visibility—Allows the user to determine the current system state and what options may be available.
- Give feedback—Provide coherent, transparent information on the results of any election by the user.
- Presenting a conceptual model good—Give the user the opportunity to build a true picture of how the system fit together, the relationship between different parts of the system and how the transition between each state takes place.
- Offer good mapping—Strive for clear, natural relationships between the choices made by the user and the results they yield.

In addition to the above standards are a range of advice and guidance to achieve high usability impaired. Broadly speaking, these deals to work on consistency, enables the regular user to use shortcuts, offer informative feedback, and endeavour to prevent errors and help the user to quickly recover from them.

2.2 *Interaction design*

Interaction design is about involving people in the systems development process to understand their needs, develop prototypes based on accepted guidelines and case studies, and refine them based on the results given in the evaluation. Interaction design follows an iterative development model therefore requires that the above description of the design process is repeated several times to achieve the best possible outcome (Preece 2002).

The goal of interaction design is to develop useful products, that the product or service is easy to learn, efficient to use and user satisfaction. User is involved in the design process. The concept of human-computer interaction affects the design, development and implementation of interactive computer systems to humans by studying the environment surrounding people. Interaction creates a drawing of the system that specifies the user's needs in terms of:

- Required functionality.
- How access to this functionality is given.
- How content is presented.
- The system's different states.
- Help and feedback information.
- How the system interacts with other resources in the user's context.

2.3 *User Centred Design*

User Centred Design (UCD) is an approach that supports the entire development process with user-centred activities, in order to create applications which are easy to use and are of added value to the intended users (Maguire 2001). It is a philosophy that places the person at the centre; it is a process that focuses on cognitive factors (such as perception, memory, learning, problem-solving, etc.) as they come into play during peoples' interactions with things. UCD seeks to answer questions about users and their tasks and goals, and then use the findings to drive development and design.

3 METHOD

By evaluating existing social networking sites and functionality, and do a needs analysis with a group of potential users, the most desirable features highlighted and possible standards in terms of user interface are analyzed to create an analytical foundation to function and interface design for a mobile based service. Because the service is evolving, technology-centric to positioning using GPS and Google Maps Web service should also identify existing services interfaces are analyzed in details, and be associated with basic functions of social networks to create a new innovative service.

By using theories of interaction design for mobile devices, and user-centred method based on innovation around existing successful services to create mobile applications prototype of a user interface for mobile devices as possible, make a simple social network in conjunction with a positioning service and sensor. Interviews of users who already use social networks on the web must provide a basis for the functionality that is desired by a needs analysis. This can also help with guidance on how a user interface can be created as easily open up new communication channels and means of communication to the user.

A usability test for the most requested functionality must be made and completed with a modified WAMMI questionnaires related to the usability test (WAMMI 2000). WAMMI measured users perception; understand ability (that bear on the users' effort for recognising the logical concept and its applicability); learnability (that bear on the users' effort for learning the application); operability (that bear on the users' effort for operations and operation control); attractiveness (the capability to be liked by the user).

By recording the user's behaviour can be observed and user behaviour analysis. Users' perform together with WAMMI results are then used to identify-your shortcomings in the current prototype and possible improvements to be highlighted. The project is in other words, be implemented as a first step in the iterative interaction model. Usability testing should be done by a PC environment where the user interacts with an illustrated mobile phone by clicking the buttons on it. By "record" the mouse cursor moves and clicks the user's behaviour can be analyzed at a later stage.

The user interface is developed to take into account users' preferences, the guidelines and standards that exist for the design of user interfaces, and psychological factors. The test version of the user interface is implemented as a simple interactive "picture-games" in the application environment, where the user navigates in the mobile prototype by clicking on an illustrated mobile phone's buttons. The mobile as the application is tested on a Nokia 6200 Classic, then this is something of an industry standard in terms of keypad and the most common screen resolution (144×190) on the market in current situation.

3.1 *Prototype of user interface*

The product to be developed is a social networking service that is able to communicate with others as part of a social network. The product has a vision to enrich the user's everyday life, and help to connect and communicate with friends and acquaintances. The product will also make the user more efficient in their daily lives by making it possible to quickly find out where various people are so that you can quickly find these people, or coordinate the work.

The first prototype of the service must include a basic level of functionality has been identified as the most important needs-based analysis:

- Navigable map function.
- Ability to show different groups/networks of people.
- Ability to contact them via message/SMS/MMS.
- Ability to see where a certain person is on the map.
- Ability to see their own location.

Since the service being developed is a combination of a social network and a GPS-based map service, it is important to weave together a map with easy ways to communicate. Mobile phones used as a development platform is a Nokia 6200. The assumption is logically reasonable to expect the user to "scroll" the map image using this governing cross so therefore the prototype has been built according to this principle. When the application starts, the user's current location, which is illustrated with a blue sphere, is linked to the user with lines. This makes it easier for the user to navigate to them (see Figure 1).

The phone has two blue buttons in the election-related display and uses them to make choices based on what is displayed in a menu in the bottom of the display in relation to these keys. By default, this prototype allows the user to select the "show"

Figure 1. Left: User's own position is illustrated by a blue sphere and the lines drawn between the user and the people who are in the user's surrounding area Right: When a person is selected, this is illustrated with an orange arrow pointing at the person.

the left button to display a menu where the user can choose which group or any network of people to be shown on the map. By scrolling up and down on the navigation cross, the user can choose from the options and confirm their choice in the left blue button, or go back with a right hand. "Find" on the main menu displays a menu that lists the names of all the persons included in the network selected via the "show". By selecting a name from the list displays the selected person directly on the map and marked as indicated by an orange arrow pointing at the person. Orange is also used as marker colour in menus to provide a consistent impression (see Figure 1).

3.2 *Usability evaluation*

In order to develop a prototype for usability testing of the mobile application, are developed as a PC application in Microsoft C #, NETs. Application illustrated a mobile phone displayed on the screen, where the user can interact with the mobile phone by clicking the phone's buttons with the mouse. By using a simulated application it is easy to make changes and user actions can be recorded by trace program and play mouse navigations on the screen and how the user clicks on the buttons on the mobile phone.

The defined tasks user solves during the usability test are the following:

1. Find out where your friend Lene Gadd 'lies.
2. To do that only your co-workers shown on the map.

3. Send a “message” to any person within the network “My University”.
4. Find out where your friend “Patrick Kruz ‘lies”.
5. Send an MMS to “Gote Von Grape”.

After the usability test, the user has to fill in the WAMMI questioner.

4 RESULTS

Usability testing of a prototype application was done on computer. The user’s behaviour was recorded to make it possible to reproduce the analysis and this presents an objective the description of tests, their output in terms of time to perform each task, observations concerning users’ behaviour when using the prototype application, and calculated WAMMI evaluation results.

4.1 *Usability test*

A clear pattern of behaviour that arise out of the fact that the prototype tested in a PC environment instead of a real cell phone is that users try to click directly through the menus on your mobile phone display instead of using the keypad on your mobile phone.

Several users tried to confirm their choice by pressing the centre of the steering cross instead of choosing the ‘selection’ button in the upper left corner. The procedure was repeated several times despite the fact that the choices made by “choice” button before.

There was only one user at all in trying to locate a person manually “scroll” across the map with the control cross. All others used the “find” function to find out where people were.

From pure observation, it seems that the link between that mark a person via the “find” function and that this in turn makes “choices” menu is visible is unclear.

After performing three tasks showed all users sign on to have memorized how to use the group filter functionality via the “view” menu in combination with “find” menu to quickly find out where a person is and communicate with this . This meant that for most users went quickly to resolve the task four and five.

Temporal data based on usability test performance are presented in Table 1.

4.2 *WAMMI*

A modified version of WAMMI questionnaire (WAMMI 2000) was used to evaluate the prototype application. Figure 2 shows the results of the WAMMI questionnaire (n = 6). In general one can

Table 1. Time per task (minute: seconds).

Subj.	Task 1	Task 2	Task 3	Task 4	Task 5
1	0:36	0:30	3:03	0:34	0:26
2	1:06	0:18	1:23	0:26	0:18
3	0:39	0:12	0:45	0:48	0:29
4	0:13	0:14	0:26	0:17	0:22
5	0:29	0:39	1:32	0:22	0:13
6	0:27	0:40	1:31	0:24	0:25
Mean	0:35	0:25	1:26	0:28	0:22

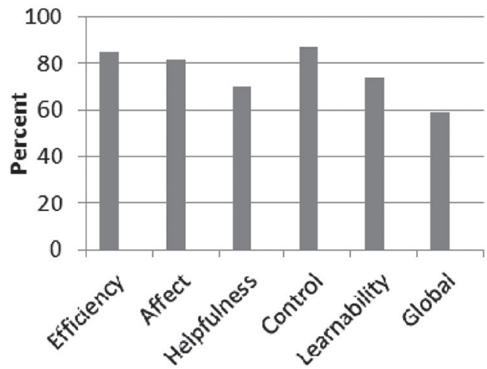


Figure 2. WAMMI score.

say that the user satisfaction regarding the user interface is high. The system does support the user tasks in an efficient manner, one can work efficiently and feels that the system is sufficiently adapted to users’ needs, and users find it easy to let the system behave in the way they want it to.

5 ANALYSIS

Performing usability testing of a mobile application in a simulated PC environment has proven to be a practical approach from the developer’s site, but confusing to users when they unconsciously try to use the application as a windows-based PC application initially. E.g. users tried to click the mouse in the phone’s menus. This should be considered as an indication that the PC-based usability testing is not suitable for mobile applications. It is probably more appropriate to film or directly observe the user’s behaviour while the latter uses a prototype of an ordinary mobile phone. Such an approach is however time-consuming and requires a lot of resources from the developer’s side. Interview group of eight people and test group of six people had it been possible, expanded to include more people, which gave a more credible results, but given the circumstances and the resources that were available is considered the number of people involved in development work still had a significant role.

Despite high subjective values of usability in WAMMI, it is clear that the helpfulness of the prototype application is the main fault. This fact is not surprising since the application does not contain all kind of help information. By constructing the menus in a manner known from conventional mobile systems, the user started quickly and that must be successful in the prototype on which control option received very high ratings from users. The current prototype should be completed with a help function, which tentatively is accessible from any menu or a button.

Users report that the application to make it effective and feel interface and functionality are attractive. However, there is more to be desired by learnability when several of the problems users had to repeat a behaviour which they performed earlier in the application.

In addition, the new functionality is implemented in the next prototype, which makes it possible to really communicate with other users through the creation of SMS, MMS, messages, etc. They should also be able to create new networks, joining the existing network and build relationships with other users, which place even more stringent requirements on the application's interface and design. Button functions should be reviewed when their meaning did not appear to provide a consistent and logical impression of observations to judge. Possibly the middle of the steering cross used as a selection button instead of the blue button at top left. The graphics illustrating that a user is selected is not sufficiently communicative and should also be reviewed. Possibly the "find" changed to "select" to make this clearer to the user.

Interaction and standardised methods have proved to be a powerful tool for building operation and interfaces based on the user's wishes and give indications of weaknesses in the software at a very early stage. This makes the interaction probably provides better results than traditional development methods.

REFERENCES

Benyon, D., Turner, P. & Turner, S. 2005. *Designing interactive systems – people, activities, contexts, technologies*. Harlow: Pearson Education Limited.

Facebook. 2009. http://www.facebook.com

Fastap 2006. Fastap redefines the mobile keypad *Mobile Europe - Europe's Wireless Magazine*, 19 July.

Google Maps 2009. http://www.maps.google.com

ISO/IEC 9241-11. 1998. *Ergonomic requirements for office work with visual display terminal – Part 11: guidance on usability*. International Organization for Standardization.

Jokela, T., Iivari, N., Matero, J. & Karukka, M. 2003. The standard of user-centered design and the standard definition of usability: Analyzing ISO 13407 against ISO 9241-11. *Proceedings of the Latin American conference on Human-Computer Interaction* (pp. 53–60). New York: ACM Press.

Jones, M. & Marsden, G. 2006. *Mobile Interaction Design*. Glasgow: John Wiley & Sons Ltd. England.

Kangas, E. & Kinnunen, T. 2005. User-centred design to mobile application development. *Communication of the ACM* 48(7): 55–59.

Kinnander, O., Bouvet, S. & Sukow, R. 1997. Handset e-mail made easier; Tegic's T9 technology cuts down on keystrokes. *Mobile Phone News*, Sept 1.

Klockar, T., Carr, D.A., Hedman, A., Johansson, T. & Bengtsson, F. 2003. Usability of mobile phones. In *Proceedings of the 19th International Symposium on Human Factors in Telecommunication*, Berlin, Germany, 197–204.

Maguire, M. 2001. Methods to support human-centred design. *Int. J. Human-Computer Studies* 55: 587–634.

Preece, J., Rogers, Y. & Sharp, H. 2002. Interaction design: beyond human-computer interaction. John Wiley.

Preece, J., Rogers, Y. & Sharp, H. 2002. *Interaction design*. New York: Wiley.

Sandberg, K.W., Kantor, T. & Pan, Y. 2009. Mobile technology as part of the research agenda on democracy. *The 13th International conference on Human-Computer Interaction*, 19–24 July 2009, San Diego, CA, USA.

Sandberg, K.W. & Pan, Y. 2009. Human Factors in Public Information Systems. (Eds.) Panayiotis Zaphiris and Chee Siang Ang. In *Human computer interaction: Concepts, methodologies, tools, and applications vol. 4*, Information Science Reference.

Sandberg, K.W. & Sundberg, H. 2009. Mobile citizen-friendly services – Two case studies in rural areas. *IADIS, International Conference: e-Society,* 25–28 February, Barcelona, Spain.

WAMMI. 2000. *Web Usability Questionnaire.* http://www.wammi.com

Zwick, C., Schmitz, B. & Kuhl, K. 2005. *Designing for Small Screens.* Ava Publishing.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluation user interface of VDU in a process control room

Karl W. Sandberg
Mid Sweden University, Sundsvall, Sweden

Yan Pan
MTO-kompetens, Sundsvall, Sweden

ABSTRACT: Varied types of Visual Display Unit (VDU) are used in modern process control rooms. The main purpose of present study is to identify and find solutions to usability problems of VDU user interface in process control room. The usability evaluation methods were heuristic evaluation, usability test, and SUMI questionnaire. Fifteen operators participated in the study. The main result from the usability test were that the time to solve the task depends on the number of false clicks and commands, and participants experienced that user interface doesn't support in an efficient manner, too many and difficult steps must be performed. To improve the user interface of VDU we have created number of guidelines; for example; create groups of the fields, buttons, text, place the buttons in functionally way, enable the user to open multiple windows, drag and drop is a dynamic operation.

Keywords: VDU, user interface, usability evaluation

1 INTRODUCTION

The process controls using computers have replaced the human machine interaction, the operator's task has become different (i.e., controlling and monitoring through the computer user interface such as keyboards and visual display units (VDUs). The hardware equipment used for presenting process data and controlling the process has been replaced with displays and controls on the VDU (Van der Schaaf 1989, Vincente 2001). This makes the user interface of the computer-based controlling system one of the most important factors for improving the task efficiency and preventing safety-related accidents (Wang & Hwang 1995, Johannsen 1997, Kawai 1997, Hix 1993, ISO 1997a, b). The simple miniaturization could increase familiarity at the cost of more complex screens because much more data and controls should be presented on a limited screen space. As a result, serious usability problems could occur, and consequently, the efficiency and safety of process control tasks could suffer as well. A several small changes of the user interface of VDU might cause confusion and might lead to malfunctioning or safety-related accidents. For that reason the designers focused on improving the user interfaces following design rules and guidelines. To avoid this, a systematic method to improve the user interface of a process control room.

The purpose of present study is to perform a usability study of alarm and case management system and drew up guidelines based on evaluation results and proposes improvements of user interface of the system.

2 METHODOLOGY

An operations control centre monitored mobile traffic in Sweden. They are responsible for guarantee that relevant operational information is reported, that fault in the mobile net will be corrected as soon as possible. The alarms must be technically analysed within a certain time interval and corrective maintenance should be started. The operator is monitored alarm and a computer supported Case Management System (CMS) using multiple VDUs.

2.1 *User characteristics*

The operators have a long experience of using the user interface to alarm and case management system, an average over three years.

2.2 *Evaluation methods*

Heuristic evaluation. The heuristic evaluation is that a group of evaluators examine an interface and determine whether it follows some usability principles, known heuristic. It has become apparent that different evaluators will find different kinds of problems and therefore should be more people perform the evaluation. The evaluators have a knowledge and understanding of usability principles (Nielson 1994).

Usability test is based on the user from the intended user can use and perform typical tasks of the user interface of VDU (Rubin 1994).

SUMI questionnaire. To measure user satisfaction, and hence assess user perceived software quality, University College Cork has developed the Software Usability Measurement Inventory (SUMI). SUMI provides an Overall Assessment and a Usability Profile which breaks the overall assessment down into five sub-scales: Affect, Efficiency, Helpfulness, Control, and Learnability (Kirakowski 1996).

3 RESULTS AND ANALYSIS

3.1 *User background*

Five women and 15 men (n = 20) participated in the study, where 13 falls in the 25–44 years old, and other in 45–65 years old.

3.2 *Heuristic evaluation*

The main qualitative results of heuristic evaluation of usability problems with user interface of VDUs to alarm and CMS are summarised (n = 16) in Table 1.

3.3 *Usability testing*

Usability test was carried out directly in the workplace, using mobile usability equipments. Two real tasks was simulated and performed by eight participates. Average times for user testing are presented in Table 2, and number of errors during the user tests (false clicks and commands), are presented in Table 3.

The most striking result revealed by analysis of the results from the usability test is: The time to solve the tasks depends on the number of false clicks and commands, i.e. the more false clicks and commands; the longer it takes to complete the task.

3.4 *SUMI questionnaire*

Figure 1 shows the results of the SUMI questionnaire (n = 16). In general one can say that the user satisfaction regarding the user interface is too low and corrective action is needed. Some more detailed conclusion was: the system doesn't support the user tasks in an efficient manner, many and too difficult steps must be performed. Consequently, one cannot work efficiently and feels that the system is insufficiently adapted to users' needs, the messages are often not clear and understandable to provide help in solving a problem, and users find it difficult to let the system behave in the way they want it to.

4 IMPROVEMENT OF USER INTERFACE

Based on different methods to measures the usability; heuristic evaluation, usability testing, and SUMI questionner of the alarm and case

Table 1. The heuristics problems.

Consistency	Difficult navigation forward and backward arrow. Logical order is not fulfilled.
Display issues	Boundaries among information and interface elements on a screen are ambiguous
Memory issues	There are too many abbreviations and acronyms
Simplicity	A screen includes too many interface elements
User control	Too many "clicks" to create a case. Involved nodes should be included in the "scope" of the alarms. Some take-offs and tests can be automated. Too much information on the screen. High Lightning of important information. Some status information is not clear.
Clearly marked exits	How the user goes through the system? With the "Next" button and the information presented continuously on the screen so there is no problem with the bearings in the system. Simplify navigation with "Previous"—and "Next" buttons.
Shortcuts—shortcuts	Too few shortcuts.
Error messages	Some messages give mixed messages and unclear. Present the error messages displayed on the screen, in easily understandable terms.
Help and documentation	The system should provide help when there is no other option.

Table 2. Mean times.

Task	Mean	Standard deviation
1	25.72 min	5.05 min
2	16.31 min	5.25 min

Table 3. Number of false clicks and commands.

Task	Mean	Standard deviation
1	15.21	2.05
2	10.23	1.67

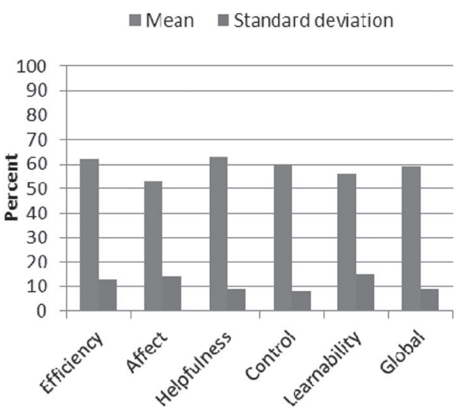


Figure 1. SUMI scores.

and management system, and using guidelines to redesign user interface of VDU. In this part we will present a solution of usability problems of user interface of VDU.

In order to find a good logistic, the VDU was divided in three process fields (see Figure 2):

1. Start-up
2. Operation
3. Final

4.1 *Solution of usability problems*

Some of the major changes that have to be done to improve the user interface of VDU of the alarm and case and management system in the process control room are presented in Table 4.

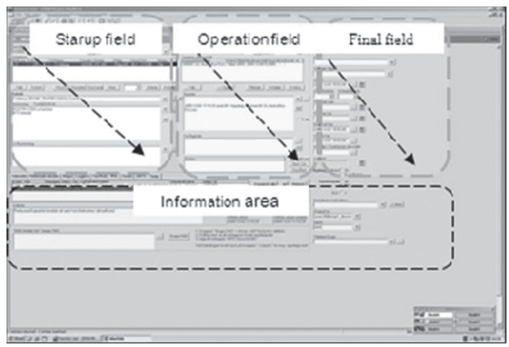


Figure 2. Organisation of user interface of VDU.

Table 4. Summary of problems and solutions of user interface.

Problems	Solutions and motivations
Grouping	Box structured into distinct groups is easier to understand than if they spread evenly across the screen. Create groups of the fields, buttons, text, etc. that belong together in terms of importance. Humans rapidly and automatically create groups of things that are close to each other. If the user can create categories of information on screen, she can benefit from the image quicker and work more efficiently.
Location and consistence	Place the buttons so that the same functionality found in the same place in all screenshots. Humans are very good to perceive small differences.
The first field in the tab order	When you open a form that can be updated in order to update only the first field should be in focus so the user can immediately complete the form. The user can concentrate on the task rather than looking for where she should start Automate the procedure so the user can directly take up its task.
Toggle images on the screen	If it is necessary to switch the image, ensure that the user has a good overview of where they are and how they can move on. The user loses easily removed at the picture changes. The user is in when questions like “Where am I?”, “How did I get here?”, “Why did I get here?” The user experience, however, that she is in control if she has a clear view and approach to the whole. Facilitate the user to quickly focus what is important in the new image.

(*Continued*)

Table 4. (*Continued*).

Problems	Solutions and motivations
Multiple windows on screen display	Avoid stimulating the user or to enable the user to open multiple windows at the same time. Studies show that users lose if she has many windows on the screen simultaneously.
Place of dialogue	In a new window or dialog is opened so they should be placed near the object that opened them. By placing the new window, close to the object that opened the window so the user can easily connect them. This is especially important if the user often gets interrupted in their work.
	When you place a new window on the screen you should be careful not to hide important information, because then the user must remember that short term memory which is borne in vain.
Use of drag and drop a shortcut	Drag and drop is a powerful way to implement complex tasks when using the mouse, but not the only way to act a task. Since drag and drop is a dynamic operation, users need dynamic feedback on what the outcome of the operation will be. The mouse pointer should tell the user when no drag and what has drawn.
Use the button when the function is the primary user	Use the button in the image if it is a feature that is the primary user of the current image. Use a menu selection on the performance is secondary. It is faster to read by vertically placed buttons, since we can skip the end of the button text. As we read horizontally text to the end and then be able to find the next button. The user has previous knowledge and experience of filling in forms and can quickly see what information is expected in the various fields.
Group menu	Grouping of menu items that belong together in terms of importance to the user. Use menu separator to distinguish the different groups. Put the groups that are generally used only in the menu list. If the user can create categories of information on screen, she can benefit from the image quicker and work more efficiently. Sort menu selections in a logical order for the user.

5 CONCLUSIONS

This study proposed a practice of developing usable user interfaces of VDUs. Various usability methods were used; heuristic evaluation, usability test, questionnaire, and user interview, and design guidelines.

During the usability evaluation of the user interface of VDUs of the alarm and the case and management system of a process control room, we discovered several problems. We have also found a way to improve the user interface by using the design principles and guidelines, through divided the process phases in three fields.

ACKNOWLEDGMENT

This research was supported in part by the Telia-Sonera, Sweden.

REFERENCES

Han, S.H, Yang, H. & ImA, D-G. 2007. Case study of a steel manufacturing company. *International Journal of Industrial Ergonomics* 37: 383–393.

Hix, D. & Hartson, H.R. 1993. *Developing user interfaces: ensuring*. Usability Through Product & Process. Wiley, New York.

Hollywell, P.D. & Marshal, E.C. 1994. An experiment to support the design of VDU-based alarm lists for power plant operators. In N. Stanton (Ed.) *Human factors in alarm design*. Taylor & Francis, London, pp. 31–44.

ISO 1997a. *ISO 9241: Ergonomics requirements for office work with Visual Display* Terminals (VDU), 17 parts. Geneva: International Standards Organisation.

ISO 1997b. *ISO 9241 - 11: Ergonomic requirements for office work with Visual Display Terminals (VDU). Part 11. Guidelines for specifying and measuring* u*sability*. Geneva: International Standards Organisation.

Johannsen, G. 1997. Conceptual design of multi-human–machine interfaces. *Control Engineering Practice* 5(3): 349–361.

Kawai, K. 1997. An intelligent multimedia human interface for highly automated combined-circle plants. *Control Engineering Practice* 5(3): 401–406.

Kirakowski, J. 1996. The software usability measurement inventory: background and usage. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.) *Usability Evaluation in Industry.* Taylor & Frances, UK.

Maguire, M. 2001. Methods to support human-centred design. *Int. J. Human-Computer Studie,* 55: 587–634.

Mayhew, D. 1992. *Principles and guidelines in software user interface design*. Prentice-Hall, Englewood Cliffs, NJ.

Nielsen, J. 1994. *Heuristic evaluation*. In: J. Niesen, R.L. Mack (Eds.) *Usability inspection methods*. Wiley, New York.

Rubin, J. 1994. *Handbook of usability testing: How to plan, design and conduct effective tests.* John Wiley & sons, inc.

Van der Schaaf, T.W. 1989. *Redesigning and evaluating VDU graphics for process control: cognitive ergonomics applied to the operator interface*. In G. Salvendy & M.J. Smith (Eds.) *Designing and Using Human–Computer Interfaces and Knowledge-based Systems*. Elsevier, Amsterdam.

Vincente, K.J. 2001. How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human–Computer Studies* 54: 831–846.

Wang, Y. & Hwang, S. 1995. An experimental study on the information design of CRT display in process control rooms. *International Journal of Industrial Ergonomics* 15: 459–469.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Preference-based usability evaluation of mobile interactive advertisings

Ya-Li Lin & Chih-Hsiang Lai
Department of Statistics, Tunghai University, Taichung, Taiwan

ABSTRACT: This research focuses on the current trends in text-based advertisings as deliver to mobile phones. The user-centered usability of mobile advertisings on the simulated interfaces is evaluated. A preference-based experiment was used as a tool of assessing usability. Forty-three undergraduate and graduate students (21 females and 22 males) participated in the experiment. The design factors of text-based mobile advertising consist of presentation mode (rapid serial visual presentation (RSVP), leading, and static), position (Top, Bottom), speed (80, 170, and 260 WPM), format (keyword-by-keyword, phrase-by-phrase, and sentence), and luminance contrast (2/3, 4/5, and 8/9). The results indicate that most users prefer mobile advertising placing on the top of screen, RSVP presentation with phrase-by-phrase format at speed of 170 WPM, Leading presentation with sentence format at speed of 170 WPM, and color combinations of blue series (foreground) on white (background) with luminance contrasts 2/3, 4/5, and 8/9.

Keywords: mobile advertising, rapid serial visual presentation (RSVP), leading format, subjective preference

1 INTRODUCTION

Mobile Advertising ("MA" for short) is a rapidly growing sector providing brands, agencies and marketers the opportunity to connect with consumers beyond traditional and digital media and directly on their mobile phones (MMA, 2008). Mobile interactive advertising refers to advertising or marketing messages delivered to portable devices, either via a synchronized download or wirelessly over the air (IAB, 2008a). It is a form of advertising via mobile (wireless) phones or other mobile devices (Wikipedia, 2008). Mobile interactivity is in some ways similar to the PC-based internet, and these similarities will speed advertisers' ability to take advantage of the mobile medium.

The early studies of Rapid Serial Visual Presentation (RSVP) were done with rolls of paper that were visible through narrow window or filmstrips (Forster, 1970). This method eliminates the need for eye movements during reading. Laarni's results (2002) indicated that reading rate varied significantly among presentation methods (Normal page, RSVP, Vertical scrolling, Leading, Teletype, and Window). Mills & Weldon (1987) reviewed several studies including dynamic presentation methods for computer screen. Öquist and Goldstein (2003) indicated that RSVP gave the best results for short text in their evaluation was encouraging since thee typical texts read on a mobile device was likely to be short. RSVP could be used effectively on mobile devices with much smaller screens than the one used in their evaluation and be seen as one small step towards an improved readability on mobile devices.

Lin & Shiao (2005) examined the effects of presentation mode (RSVP, and Leading), presentation speed (70 and 140 WPM), and color combinations with luminance contrasts 2/3, 4/5, and 8/9 on visual performance and visual fatigue of small screen space. Results indicated that the interaction of presentation speed and dynamic display would significantly affect visual performance on the small screen space of surfing the homepage of traveling Website. It also showed that visual performance is the best for Leading mode at the presentation speed of 140 WPM, but luminance contrast is not a significant effect on visual performance. The key findings in Lin & Lin (2006) indicated the two-factor interaction of advertising size and ad contents' type as well as the interaction of gender and advertising position would significantly affect the likelihood of click-through on Web advertisings. Lin & Chen (2009) focused on the click-through rate for the advertising effectiveness to examine the effects of design factors on animated online advertisings. Their findings showed order effect, two-factor interactions of ad type and presentation position as well as presentation position and animation length reach statistically significant. In addition, the report of IAB (2008b) indicated that text-based advertising remains the most common creative format on both the Web and other applications like Short Message Services (SMS). Mobile Marketing Association (MMA for short) has published mobile advertising guidelines, but it is difficult to keep such guidelines current in such a fast-developing area. Leveraging accepted Web advertising best practice will facilitate building a successful mobile advertising business. At the same time, the user experience, interactivity, and expectations of consumers on the mobile Web differ from their PC counterparts, and simply transplanting PC-optimized advertising onto mobile device is unlikely to yield optimal results.

As touch screen interfaces become more and more popular, especially mobile display advertising, an important usability problem is how the best to display text-based mobile advertising on touch screen interface. One possibility is to use dynamic presentation methods involving the movement of text on the screen. These methods have been shown to be viable alternatives for displaying text information on small screen of non-PC devices or limited screen space of PC-based devices (Juola, et al. 1982, Juola, et al. 1995, Muter, 1996). These studies focused on investigating users' visual performance, readability, and reading comprehension of dynamic or continuous texts. However, it lacks of providing the guideline of the user experience, interactivity, and expectations of consumers on the mobile Web. This research will focus on the current trends in text-based advertisings as delivered to mobile phones. The user-centered usability of text-based mobile advertisings on the simulated interfaces will be evaluated and introduced as follows.

2 DESIGN OF THE SIMULATED MOBILE ADVERTISINGS INTERFACE

The simulated system would contain text-based mobile advertisings and mobile Web contents. The size of text-based mobile advertising is a large image banner with one line of ad texts up to 18 characters maximum. The font type of mobile advertising uses "New Thin Ming Ti" and font size 12.

RSVP mode is defined as the text presents serially one word of several words at a time in rapid succession at a single visual location (Laarni, 1982). The contents of text-based mobile advertising were moved from bottom to top in rapid succession at a single visual location on the simulated mobile phone interface. Leading mode is defined as that text presents from the right to left continuously along a single line.

Top position is defined as the mobile advertising placed on the top of page, that is, Mobile Web Banner (MMA, 2008). Bottom position is defined as the mobile advertising placed on the bottom of page, that is, Mobile Web Poster. Based on the study of Juola et al. (1995), presentation speeds were determined as 80, 170, and 260 Words Per Minutes (WPM for short). Chinese keyword is a word used to find useful results in internet search (e.g. Walkman). Chinese phase is a sequence of words intended to have some meanings (e.g. Image Size). Chinese sentence is considered to be suitable for presenting Chinese text on single-line display. The consideration of color combinations (R, G, B)/(H, S, B) based on luminance contrasts of 2/3, 4/5, and 8/9 were used in this study.

The prototype of simulated mobile phone interface used in the preference-based experiment is shown in Figures 1–3. Figures 1–2 illustrate three modes of presentation: static and dynamic (RSVP and Leading mode) text-based mobile advertisings. Figure 3 illustrates two positions of presentation: top and bottom of the page. In addition, the specifications of color combinations ((R, G, B)/

Figure 1. Illustration of static text advertising (left-hand side denotes the screen before MA was clicked and right-hand side denotes the screen after MA has been clicking though).

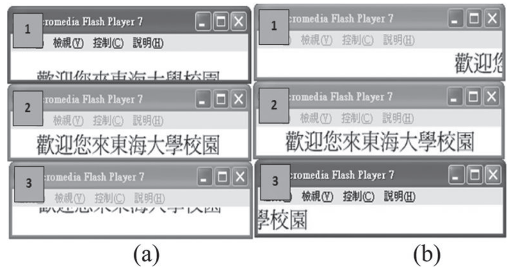


Figure 2. Illustration of dynamic presentation modes: (a) RSVP text advertising moving from bottom to up, (b) Leading text advertising moving from right to left.

(H, S, B)) based on three luminance contrasts of 2/3, 4/5, and 8/9 used in this study are shown in Table 1–3.

The recommendations on how to present visual information on screens of ISO 9241-12 (1998) are based on seven guiding principles: clarity (information should be conveyed quickly and accurately), discriminabilty (information should be able to be distinguished accurately), conciseness (provide only the information necessary to complete the task), consistency (present the same information in the same way throughout the application), detectability (direct the user's attention to the information required), legibility (information should be easy to read), and comprehensibility (the meaning should be clearly understandable). In addition, ISO 9241-12 (1998) also suggest that participants are able to make out the information on screen clearly about luminance contrast getting up to be over 2/3.

(a) (b)

Figure 3. Illustration of positions of text-based mobile advertising: (a) top of page, (b) bottom of page.

3 RESEARCH METHODOLOGY

The design of preference-based mobile advertising interface for mobile users was developed. The subjective preference questionnaire was implemented associated with a simulated mobile phone interface. The subjective preference questionnaire is a structured field of usability assessment. It is useful in the early stages of user-centered design development.

Table 1. Illustration of color combinations of foreground on white background ((R, G, B)/(H, S, B) = (255, 255, 255)/(0, 0, 100)) based on luminance contrast 2/3.

Color combination	Foreground	Foreground: (R, G, B)/(H, S, B)	Color difference
	Olivedrab	(63, 84, 0)/(75, 100, 33)	106
	Dark blue	(0, 63, 84)/(195, 100, 33)	108
	Sienna	(84, 42, 0)/(30, 100, 33)	107
	Gray	(84, 84, 84)/(0, 0, 33)	104

Table 2. Illustration of color combinations of foreground on white background based on luminance contrast 4/5.

Color combination	Foreground	Foreground: (R, G, B)/(H, S, B)	Color difference
	Darkolivegreen	(38, 51, 0)/(75, 100, 20)	126
	Teal	(38, 51, 195)/(195, 100, 20)	127
	Maroon	(51, 26, 0)/(30, 100, 20)	127
	Dimgray	(51, 51, 51)/(0, 0, 20)	124

Table 3. Illustration of color combinations of foreground on white background based on luminance contrast 8/9.

Color combinations	Foreground	Foreground: (R, G, B)/(H, S, B)	Color difference
	Darkgreen	(21, 28, 0)/(75, 100, 11)	139
	Midnigh blue	(0, 21, 28)/(195, 100, 11)	141
	Darkred	(28,1, 0)/(0, 30, 11)	188
	Black	(28, 28, 28)/(0, 0, 11)	138

Table 4. Illustration of color combinations of foreground on white background ((R, G, B)/(H, S, B) = (28, 28, 28)/(0, 0, 11)) based on luminance contrast 8/9.

Color combinations	Foregound	Foreground: (R, G, B)/(H, S, B),	Color difference
	Chartreuse	(102, 255, 0)/(75, 100, 20)	141
	Dodgerblue	(0, 68, 255)/(75, 100, 20)	142
	Red	(255, 0, 25)/(75, 100, 20)	195
	Black	(255, 255, 255)/(0, 0, 100)	138

3.1 *Participants*

Forty-three undergraduate and graduate students (21 females and 22 males) coming from Tunghai University voluntarily participated in the preference-based experiment. They have to express their subjective preference for each mobile advertising using a preference questionnaire based on the Likert's ten-point scale from 1 (extremely dislike) to 10 (extremely like). The distribution of age ranged from 22 to 30 years old (mean age of 24.7 years and standard deviation of 1.73 years). They all had normal vision or corrected vision reaching at least 0.8 and no color-blindness.

3.2 *Apparatus and materials*

The simulated system of mobile advertising was constructed in a Pentium IV desktop computer (CPU1.62GHz, 896MB RAM) with Microsoft Internet Explore 6.0, a 17-inches TFT-LCD monitor (1280 × 1024 pixels). Macromedia Flash MX 2004 was used to design the simulated interface. SWiSHMax (http://www.SWiSHzone.com) was used to design the text effects of static and dynamic mobile advertising on the interface of the simulated system. Ulead PhotoImpact 11 would be used to illustrate the luminance contrast of showing text mobile advertising (foreground) and background area.

3.3 *Statistical analysis*

Chi-squared homogeneous test is used to test if the favorite percentages of k samples are similar. Researcher's hypothesis (H_a) is that at least one of k proportions is not equal. Therefore, null hypothesis (H_0) is the favorite proportions of k samples equal. Test statistic is $\chi^2 = \Sigma_{i=1}^{k}(O_i - E_i)^2 / E_i \sim \chi_{(k-1)}^2$ where the observed frequency denoted as notation O_i and the expected frequency denoted as notation E_i. Reject H_0 if $\chi^2 > \chi_{(k-1)}^2$ at $\alpha = 0.05$ (or P-value $< \alpha$).

4 RESULTS OF SUBJECTIVE PREFERENCE

There are 49% of users choosing Static mode as their favorite percentage of presentation modes, 21% of users choosing RSVP mode, and 30%

of users choosing Leading mode. The results of Figure 4 (a) and Table 5 indicate that the percentage of favoring Static mode is higher than the other modes (RSVP and Leading), however, it lacks of significant difference among three display modes ($\chi^2 = 5.209$, *P-value* = 0.074).

There are 88% of users choosing Top of the page as their favorite percentage of presentation positions, and 12% of users choosing Bottom of the page. The results of Figure 4 (b) and Table 5 indicate that the percentage of favoring the Top position is higher than the one of Bottom position. The proportional differences of subjective preference between Top and Bottom of the pages are statistically significant ($\chi^2 = 25.326$, *P-value* = 0.000).

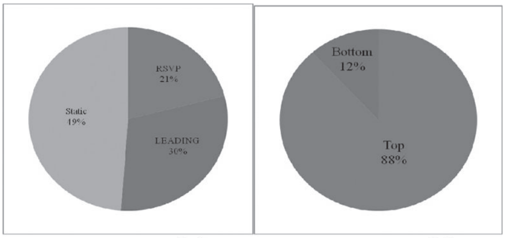
The favorite percentage of presentation speed under the condition of RSVP mode and sentence format is 7% of users choosing presentation speed of 80 WPM, 60% of users choosing presentation speed of 170 WPM, and 33% of users choosing presentation speed of 260 WPM. The results of Figure 5 (a) and Table 5 indicate that the percentage of favoring the presentation speed of 170 WPM is higher than the ones of 80 and 260 WPM under the condition of RSVP mode with sentence format. The proportion differences of subjective preference among 80, 170 and 260 WPM are also statistically significant ($\chi^2 = 18.465$, *P-value* = 0.000).

The favorite percentage of presentation format under the condition of RSVP mode and 170 WPM is 23% of users choosing keyword-by-keyword

Table 5. Frequency distribution of subjective preference questionnaire.

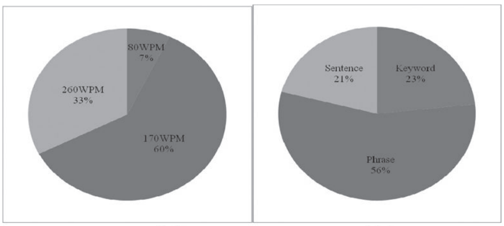
Factor	Level	Favorite frequency (Percentage)	*Chi-squared statistic*	*P-value*
Presentation Mode	RSVP	9 (21%)	5.209	0.074
	Leading	13 (30%)		
	Static	21 (49%)		
Ad Position	Top	38 (88%)	25.326	0.000*
	Bottom	5 (12%)		
Presentation Speed Based on RSVP Display	80 WPM	3 (7%)	18.465	0.000*
	170 WPM	26 (60%)		
	260 WPM	14 (33%)		
Presentation Format Based on RSVP Display	Keyword	10 (23%)	9.814	0.007*
	Phrase	24 (56%)		
	Sentence	9 (21%)		
Presentation Speed Based on Leading Display	80 WPM	16 (37%)	18.744	0.000*
	170 WPM	25 (58%)		
	260 WPM	2 (5%)		
Presentation Format Based on Leading Display	Keyword	6 (14%)	13.163	0.001*
	Phrase	12 (28%)		
	Sentence	25 (58%)		
Luminous Contrast 2/3(Background: White)	(F)Olive drab	8 (19%)	19.047	0.000*
	(F)Dark blue	23 (53%)		
	(F) Sienna	5 (12%)		
	(F)Gray	7 (16%)		
Luminous Contrast 4/5(Background: White)	(F)Dark olive green	4 (9%)	18.116	0.000*
	(F)Teal	22 (51%)		
	(F)Maroon	6 (14%)		
	(F)Dim gray	11 (26%)		
Luminous Contrast 8/9 (Background: White)	(F)Dark green	7 (16%)	16.442	0.001*
	(F)Midnight blue	22 (51%)		
	(F)Dark red	5 (12%)		
	(F) Black	9 (21%)		
Luminous Contrast 8/9 (Background: Black)	(F)Chartreuse	14 (33%)	15.3952	0.000*
	(F)Dodger blue	0 (0%)		
	(F)Red	4 (9%)		
	(F)White	25 (58%)		

Note: "*" denotes *chi-squared statistic* reaches statistically significance at $\alpha = 0.05$.



(a) (b)

Figure 4. Pie charts of favorite percentage for: (a) presentation mode and (b) presentation position.



(a) (b)

Figure 5. Pie charts of favorite percentage of: (a) presentation speed with RSVP and sentence format, and (b) presentation format with RSVP and 170 WPM.

format, 56% of users choosing phase-by-phase format, and 21% of user choosing sentence format. The results of Figure 5 (b) and Table 5 indicate that the percentage of favoring the phase-by-phase format is higher than the ones of keyword-by-keyword and sentence formats under the condition of RSVP mode and 170 WPM. The proportional differences of subjective preference among keyword-by-keyword, phase-by-phase, and sentence are statistically significant ($\chi^2 = 9.814$, *P-value* = 0.007).

The favorite percentage of presentation speed under the condition of Leading mode and sentence format is 37% of users choosing presentation speed of 80 WPM, 58% of users choosing presentation speed of 170 WPM, and 5% of users choosing presentation speed of 260 WPM. The results of Figure 6 (a) and Table 5 indicate that the percentage of favoring the presentation speed of 170 WPM is higher than the ones of 80 and 260 WPM under the condition of Leading mode with sentence format. The proportional differences of subjective preference among 80, 170 and 260 WPM are statistically significant ($\chi^2 = 18.744$, *P-value* = 0.000).

The favorite percentage of presentation format under the condition of Leading mode and 170 WPM is 14% of users choosing keyword-by-keyword format, 28% of users choosing phase-by-phase format, and 58% of users choosing sentence format. The results of Figure 6 (b) and

(a) (b)

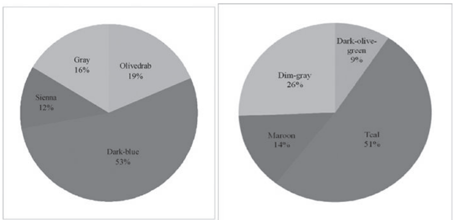
Figure 6. Pie charts of favorite percentage of: (a) presentation speed with Leading and sentence format, and (b) presentation format with Leading and 170 WPM.

Table 5 indicate that the percentage of favoring the sentence format is higher than the ones of keyword-by-keyword and phase-by-phase formats under the condition of Leading mode and 170 WPM. The proportional differences of subjective preference among keyword-by-keyword, phase-by-phase, and sentence are statistically significant ($\chi^2 = 13.163$, *P-value* = 0.001).

The favorite percentage of color combination under the condition of luminance contrast 2/3 and white background is 53% of users choosing dark blue foreground, 19% of users choosing olivedrab foreground, 16% of users choosing gray foreground, and 12% of users choosing sienna foreground. The results of Figure 7 (a) and Table 5 indicate that the percentage of favoring the color combination of dark blue foreground and white background is higher than the ones of the other color combinations under luminance contrast 2/3. The proportional differences of subjective preference among the four color combinations under luminance contrast 2/3 are statistically significant ($\chi^2 = 19.047$, *P-value* = 0.000).

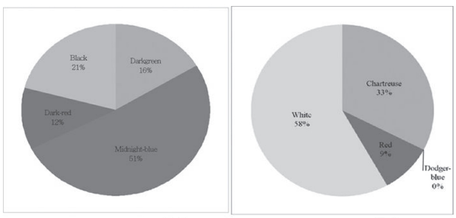
The favorite percentage of color combination under the condition of luminance contrast 4/5 and white background is 51% of users choosing teal foreground, 26% of users choosing dimgray foreground, 14% of users choosing maroon foreground, and 9% of users choosing dark olive green foreground. The results of Figure 7 (b) and Table 5 indicate that the percentage of favoring the color combination of teal foreground and white background is higher than the ones of the other color combinations under luminance contrast 4/5. The proportional differences of subjective preference among the four color combinations are statistically significant ($\chi^2 = 18.116$, *P-value* = 0.000).

The favorite percentage of color combination under the condition of luminance contrast 8/9 and white background is 51% of users choosing midnight blue foreground, 21% of users choosing black foreground, 16% of users choosing dark green foreground, and 12% of users choosing



(a) (b)

Figure 7. Pie charts of favorite percentage of: (a) luminance contrast 2/3 based on white background, and (b) luminance contrast 4/5 based on white background.



(a) (b)

Figure 8. Pie chart of favorite percentage of: (a) luminance contrast 8/9 based on white background, and (b) luminance contrast 8/9 based on black background.

dark red foreground. The results of Figure 8 (a) and Table 5 indicate that the percentage of favoring the color combination of midnight blue foreground and white background is higher than the ones of the other color combinations under luminance contrast 8/9. The proportional differences of subjective preference among the four color combinations are statistically significant ($\chi^2 = 16.442$, $P\text{-}value = 0.001$).

The favorite percentage of color combination under the condition of luminance contrast 8/9 and black background is 58% of users choosing white foreground, 33% of users choosing chartreuse foreground, 9% of users choosing red foreground, and 0% of users choosing dodger blue foreground. The results of Figure 8 (b) and Table 5 indicate that the percentage of favoring the color combination of white foreground and black background is higher than the ones of the other color combinations under luminance contrast 8/9. The proportional differences of subjective preference among the four color combinations are statistically significant ($\chi^2 = 15.395$, $P\text{-}value = 0.000$).

5 CONCLUSIONS

The summaries made by the results of subjective preference as following: (1) Most users prefer mobile advertising placing on the top of page to bottom one; (2) Most users prefer presentation speed of 170 WPM to 80 and 260 WPM under the condition of RSVP and sentence and Leading and sentence; (3) Most users prefer phrase-by-phrase format under RSVP and 170 WPM, however, they prefer sentence format under Leading and 170 WPM; and (4) Most users prefer blue series foreground to others based on the white background whatever the luminance contrasts are 2/3, 4/5, and 8/9.

REFERENCES

Foster, K.L. 1970. Visual perception of rapidly presented word sequences of varying complexity. *Perception & psychophysics* 8: 215–221.

Interactive Ad Bureau 2008a. *A Mobile Advertising Overview*: 4–6.

Interactive Ad Bureau 2008b. *IAB Internet Ad Revenue Report*: 8–10.

ISO 1998. *ISO 9241-12 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 12: Presentation of information*. International Organization for Standardization (ISO). http://www.iso.org/iso/home.htm

Juola, J.F., Tiritoglu, A. & Pleunis, J. 1995. Reading text presented on a small display. *Applied Ergonomics* 26(3): 227–229.

Juola, J.F., Ward, N.J. & McNamara, T. 1982. Visual search and reading of rapid serial presentations of letter strings, words and text. *J. Exper. Psychol:General 111*: 208–227.

Laarni, J. 2002. Search for optimal methods of presenting dynamic text on different types of screens. *ACM International Conference Proceeding Series* 31: 219–222.

Lin, Y.L. & Chen, Y.W. 2009. Effects of ad type, position, animation length, and exposure times on click-through rate of online advertisings. *Computers & Industrial Engineering* 57(2): 580–591.

Lin, Y.L. & Lin, K.C. 2006. Effects of ad size, position, type, and user's gender on the click-through rate of Web advertisement. In *Proceedings of the APCHI 2006(7th Asia-Pacific Conference on Computer Human Interaction),* October 11–14. Taipei.

Lin, Y.L. & Shiao, S.H. 2005. Visual performance of dynamic display design on the small screen for travel homepage. *Journal of Ergonomic Study* 7(2): 105–112.

Mills, C.B. & Weldon, L.J. 1987. Reading text from computer screens. *ACM Computer Surveys* 19(4): 329–358.

Mobile Marketing Association. 2008. *Mobile Advertising Guidelines*: 3–10.

Muter, P. 1996. Interface design and optimization of reading of continuous text. *Cognitive Aspects of Electronic Text Processing*: 161–180.

Öquist, G. & Goldstein. M. 2003. Towards an improved readability on mobile devices: evaluation adaptive rapid serial visual presentation. *Interacting with Computer* 15: 539–558.

Wikipedia. 2008. Mobile advertisings. Retrieved from *http://en.wikipedia.org/wiki/Mobile_ad*

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Discriminant analysis for detection of low usability web pages

Noboru Nakamichi
Nanzan University, Seto, Japan

Toshiya Yamada
The Graduate University for Advanced Studies, Tokyo, Japan

Tomoko Matsui
The Institute of Statistical Mathematics, Tokyo, Japan

ABSTRACT: The purpose of this work is to reduce the cost of the web usability evaluation by usability testing. The cost will reduce by detecting low usability web pages. We analyzed empirically to find detectable metrics from the quantitative data including eye movement. We investigate the relation between the quantitative data about the behavior of users and the web usability evaluated by subjects. The subjects examined 263 web pages in total. The Logistic Regression Analysis (LRA) is used to detect the low usability web pages. The metric is defined using the browsing time, the amount of wheel rolling and the moving distance of gazing points which are selected by the stepwise selection method based on the Area Under the ROC Curve (AUC). When setting the cut-off value of 20 pages with low usability, the highest detection rate of the web pages with low usability is 0.65 by LRA. From the results, if the moving speed of gazing points is high, the possibility of low usability web page is high.

Keywords: gazing point, eye information, web usability, linear discriminant analysis, logistic regression

1 INTRODUCTION

The usability of web sites is so important that it can influence the amount of sales, because users are unwilling to read web pages with low usability [Kelly and Emily 2002]. To create easy-to-use web pages, usability evaluation is required [Nielsen 2000]. Among various methods for usability evaluation, usability testing [Joseph and Redish 1993] is widely used. Usability testing is a method of discovering problems based on operation of the application by subjects. It tends to discover serious problems leading to trouble, and other problems which evaluators cannot discover. However, the evaluators take time for analyzing recorded data such as a user's utterance data and VTR.

Supporting methods based on quantitative data of users' behavior are proposed in usability evaluation. Conventionally, specialists evaluated web usability using quantitative data based on experience and knowledge. For example, when using gazing points, a specialist finds areas where gazing crosses and returns. Then the specialist analyzes a users' behavior in those areas, and evaluates usability problems on a web page. Specialists can evaluate web usability for details. However it couldn't reduce cost. Companies which change a website frequently need to improve the efficiency of evaluation. Therefore, they need to first detect low usability pages on a website and then quickly narrow down the objects of usability evaluation.

Quantitative data about user's behavior may support detection of low usability Web pages without special knowledge. We hypothesized that characteristics of various quantitative data used for evaluation are related to Web usability. If low usability Web pages are detectable from quantitative data, the efficiency of usability evaluation will increase without additional new work. However, it was not clear which data would be effective in detection of low usability Web pages among various quantitative data.

We experimented to investigate the relation between quantitative data viewing user's behavior and web usability evaluation by subjects for detecting low usability web pages. We define low usability pages as "hard to use" pages subjects choose using the questionnaire. Variable selection result by logistic regression, we showed clearly that the browsing time, wheel rolling and the moving distance of gazing points are effective in detecting low usability Web pages.

2 RELATED WORKS

In this chapter, we describe methods of supporting evaluation objectively using quantitative data such as browsing time, mouse movement and eye movement.

Laila [Laila and Fabio 2002] analyzed the execution situation of a task from a users' operation event recorded by Java script. He supported the analysis based on quantitative data, such as page reference time and task execution time. He analyzed the usability of a Web page based on task execution time totaled for every Web page. Okada [Okada and Asahi 1999] developed the GUI-TESTER which extracts a common operation pattern from two or more users' operation history. If the tool is used, the operation pattern for mistaken operation can be extracted. And when the moving distance of a mouse cursor is long and an operation time interval is long, they suggest a possibility that a screen layout is bad.

WebTracer [Nakamichi et al. 2007] can collect the operation log of users on the Web pages. Collectable data include the information on user's sight line (the coordinates of the gazing point on the computer screen), operation log of a mouse, and the displayed screen images, together with their time information. The data collected by WebTracer characterize Web pages and have the possibility of being used for supporting the usability evaluation. However, the relation between such data and the problems in the Web usability was merely an example of the characteristics of the Web pages. Quantitative evaluation of the relation to the usability of Web pages was not done.

Heatmaps from user eyetracking studies based on fixations are used for observing in detail. Nielsen found that user's main reading behavior was fairly consistent across many different sites and tasks. [Nielsen 2006]. Eyetracking visualizations show that users often read Web pages in an F-shaped pattern: two horizontal stripes followed by a vertical stripe. Specialists may evaluate using eyetracking heatmaps. However it doesn't lead to the cost reduction of the web usability evaluation.

Eye-tracking methodologies are applying in the domain of Web search because gaze can be used as a proxy for a user's attention. Eye-tracking measures include pupil dilation, fixation information, and sequence information such as scan paths [Guan and Cutrell 2007]. They relied on measures related to gaze fixations with a minimum threshold of 100 ms in areas of interest. They found that as they increased the length of the query-dependent contextual snippet in search results, performance improved for informational queries but degraded for navigational queries. Analysis of eye movements showed that the decrease in search performance was partially due to the fact that users rarely looked at lower ranking results. Matsuda measured user's eye movements during web search tasks to analyze how long users spend on each result of the results pages [Matsuda et al. 2009]. They found the results displayed on the bottom of the page were viewed for a shorter time than the results displayed on the top of the next page.

In these conventional researches, the specialist had discriminated usability only using certain quantitative data. However, the effectiveness of combinational data was not verified statistically.

3 EXPERIMENT OF USABILITY TESTING

3.1 *Quantitative data of user's behavior*

Browsing, mouse movement, and eye movement are the quantitative data about user's behavior mainly used for web usability evaluation. This experiment recorded the following 6 quantitative data for every Web page:

- Browsing time (sec): Time since the user begins to see a certain page, until it changes to another page
- Moving distance of mouse (pixel)
- Moving speed of mouse (pixel/sec)
- Wheel rolling (Delta)
- Moving distance of gazing points (pixel)
- Moving speed of gazing points (pixel/sec)

Gazing point is the point at the intersection of the user's look with the target screen.

3.2 *Experimental environment*

The experiment environment used by this research is as follows:

- Display: 21 inches (Viewable screen size: H30 × W40 cm)
- Device for measurement of sight line: NAC, EMR-NC (View angle: 0.28, resolution on the screen: approx. 2.4 mm)
- Recording and playing of sight-line data: WebTracer (Sampling rate: 10 times per second)

WebTracer [Nakamichi et al. 2007] is an environment for recording and analyzing the user's operations in Web pages.

3.3 *Experimental procedure*

We experimented with usability evaluation in the following procedures to five tasks. Subjects are 10 frequent users of the Internet. They have never visited the sites used in the experiment.

Table 1. t-Test of low usability pages and other pages every kind of quantitative data.

Quantitative data for each pages	Evaluation result by subjects				t-Test (significance probability P)
	Low Usability: 20 pages		Others: 243 pages		
	Average	S.D.	Average	S.D.	
Browsing time (sec)	16.9	12.4	11.4	10.2	0.06579
Moving distance of mouse (pixel)	1197.0	712.8	1027.7	1052.9	0.33699
Moving speed of mouse (pixel/sec)	91.9	67.6	105.1	73.0	0.41283
Wheel rolling (Delta)	570.0	949.6	229.1	529.8	0.12863
Moving distance of gazing points (pixel)	8723.5	5574.7	4608.8	3841.7	*0.00402*
Moving speed of gazing points (pixel/sec)	546.2	140.1	419.6	155.9	*0.00081*

S.D. : Standard Deviation

We requested the subject to perform five tasks of looking for the starting salary of a master from the site of five companies, as a main experiment.

Procedure 1: The Web page for an experiment linked to the top page of each company is displayed by a subject. And the experiment is started from the time of a subject clicking the link.
Procedure 2: While subjects are doing the tasks, several types of quantitative data are recorded using WebTracer.
Procedure 3: The Web pages that subjects visited are displayed. We requested the subject to choose the ease of use for every visited Web page from the following five levels. We defines a low usability page as a page that a subject choose "hard to use" from five levels of the questionnaire.

1. hard to use
2. relatively hard to use
3. relatively easy to use
4. easy to use
5. don't know

Procedure 4: We reproduce the operation history recorded by WebTracer, and a subject checks all the visited Web pages. At that time, we interviewed the subjects about the situation of their search.

3.4 *t-Test of low usability pages and others*

We recorded the quantitative data for 275 pages which the subject visited. We were not able to record correctly about 12 pages of them. The cause is a frequent blink and head movement. Moreover, there were eight pages which the subject answered "don't know" about the usability of the Web page. We measured the quantitative data in 263 pages except these pages.

Following the answers in the experiment, we classified the cases into two types: the cases with low usability and the other cases. If the subject in a case considered a Web page "hard to use" then we regard this as a case with low usability. If the subject answered "relatively hard to use" or "relatively easy to use" or "easy to use" then we regard this as another case.

We postulated a difference between the quantitative data for "low usability" and "others" pages which would allow detection of "low usability" pages. For each type of data, we performed a statistical test with the hypothesis that there is a difference between the two classes of cases.

Table 1 shows the average values and standard deviations for each type of quantitative data, together with the results of the above statistical tests. The results of the statistical tests in Table 1 show that the mean of each type of operation data for the cases with low usability is statistically different from that for the cases with others. However, the moving distance of the gazing point and the moving speed of the gazing point have both significantly different variances and means. This result showed clearly that the quantitative data measures that are effective in Web usability evaluation are the moving distance of the gazing point and the moving speed of the gazing point. In addition, on the low usabil-ity pages, we found that the moving distance of the gazing point was long, and the moving speed of the gazing point was high.

4 DISCRIMINANT ANALYSIS OF LOW USABILITY PAGES

We analyzed the data to detect low usability web pages using logistic regression analysis. Detecting for low usability pages/others shows in Table 2. We make optimal logistic model under step-down procedure based on a Receiver Operating Characteristic curve (ROC Curve) and an

Table 2. Detecting for low usability pages/others.

		Actual condition	
		Low usabilit (20 pages)	Others (243 pages)
Test result	Low usability (Positive)	True Positive (TP)	False Positive (FP)
	Others (Negative)	False Negative (FN)	True Negative (TN)

area under the ROC curve (AUC). Criteria AUC value is average AUC value of 10 data sets. Optimal logistic model such as more effective one in few variables.

We set cut-off value is 20/263 (= low usability pages/all pages) as discrimination threshold. Cut-off value usually is set 0.5 in disciriminant analysis. But low usability pages are fewer than others in the experiment. It is difficult to detect low usability pages by number of data bias. We set cut-off value is 20/263 in logistic regression. And the case with cut-off value = 20/263 was compared with the case with cut-off value = 0.5.

4.1 *Variable selection by logistic regression*

We select variable to make logistic model for detection of low usability pages from all recorded quantitative data. We reduce variables for making optimal logistic model under backward elimination method based on a Receiver Operating Characteristic curve (ROC curve). We use area under the ROC curve (AUC) for backward elimination criterion. The AUC is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. Criteria AUC value is average AUC values of 10 data sets which were selected by the random sampling. Criteria mean correct ratio and mean true positive ratio are average of 10 data sets too. A data sets is composed training data set which selected by the random sampling of 10 pages from low usability pages and 121 pages from other pages and test data set which remaining pages in each groups.

In delete variable terms, we change cut-off value from 1/263 to 262/263. And we delete the variable of highest AUC. Variable selection step is showed in Table 3. Finally, we get logistic model composed browsing time, moving distance of gazing point and wheel rolling as follows.

$$\text{Logit}(p) = a + b1(\text{Browse}) + b2(\text{E\_dis}) + b3(\text{Wheel})$$

Fig. 1 is ROC curves with Specificity ratio (= TN/(FP+TN)) as the X-axis, the Sensitivity ratio (= TP/(TP+FN)) as the Y-axis, and ROC curves in each step plotted on it. From Fig.1 and Table 3, variables selection logistic model with cut-off value = 20/263 is the nearest 1.0.

4.2 *Comparison between logistic model based on selection model*

We compare logistic model with the performance of Linear Discriminant Function (LDF). When comparing it, it compared it in case of all variables model and variable selection model. Moreover it compared it in case of cut-off value = 0.5 and cut-off value = 20/263. Table 4 is detection result of logistic regression.

Logistic model of variables selection with cut-off value = 20/263 is highest mean TP ratio. Logistic model based on estimated intercept and regression coefficient by training data which were selected by the random sampling as follows.

$$\begin{aligned}\text{Logit}(p) = &-3.5659778 - 0.1356705\,(\text{Browse}) \\ &+ 0.0004381\,(\text{Eye\_dis}) \\ &+ 0.0006303\,(\text{Wheel})\end{aligned}$$

Mean TP ratio is important in web usability evaluation for detecting low usability pages. So logistic model of variables selection with cut-off value = 20/263 is the most effective. The performance of variables selection model is better than that of all variables model. The result of mean TP

Table 3. Variable selection by meanAUC.

Step1: Deleted variable is moving speed of mouse.

Variables	Mean C	Mean TP	Mean AUC
ALL	0.719	0.530	0.716
Browse (Browsing time)	0.722	0.570	0.733
M_dis (Moving distance of mouse)	0.729	0.600	0.755
M_sp (Moving speed of mouse)	0.734	0.570	0.766
Wheel (Wheel rolling)	0.705	0.540	0.721
E_dis (Moving distance of gazing point)	0.735	0.550	0.721
E_sp (Moving speed of gazing point)	0.720	0.560	0.713

Mean C: Mean Correct ratio = (TP + TN)/ALL
Mean TP: Mean True positive ratio = TP
Mean AUC: Mean Area under the ROC curve

Step2: Deleted variable is moving distance of mouse.

Variables	Mean C	Mean TP	Mean AUC
All—M_sp	0.734	0.570	0.766
M_sp—Browse	0.738	0.610	0.774
M_sp—M_dis	0.733	0.620	0.808
M_sp—E_dis	0.716	0.550	0.755
M_sp—E_sp	0.762	0.550	0.784
M_sp—Wheel	0.728	0.580	0.758

Step3: Deleted variable is moving speed of gazing points.

Variables	Mean C	Mean TP	Mean AUC
All—M_sp—M_dis	0.733	0.620	0.808
M_sp—M_dis—Browse	0.723	0.620	0.808
M_sp—M_dis—E_dis	0.707	0.620	0.798
M_sp—M_dis—E_sp	0.763	0.650	0.828
M_sp—M_dis—Wheel	0.727	0.630	0.806

Step4: Deleted variable is none.

Variables	Mean C	Mean TP	Mean AUC
All—M_sp—M_dis—E_sp	0.763	0.650	0.828
M_sp—M_dis—E_sp—Browse	0.735	0.540	0.760
M_sp—M_dis—E_sp—E_dis	0.740	0.490	0.673
M_sp—M_dis—E_sp—Wheel	0.754	0.630	0.820

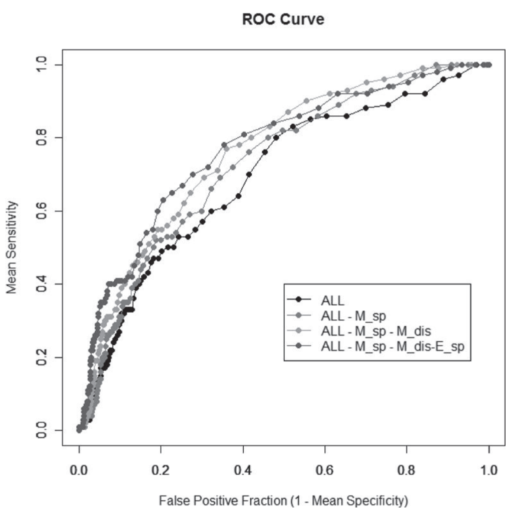


Figure 1. ROC curves in each step.

Table 4. Detection result by logistic regression.

Variables	All	All	Selection	Selection
Cut-off value	0.5	20/263	0.5	20/263
Mean Correct ratio	0.910	0.719	0.917	0.763
Mean TP ratio	0.030	0.530	0.030	*0.650*
AUC	0.716	0.716	0.828	0.828

Table 5. Detecting result of a training data by logistic model.

		Actual condition	
		Low usability (10 pages)	Others (121pages)
Test result	Low usability	7 pages	34 pages
	Others	3 pages	87 pages

ratio is greatly different according to the selection of cut-off value.

Detecting result of a training data by the logistic model is shown in Table 5. From the result, we confirm the performance of logistic model of variables selection with cut-off value = 20/263.

5 DISCUSSION

5.1 *User's behavior in low usability pages*

From variable selection result, browsing time, moving distance of gazing point and wheel rolling are selected. The moving speed of the gazing point has both significantly different variances and means in t-Test result of Table 1. Browsing time and moving distance of gazing point that composed moving speed of the gazing point were selected.

We focused subject's behavior in a low usability web page is shown in Fig. 2. The subject's gazing point goes to various parts of the Web page from Fig. 2. When focus to trajectory of gazing point and mouse cursor, the subject's gazing point goes to various parts of the Web page compared with mouse cursor. When focus to gazing point, gazing points are centered on many parts of the Web page. Sampling rate of gazing point is 10 times per second, however a part of moving distance of gazing point is long. From the subjects, the following two kinds of comment were obtained in the Web page.

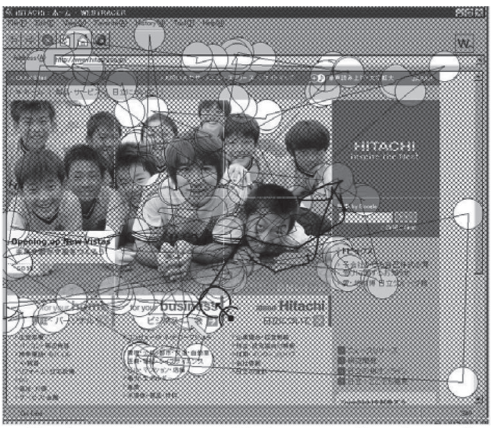
"I couldn't easily find the link which leads to the objective information."

"I got lost because the menu layout is bad."

It is difficult to discover bad layout problem on the web page only by centered part of gazing points. You have the possibility to discover bad layout pages from Websites using the moving speed of the gazing point for each pages. Moreover using centered part of gazing points, you have the possibility to discover cause of bad layout from Web pages.

5.2 *Advantage of applying in usability testing*

Realistic approach for applying to usability testing is setting of cut-off value. From detection result of discriminant analysis in Table 5, setting of cut-off value is important for detecting low usability pages. We set cut-off value = 20/263 because we had already known the ratio of low usability web pages and recorded all pages from the experimental result as prior probability. Since the cost which Web usability evaluation takes is increasing in software development, increasing efficiency is required. Before a specialist analyzes the quantitative data of user's behavior, the specialist needs to evaluate all Web pages which the user visited to discriminate low usability pages.



○ : Gazing point, ——: Trajectory of gazing point

Figure 2. User's behavior in a low usability web page.

A specialist set cut-off value based on limited cost for detecting pages with high possibility that usability is low. If the time for the evaluation is decided, the number of pages that the specialist evaluates during that time is foreseeable. From ROC curves of Fig. 1, the specialist can detect low usability pages that should be evaluated using cut-off value based on the number of pages. We analyzed the relation between detectable rate of low usability web pages and extracted number of pages. Fig. 3 shows Mathematical line based on logistic model of variables selection with cut-off value = 20/263. From the result, if you want to detect 80% of A, it is necessary to evaluate 2/5 of all recorded pages. We expect that setting of cut-off value based on limited cost can lead to cut evaluation costs by setting only the detected low usability pages as the evaluation target.

6 CONCLUSION

We experimented to investigate the relation between the quantitative data viewing behavior

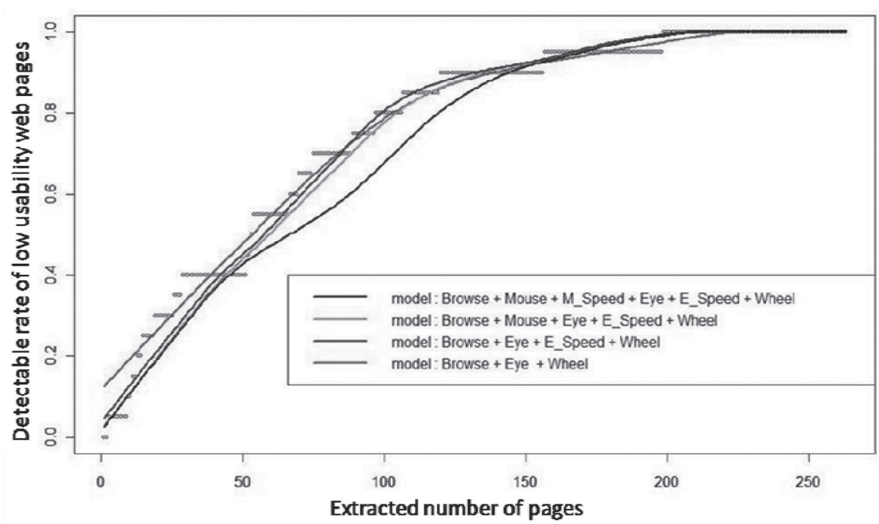


Figure 3. Mathematical line based on logistic model.

of users and web usability evaluation by subjects. We recorded 20 low usability pages include in 263 pages subjects visited. As a result of this variable selection based on logistic regression, we showed clearly that the browsing time, the wheel rolling and the moving distance of gazing points are effective in detecting low usability web pages. Logistic model of variables selection with cut-off value = 20/263 is highest mean TP ratio. Mean TP ratio is important in web usability evaluation for detecting low usability pages. From the results, if the moving speed of gazing points is high, the possibility of low usability web page is high.

These quantitative data which we evaluated are already used in existing usability evaluations. We must also clarify the behavior of a subject on a page with features that cannot be discriminated, to remove it, and to investigate higher discrimination. And discriminating the relevance of web search results using pupil size as new quantitative data is starting [Oliveira et al. 2009]. There might be a possibility that the mean TP ratio can be improved in addition by adding such new data.

ACKNOWLEDGMENTS

This research was partially supported by Nanzan University Pache Research Subsidy I-A-2 for the 2010 academic year.

REFERENCES

Guan, Z. & Cutrell, E. 2007. An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 417–420.

Joseph, S.D. & Redish, A.P. 1993. *A Practical Guide to Usability Testing.* Ablex Publishing.

Kelly, G. & Emily, G. 2002. Web redesign. Pearson Education.

Laila, P. & Fabio, P. 2002. Intelligent analysis of user interactions with web applications. In *Proceedings of the 7th international conference on Intelligent user interfaces*, 111–118.

Matsuda, Y., Uwano, H., Ohira, M. & Matsumoto, K. 2009. An Analysis of Eye Movements during Browsing Multiple Search Results Pages. In *Proceedings of the 13th International Conference on Human-Computer Interaction*, Springer-Verlag, 121–130.

Nakamichi, N., Sakai, M., Hu, J., Shima, K. & Matsumoto, K. 2007. Webtracer: A new web usability evaluation environment using gazing point information. In *Electronic Commerce Research and Applications* 6(1): 63–73, Elsevier Science.

Nielsen, J. 2000. *Designing web usability.* New Riders.

Nielsen, J. 2006. F-shaped pattern for reading web content. Jakob Nielsen's Alertbox, retrieved from *http://www.useit.com/ alertbox/reading_pattern.html*

Okada, H. & Asahi, T. 1999. GUITESTER: a log-based usability testing tool for graphical user interfaces.. *IEICETrans. on Infor-mation and systems* E82-D(6): 1030–1041.

Oliveira, T.P.F., Aula, A., Russell, M.D. 2009. Discriminating the relevance of web search results with measures of pupil size. In *Proceedings of the 27th international conference on Human factors in computing systems*, 2209–2212.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A conceptual framework to determine usability criteria for family of products for older Malaysians

Hassan Zareei & Rosnah Mohd. Yusuff
Faculty of Engineering, Department of Mechanical and Manufacturing Engineering, Serdang, Malaysia

Sharifah Norazizan SAR & Mohd. Rizal Hussain
Institute of Gerontology, University Putra Malaysia, Serdang, Malaysia

ABSTRACT: This study proposes a conceptual framework which has three steps to derive and rank usability criteria for family of products related to Instrumental/Enhanced Activity of Daily Living (I/EADL). The three steps include, measuring level of functional disability by a modified IADL scale to obtain the kind and level of disability and identify related family of products, developing an initial list of recognized usability factors, and conducting Focus Group Discussion (FGD) process to develop and rank a complete list of usability criteria for family of products to support I/EADL. The proposed method can be used to evaluate product usability and create a helpful link between level of usability and IADL scale, thereby making better decision in purchasing each family of product for older Malaysians.

Keywords: usability criteria, elderly user, Malaysia, I/EADL, FGD, family products

1 INTRODUCTION

Nowadays, the ageing population is a universal fact touching both developed and developing countries. Elderly people are a distinct subgroup with special human factors needs wishing to continue to live independently. In designing products, taking into consideration the requirements of users with extraordinary needs in daily routines would result in a more widely useful product. However, many new systems and products were rejected by older users because of usability problems that means older people should be taken into account in usability study when designing new products and technologies.

2 BACKGROUND

Holzinger et al. (2008) claimed that increasing diversity of elderly populations, both culturally and educationally, requires the development of a specific set of criteria. They studied two different aspects of usability for the elderly; passive and active interaction in area of information technology. The second category, which referred as active technology, required not only the established standard but also those usability criteria which applied to individuals with special needs.

In this way the consumer-older person who uses a product for a long time-not only must be the ultimate evaluator of whether a product is satisfactory but also is in the best position to propose factors to be considered in developing evaluation criteria for a specific product. Failure to consider such factors is likely to result in the giving up of products by elderly persons, and more dissatisfaction, frustration, and economic waste associated with such abandonment (Arthanat et al. 2007). In order to utilize properly family of products in society especially for senior citizens, products must be properly designed for the end user. One way to achieve this goal is through effective and efficient usability model.

2.1 *Issues and challenges in general*

As Tinker (1997) stated, it is clear that the differences among the older people make it difficult to generalized. Ageing affects most of physical and cognitive functions of this group that made the design of products rather challenging experience. Characteristics such as different degrees of disabilities and impairments (Jorge 2001) can make it an almost impossible task, and the varying degrees of cognitive functioning that challenge the ways in which to promote easy learning and defeat development of usable interfaces for them. It is evident that some elderly problems in this context related to health or medical problems that are not fixable by any human factors interventions.

However, as pointed out by Hawthorn (2000) it is also important to remember that some older people have just minor disabilities. Although, by applying a "Dynamic Diversity" approach (Gregor et al. 2002) technology development focuses both on differences between individuals and on variations within the same individual according to the user's ageing process. In reality, it is rarely possible to collect all criteria to design a product that satisfies all intercultural groups.

2.2 *Special issues for older Malaysians*

Malaysia is characterized by a multi-racial, multi-religious, multi-cultural and multi-ethnic population. The majority of Malaysia's people consist of Malays and other Bumiputera groups make up 65% of the population, Chinese 26%, Indians 8% and other unlisted ethnic groups 1% (Department of Statistics, 2000). Although English is widely used in Malaysia, the percentage of elderly people who can use English is lower in rural areas. So, they cannot use many technologies such as websites and computer application because of incompetency in the English language.

Syariffanor (2006) has pointed out cultural lifestyle has significant influence to the level of acceptance and using technology among the multi-racial communities in Malaysia, for example, some people may not use the appliance at all in rural area but some in urban use extensively. The differences in cultural lifestyle also influence their social activities. For instance, Chinese elderly people like to join senior citizens' clubs to socialize while Malays elderly people, especially those who live in rural area, participate in community activities and religious activities organised by the mosques in their residential area.

Ismail et al. (2001) have pointed out the majority of elderly Malaysians lose their main source of income once they retire. On average, Malaysian retirees feel the amount they receive is not sufficient to cover their household expenses. Retirees in central Peninsular Malaysia tend to be better off while unskilled workers and low income people have a lower pension than average. The difference between high and low incomes is quite big, for instance, the high income retirees have four times more retirement income than those with a low income (AXA Retirement Scope 2008). Thus, product use could be affected due to the income amount, namely high income group can use more appliances while poor people use less.

Malaysian elderly do not have much educational achievement (Department of Statistics 1991) due to the limited educational facilities, which is explained by the economic development of the nation in the early decades of the 20th century. According to 1991 censes, 63% of the elderly received no schooling at all. However, this represents an improvement in comparison to 1970 and 1980, in which the percentage of those with no education was 75 and 73%, respectively. However, future cohorts of elderly will be more educated due to education became widely available. Among the three major groups, the Indian senior citizens were better educated compared to the other two ethnic groups. About half of the Indians had received some education compared to 34% for Bumiputeras and 42% for Chinese (Ong 2001).

3 CONCEPTUAL FRAMEWORK

Conceptual framework exposes features or requirements that need to be represented in a model. Based on the issues above, we developed a conceptual framework (Fig. 1). In the next sub section we outline the main components of proposed framework.

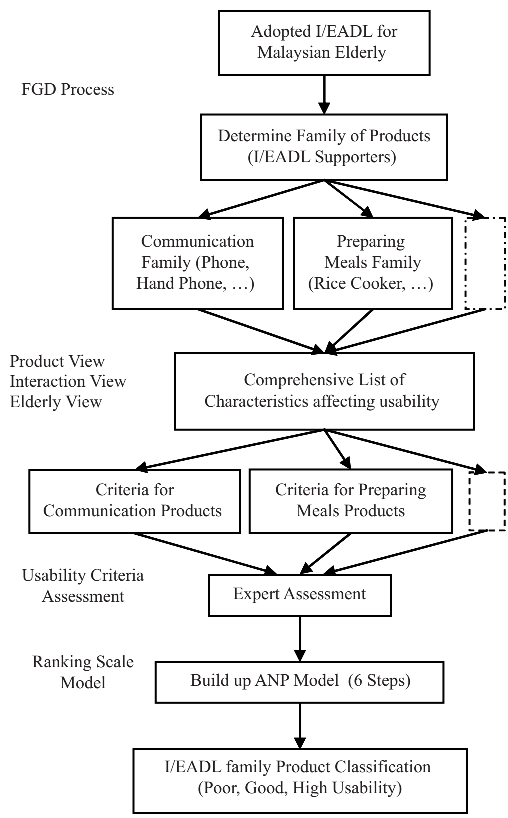


Figure 1. Conceptual framework and links between components.

3.1 *Criteria for recruiting participants*

Due to wide diversity in characteristics of older Malaysians, they must be picked with care to illustrate those characteristics which the researchers believe important. As pointed out by Hardy et al. (2002) elderly have a fundamental distrust and a very limited understanding of the underlying technical concepts, leading a reluctance to experiment. Besides, age related factors can also make inaccurate self-reporting questionnaires (Park 2000).

In this study, a total of 50 elderly aged 60 to 74 years able body from three main ethnic groups (Malay, Chinese and Indian) are recruited purposively on a voluntary basis. The elderly people who are uncommunicative or unwilling to join to study are excluded.

3.2 *Adopted IADL*

To guarantee independent living for elderly people we need to go beyond IADL. We derived list of common daily activities and family of products concerning especially adopted I/EADL with the culture and life-style of elderly Malaysians. There are two definite categories of instrument for elderly people. For each category, different usability criteria must be considered. The first category deals with mostly automatic instruments, basically applicable for care givers. This category can be called passive instruments. The second category, which we focus on, needs the active interaction of the elderly as end users. This second category is called active instruments. The established standards and taking into account needs and opinions of the elderly to define special usability criteria for active instruments will enable them to continue more independent for a longer time along with more quality in life.

3.3 *Initial list of usability criteria*

Generally, extended criteria are needed for the elderly due to their health impairments such as low vision, poor hearing, difficulty with hand tasks, memory difficulties/learning disability, or general weakness compare to normal users. The more important criteria in the aforementioned background are as follows:

- Safety, which concerns whether a product limits the risk of harm to elderly users while interaction;
- Readability, ease with which visual content can be understood;
- Learnability, the extent to which the consumer, upon initially receiving the product, can easily learn to use it and can start using it within a reasonable period of time once assembled, including whether specialized training is required;
- Minimal action, capability of the product to help users achieve their goal in a minimum number of steps;
- Minimal memory load, whether a user is required to keep minimal amount information in mind in order to achieve a specified goal (Lin et al. 1997);
- Trustworthiness, which means the faithfulness that a product offers to its users.

Moreover, the data required to collect extended criteria for Malaysian elderly due to the cultural diversity. Because, users differ across regional, their expectations of each family product make them behave primarily by their local cultural perspectives. Consequently, user reactions become more predictable and understandable when the user's cultural perspective is taken into account (Barber 1998).

3.4 *FGD*

Focus groups are an effective and relatively inexpensive way to achieve people's perceptions and attitudes about any particular product or concept. In this case, there is no direct interaction with the product. All the data from a focus group are in the form of self-reported metrics.

To get more reliable data, selected participants are divided in six FGDs consisting of two Malay groups (Male and Female), two Chinese groups, and two Indian groups, based on more common disabilities by the modified IADL scale. The study continue with related family products in each panel. All FGDs are conducted by a moderator depending on the ethnicity and assisted by two note-takers. It is evident that focus groups are ideal in examining the cultural perspective of participants.

In each focus group, representative participants regarding their experiences and challenges describe products that they have used to support daily activities.

The FGD is only a feedback to the questionnaire and still we need to rank the usability criteria using ANP Model.

3.5 *ANP model*

ANP models have two parts: the first is a control hierarchy or network of objectives and criteria that control the interactions in the system under study; the second are the many sub-networks of influences among the elements and clusters of the problem, one for each control criterion. The criteria obtained in previous part will be the foundation of the ANP analysis method. The second part adopts proper key criteria through experts' questionnaires as the major basis for constructing ANP model.

In summary, following steps will be conducted to build up ANP model:

Step 1: Construction of the network hierarchy;
Step 2: Questionnaire surveys and expert preference;
Step 3: Establishment of pairwise comparison matrixes;
Step 4: Consistency test;
Step 5: Computations of super matrixes;
Step 6: Selection of most optimal options; and to determine which usability criteria contribute more in decision making regarding the best product selection.

The power of the Analytic Network Process (ANP) lies in its use of ratio scales to capture all kinds of interactions and make accurate predictions, and, even further, to make better decisions. So far, it has proven itself to be a success when expert knowledge is used with it to predict sports outcomes, economic turns, business, social and political decision outcomes (Saaty 2006).

4 CONCLUSION

There are a lot of factors affecting usability of product. This paper proposed a conceptual framework for identifying a set of classified usability criteria for each family of products related to I/EADL. Applying ranking method such as ANP for each set of usability criteria would be helpful for manufactures to design product more usable and acceptable by the elderly.

REFERENCES

Arthanat, S. et al. 2007. Conceptualization and measurement of assistive technology usability. *Disability and Rehabilitation: Assistive Technology* 2(4): 235–248.

AXA Retirement Scope. 2008. Retirement Results for Malaysia. *http://www.retirement-scope.axa.com/lib/rs/uploads2/asiepacifique/malaisie/AXA_report_malaysia_2008.pdf*

Barber, W. & Badre, A. 1998. Culturability, the merging of culture and usability. In: *Proc. of the 4th conference on Human Factors and the Web.* NJ: Basking Ridge.

DOSM 2000. *Census of Population and Housing Malaysia*. Kulala Lumpur: Deptartment of Statistics.

DOSM 1991. *Senior Citizens and Population Ageing in Malays: Population Census Monograph Series.* Kuala Lumpur: Deptartment of Statistics.

Gregor, P., Newell, A.F. & Zajicek, M. Designing for dynamic diversity: interfaces for older people; *Proc. of the fifth international ACM conference on Assistive technologies, 2002*. New York: ACM Press: 151–156.

Hardy, M. & Baird, C. 2002. Is it all about aging: Technology and aging in social context. In N. Charness & W. Schaie (eds), *Impact of technology on successful aging*. New York: Springer.

Hawthorn, D. 2000. Possible implications of aging for interface designers. *Interacting with Computers* 12: 507–528.

Holzinger, A. et al. 2008. Investigating usability metrics for the design and development of applications for the elderly. *ICCHP 2008, Springer-Verlag*: 98–105.

Ismail, N. & Gabda, D. 2001. *Social security scheme and the elderly.* In *Social Security for the Elderly.* Kuala Lumpur: Prentice Hall: 77–89.

Jorge, J.A. 2001. Adaptive tools for the elderly: new devices to cope with age-induced cognitive disabilities. *Universal Accessibility of Ubiquitous Computing: Providing for the elderly*; *Proc. of the 2001 EC/NSF Workshop, 22–25 May 2001.* Portugal: Alcácer do Sal. WUAUC'01. New York: ACM Press: 66–70.

Lin, H.X. et al. 1997. A proposed index of usability: A method for comparing the relative usability of different software systems. *Behaviour and Information Technology* 16: 267–277.

Ong, F.S. 2001. Ageing in Malaysia national policy and future direction. *http://adrf.trf.or.th/Archive_data/Aging-in -Malaysia.doc.* Accessed on 15 Dec 2008.

Park, D. & Schwarz, N. 2000. *Cognitive Aging: A Primer*. In Psychology Press: 238. Hove: Taylor and Francis Group.

Saaty, T.L. 2006. *Decision Making with the Analytic Network Process.* New York: McGraw Hill.

Syariffanor, H. 2006. Localising the User Interface for Elderly People in Malaysia. *http://www-edc.eng.cam.ac.uk/~jag76/hci_workshop06/hisham.pdf*. Accessed on 20 Nov 2008.

Tinker, A. 1997. *Older people in modern society*. London: Longman.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Design better user experience for PPE: Thinking and recommendations

Yajun Zhang & Zhaoxia Jin
Honeywell Integrated Technology Co. Ltd., China

ABSTRACT: It is well known that Personal Protective Equipment (PPE) should not just design to be functional but also should be proactively designed to be comfortable, fit and providing better user experience to meet the needs of variety of users. In this paper, we have attempted to summarize and share some findings and insights on key aspects of PPE user experience. Some typical examples of problems and interesting findings of using Personal Protective Equipment were revealed from our product evaluation studies and design practices. In addition, we further discussed our suggestions and recommendations on these key user experience aspects. The findings and recommendations disclosed in this paper are not only limited to the product type, but also applicable to many other products.

Keywords: PPE, user experience, comfort, fit

1 INSTRUCTION

Personal Protective Equipment (PPE), which includes variety of devices and garments such as gloves, hard hat, goggles, vest, earplugs and respirators, is designed to protect employees from serious workplace injuries or illnesses resulting from contact with chemical, radiological, physical, electrical, mechanical, or other workplace hazards (OSHA, 2002).

There is a growing awareness that the products should not just design to be functional, meeting safety standard, but also should be designed to be comfortable, fit and providing better user experience by meeting the various local needs. Also as Abeysekera (1992) reported, needs related to user experience of PPE contain: 1) control of additional physiological loads, 2) thermal comfort, 3) well-fitting and adjustable, 4) minimum the interference of sensations, 5) freedom for movement, 6) maintain work performance, 7) wearing more than one PPE, 8) ease for putting on and taking off, and 9) good aesthetics design. To design for better user experience, it is critical to know the common issues that users may be facing while using PPE and its implications. In this paper, with the intention of sharing some findings and insights on designing better user experience, some typical examples of problems and interesting findings of using Personal Protective Equipment were revealed and discussed from our product evaluation studies and design practices. In addition, we have discussed our suggestions and recommendations on these key user experience aspects.

2 TYPICAL PPE USER EXPERIENCE ISSUES AND SUGGESTIONS

2.1 *Defining target user*

Various characteristics of the target user, e.g., race, age, sex, body profile, will determine many of the properties of the products. For instance, Lili Du and Ziqing Zhuang (2008) found Chinese civilian adults have shorter face length and nose protrusion and larger face width and lip length compared to American, which means a face piece respiratory mask which fits Caucasian may not be suitable for Chinese. This kind of knowledge should be gained before design begins and needs to be kept in mind during whole development process.

Suggestion: An early step of any design is to find out who you are designing the product for. Understanding your end users and their characteristics is crucial as it can make a huge impact on designs. User studies and voice of customer studies are suggested to be carried out to know more and deeper about your target users. To better formulate these understandings, and carry the knowledge and information further for design and development, you can create the user personas. Persona is an archetype of a user and is described in terms of needs, goals and tasks (Kari et al. 2004). In this way it gives the 'precision' required in design work. In most cases, personas are synthesized from data collected from interviews with users. For each major user group it is better to develop a separate persona.

2.2 *Ill-fitting*

Ill-fitting PPE is a common issue—either too tight or too loose. Some products are too loose for some populations, e.g., Chinese. Imagine a kid is trying to put adult's clothes on—everything is way too big and too loose. This problem is very easy to understand since there is obvious body profile difference between different populations, e.g., Asian vs. Caucasian. Yang and Shen (2008) reported both Chinese genders (270 males and 191 females) had a significantly smaller anthropometric value of facial width than those of American.

Fit properly means neither too tight nor too loose.

Taking gloves for example, gloves should be the right size, and fit comfortably—too tight may increase the load on the hand and will result in fatigue and loss of motion range, while too loose will loose dexterity and impede work. Another example is footwear. When shoes are too tight, they will cause feet swelling, pain and discomfort. When they fit loosely, they will cause slip, blisters. Both situations will make the footwear cumbersome and potentially cause injury beyond design intention.

Suggestion: Design for fit which means the PPE product should not impose any obstruction to the user in the execution of their tasks, nor discount the product's performance, and without any additional risks being generated (Adalbert Pasternack, 2001). To achieve this goal, anthropometric data need to be leveraged and well considered during design and development. The relevant publications and databases need to be consulted to ensure the correct critical dimensions. In addition, anthropometric data cannot always be applied directly to the design of the products. Further calculation and certain level of adaptation are usually needed, supported by generally applicable measurement methods.

2.3 *Incorrect position*

Another problem which is closely associated with fit is sometimes key components do not sit on the right position of users' body even though the overall size of the products is snuggly fit. One of the typical examples of incorrect positioning is the matching between the supporting area and relative body area. In the case of waist mounted product like Powered Air Purifying Respirator (PAPR), the waist support of some products sit lower than participants' waist. If the waist support is not at the correct position, it won't achieve the design intention—to evenly distribute the load around the waist and provide more comfort. Another typical example is some key safety components like back D-rings for harness were much lower than the standard position, which could cause great danger due to safety standards (shown in Figure 1).

Suggestion: To modify the product size to fit another population by simply scaling up and down all dimensions or total size of a product as a whole could result in some awkward issues, e.g., positions of some key components may be moved to a wrong place. To avoid this issue, it is suggested to define the reference points first before starting modifying the size of the product. Usually these points will be in related to body measurement landmarks (key body features, bones or bone joints).

2.4 *Poor fit effectiveness*

Fit effectiveness means that the desired population is accommodated without wasted sizes or wasted accommodation regions (Robinette, Kathleen M. 2007). Sometimes, strap lengths for a PPE product could be way too long for some users or populations, which lead to the excessive straps swinging around, shown in Figure 2. Even if belt keeper was provided by some products, the excessive straps could be still way too long to be managed well.

Suggestion: Fit mapping to target anthropometric data usually can also help to exam the fit effectiveness. To achieve good fit effectiveness, wasted sizes or adjustment ranges need to be dropped. Further providing different sizes of a product is another common way to improve the fit effectiveness.

2.5 *Comfort*

There are many factors that will contribute to discomfort. One typical and simple example is the

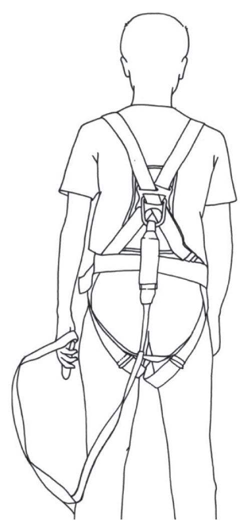


Figure 1. Incorrect position of back D-ring.

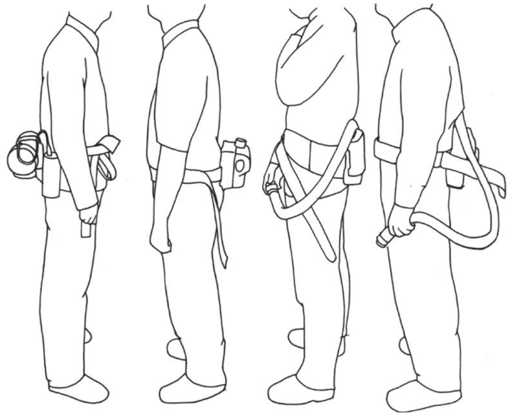


Figure 2. Poor fit effectiveness of waist belt.

weight and weight distribution. There might be differences between westerners and Asians in terms of comfort criteria, such as acceptable weight as we mentioned acceptable pressure distribution, smell, comfortable thermal-humid ranges etc. For example, Swei-Pi Wu (1997) reported the average Maximum Acceptable Weight of Lift (MAWL) values of Chinese subjects were smaller than those of occidental subjects.

Suggestion: Discomfort does not only affect the performance and decrease the productivity, but also create hazards greater than the original hazards which the PPE is expected to protect the user (Akbar-Khanzadeh and Bisesi, 1995; Roughton and Mercurio, 2002). Therefore, great attention should be paid to comfort aspect in the design and development process. To design comfortable products needs to fully understand the comfort criteria of target population and then design to meet these criteria. In addition, various assessments and evaluations can be conducted iteratively during the whole development process to identify and address discomfort issues and eventually guarantee a comfortable product.

2.6 *Context of use*

As far as the conditions of workplace are considered, the required solution may be totally different from one region to another. Considering the usage environment can be much different among regions, some issues and risks mentioned above may cause even bigger potential impact and dangers for example, the electricity power distribution pole in some areas can be very chaos with many things here and there (as shown in Figure 3); in this case the excessive straps swing around may cause more potential danger. Furthermore, Users from some regions like China usually wearing lots

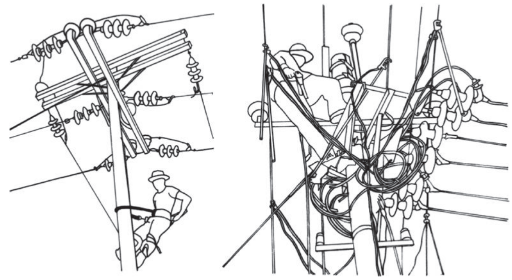


Figure 3. Ordered electricity distribution pole and chaotic electricity distribution pole.

of other accessories since one person usually will perform multiple tasks in field, which will even increase the potential risks.

Suggestion: Before the product development process starts, it is of extreme importance to fully understand the context of use i.e., the type of activity, the conditions and risks prevailing at the workplace. As defined by the standard (ISO 1997): the Context of Use consists of the users, tasks and equipment (hardware, software and materials), and the physical and social environments in which a product is used.

The analysis of the Context of Use will 1) Provides an understanding of the circumstances in which a product will be used. 2) Helps to identify user requirements for a product. 3) Helps address issues associated with product usability. 4) Provides contextual validity of evaluation findings (Martin Maguire 2001). To provide products that can work well with given context usually needs in-depth user studies and field research. When these have been conducted, some innovative and unique solutions may just come along the way.

3 SUMMARY

The aforementioned findings are just typical observations extracted and derived from PPE product evaluation studies and design practices. In the design and development process, it is critical to identify and solve these issues with right approaches. Viewing these problems individually could not ensure the designed products have satisfied user experience. It is suggested that the product manufacturer must have taken these aspects (comfort, fit and user experience) into account and incorporated them into the product design process systematically—e.g., set the user experience target, define related activities, control and measure to finally meet the initial goal.

To summarize, this paper introduced some typical examples of problems and interesting findings of using Personal Protective Equipment derived from our product evaluation studies and design practices. In addition, we have discussed our suggestions and recommendations on these key user experience aspects. The findings and recommendations disclosed in this paper are not only limited to the product type but also applicable to many other products especially in close relation to human body size and fit.

REFERENCES

Abeysekera, J. 1992. *The use of personal protective clothing and devices in the cold environment*. Preliminary study report, Division of Industrial Ergonomics, Luleå University, Sweden.

Du, L., Zhuang, Z. et al. 2008. Head-and-face anthropometric survey of Chinese workers. *Ann. Occup. Hyg.* 52(8): 773–782.

ISO 1991. *ISO 9126: Software product evaluation-2quality characteristics and guidelines for their use.* Geneva: International Standards Organisation. Also available from the British Standards Institute, London.

ISO 1997. *ISO 9241-11: Ergonomic requirements for o.ce work with visual display terminals (<D¹s). Part 11~guidelines for specifying and measuring usability*. Geneva: International Standards Organisation. Also available from the British Standards Institute, London.

ISO 1999. *ISO 13407: Human-centred design processes for interactive systems*. Geneva: International Standards Organisation. Also available from the British Standards Institute, London.

Lee Au, E.Y., Jin, Z.X. & Landry, M. 2009. Design personal protective equipment for productivity, comfort and fit: Challenges and opportunities. *IEA 2009 conference.*

Maguire, M. 2001. Context of use within usability activities. *Int. J. Human-Computer Studies* 55: 453–483.

Occupational Safety and Health Administration 2002, OSHA fact sheet: personal protective equipment. Available URL: *http://www.osha.gov/OshDoc/data_General_Facts/ppe-factsheet.pdf*

Pasternack, A. 2001. Ergonomic aspects of personal protective equipment, 87 Dräer Review 8/01.

Roark, J. 2004. The proper fit for PPE, Available URL: *http://www.infectioncontroltoday.com/articles/407/407_4b1topics.html*

Rönkkö, K., Hellman, M., Kilander, B. et al. 2004. Personas, participatory design and product development: An infrastructure for engagement. *Proceedings Participatory Design Conference 2004*, Toronto, Canada.

Robinette Kathleen, M. et al. 2007. Maximizing anthropometric accommodation and protection. Air force research laboratory human effectiveness directorate biosciences and protection division biomechanics branch, final report for April 2002 to August 2007.

Wu, S.P. 1997. Maximum acceptable weight of lift by Chinese experienced male manual handlers. *Applied Ergonomics* 28(4): 237–244.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A literature survey of conceptualizing user experience

Wonkyu Park, Sung H. Han, Jaehyun Park, Hyun K. Kim, Yong S. Park,
Youngseok Cho, Jaemin Chun & Seunghwan Oh
Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea

ABSTRACT: In recent years, Human-Computer Interaction (HCI) researchers have been interested in User Experience (UX). UX is briefly defined as all kinds of experiences gained when users interact with products/services, but it is difficult to find a specific definition that is widely accepted in HCI area. We conducted a comprehensive literature review to define UX concept concretely. A total of 247 literatures were collected from various sources. As a result, several attributes of UX were identified. The UX keeps changing as time goes by, and it covers affect and user value as well as usability. Also, several factors such as users' internal states, products' characteristics and environmental context around users influence the UX.

Keywords: user experience, literature survey, concept of UX, element of UX, trend of UX study

1 INTRODUCTION

In recent years, Human-Computer Interaction (HCI) researchers have been interested in User Experience (UX) (Law et al. 2009, Norman et al. 1995). The international organization for standardization (ISO 2009) defines UX as 'the overarching experience a person has as a result of their interactions with a particular product or service, its delivery, and related artifacts, according to their design'. The user experience includes all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur before, during and after use. This concept has been popularly studied from different perspectives: marketing, branding, visual design, and usability, but the exact definition, framework, and elements of user experience are still evolving (Wikipedia 2010b).

In this study, we conducted a comprehensive literature review to analyze the trend of UX research and define the concept of UX concretely. Since there are many different concepts in various research fields, for example in HCI, psychology, marketing and business administration, those concepts should be integrated.

2 LITERATURE SURVEY

2.1 *Collection of literature*

The literature investigating research trends and definitions of UX covers various research fields such as ergonomics, cognitive psychology, computer science, information engineering, quality engineering, design, marketing, and business administration. This literature was collected from journals, proceedings, dissertations, magazines, reports, news, books and web articles. Several keywords were used to retrieve data on the web. The keywords were user experience, usability, ease of use, user interface, affect engineering, human-computer interaction, human-machine interaction, technology acceptance model, brand image, brand equity and product life cycle.

A total of 274 articles were collected at the beginning. Most articles were proceedings papers (101) and journal papers (66).

2.2 *Selection of literature*

There were too many articles to summarize them. Furthermore, most didn't have detailed information about UX concepts. Thus, we selected only those which have concrete contents related to definitions, components and methods of UX. They were categorized into three groups: high, moderate and low usefulness. Only 38 of the 274 papers were considered to have high usefulness, while 89 had moderate usefulness.

2.3 *Analysis of literature*

We collected detailed information for the 127 papers rated as high and moderate. The information collected was author, publication year, nationality of institute, UX definition, UX components, UX measurement methods, institution type (e.g., university or company), and literature type (e.g., dissertation, proceeding paper or journal article). In addition, the major field of study, affiliation and nationality of authors were also investigated.

3 RESULTS

3.1 *Research trends of UX*

3.1.1 *Publication year*

As shown in Figure 1, more and more studies have been published since 2000. In particular, UX studies increased dramatically after 2006. Since we just collected literature which helps to conceptualize UX, more studies would be related to UX besides these.

3.1.2 *Institution*

UX research has been conducted at universities, by industry and through industry-universities cooperation. Most literature came from universities. There were 81 articles from universities, 25 articles from industry, and 16 from industry-universities. In addition, 5 articles came from others, like a government institution, research center or educational institution.

Nokia (9), IBM (6), University of Art and Design Helsinki, University of Oulu, Delft University of technology were the top institutions for UX study. Nokia, IBM, Google and Intel published about 84% literature among industry, so other companies are thought to be weak at UX study.

3.1.3 *Region*

Most literature has been published in North America (56) and Europe (47). Only 13 articles from Asia were collected. In North America, both USA and Canada actively studied UX. In Europe, Finland, Germany, England, Netherland and France were the main countries. In Asia, most research was conducted in Korea, China, Singapore and Japan.

The distribution of institution type in each region is showed in Figure 2. The ratio of industry

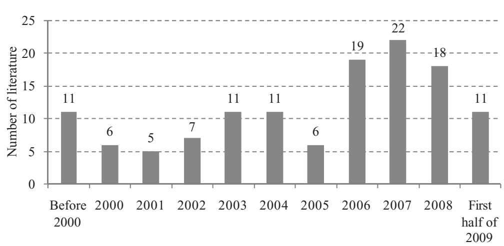


Figure 1. Distribution of literature by publication year.

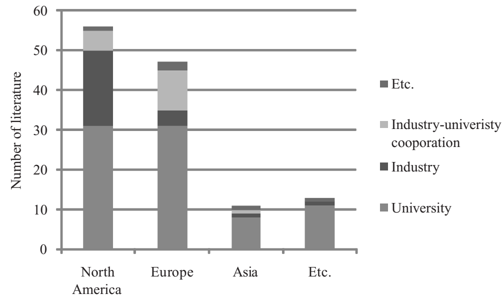


Figure 2. Distribution of institution type by region.

is high in the North America compared with the other regions, since there are many consulting firms and IT companies in North America. The higher ratio of industry-universities cooperation in Europe is primarily due to collaborative research by Nokia and universities in Finland.

3.1.4 *A major field of study*

In order to analyze authors' major field of study, we conducted an additional survey of the 240 authors. First, we classified research domains into six major fields such as ergonomics (including HCI and industrial engineering), design (including industrial design and interface design), computer science, information engineering, psychology and business administration (including marketing and advertising). As shown in Figure 3, the major fields of authors were distributed evenly. Business administration was the most popular in UX study because it includes various other fields such as a marketing, advertising and economics. Psychology was relatively rare, and psychologists mainly studied about technology acceptance models while economists studied about brand equity.

As shown in Figure 4, the ratio of authors' background varied in each region. The ratio of business administration and psychology was high in the North America, but the ratio of ergonomics and design was high in the Asia. On the other hand, there was no particular characteristic in the Europe.

3.2 *Concepts of UX*

Various concepts of UX have been defined by many researchers (Law et al. 2009, Roto 2006). They suggested different concepts based on their backgrounds. The typical concepts are as follows.

Level of UX changes continuously as time goes by (Karapanos et al. 2009). That is, users' perceptions of products/services are dependent on phases of product experience. Hassenzahl & Tractinsky

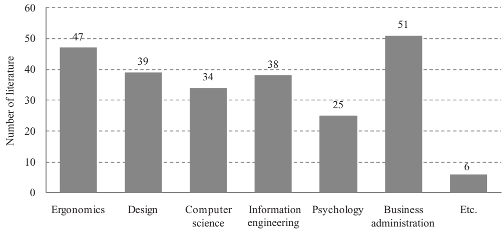


Figure 3. Distribution of literature by major fields of study.

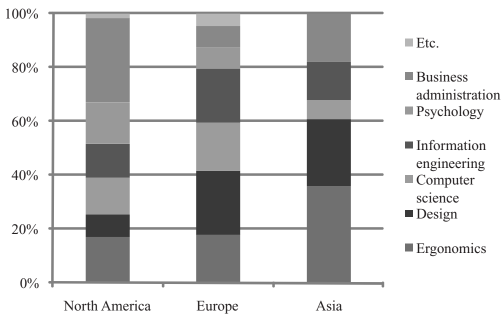


Figure 4. Ratio of authors' background in important regions.

(2006) defined the UX as 'a consequence of a user's internal state, the characteristics of the designed system and the context within which the interaction occurs'. Hekkert (2006) noted that the product experience includes 'the degree to which all our senses are gratified (aesthetic experience), the meanings we attach to the product (experience of meaning) and the feelings and emotions that are elicited (emotional experience)'. Marcus (2006) found that experience covers 'all stakeholder touch-points such as the places where a buyer, customer, user, learner, expert, advocate, staff member, journalist, or investor comes into contact with the product or service, or its sponsoring company/organization'. Rust et al. (2004) didn't mentioned UX directly, but they found that the brand equity of products has a critical effect on customer equity. Not only the brand equity, but also value equity and relationship equity influence the customer value. Arhippainen & Tähti (2003) focused on particular conditions, consisting of social factors, cultural factors and context of use. Finally, Mäkelä & Suri (2001) pointed out that the UX is 'a result of a motivated action in a certain context'. Also, at a point of time, the user's previous experiences and expectations influence the present experience, and the present experience leads to more experiences and modified expectations.

4 DISCUSSION

A variety of different concepts were integrated and key attributes were identified to specify UX. At first, we defined the scope of the experiences, and then came up with elements that consisting UX.

4.1 *Scope of UX*

Experiences are organized in a hierarchical structure composed of three levels of user interactions. As present in Figure 5, these are experience, brand experience and product/service experience.

The highest level experience, includes the other levels (Law et al. 2009). The experience consists of knowledge of, skill in, or observation of something or event (Wikipedia 2010a). This can occur without direct interactions with products. For example, communicating with friends, religious activity, exercising and making a cup of coffee are kinds of experiences (Forlizzi & Ford 2000).

Brand experience means all kinds of experiences gained when users contact a particular brand. A brand can be separated into corporate, family and individual brands (Farquhar et al. 1992). For example, Microsoft is a corporate brand, Windows is a family brand and Windows 7 is an individual brand. Law et al. (2009) said that it includes interactions with a product/service having a brand, other products/services of the brand and company itself. All the brand experiences occur whether users interact with products/services directly or not, because this can be gained by contacting the corporation itself, mass media or other users. For example, corporations give customers a good brand image through frequently showing advertisements about their ethics. Thus, customer's brand experiences are usually influenced by the product/service experience, loyalty, brand awareness, attitude to brand or brand ethics (Aaker 1994, Keller & Lehmann 2006, Krishnan & Hartline 2001, Yoo & Donthu 2001).

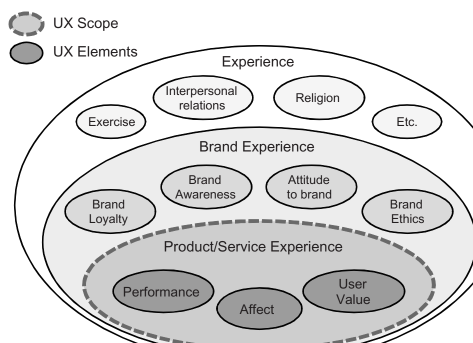


Figure 5. Scope of experiences.

Product/service experience is gained when users interact with a product/service or its group. This experience is expressed in three different ways: how easy it is to use, how good it makes user feel and how much it improves customer's value (Davis & Wiedenbeck 2001, Desmet & Hekkert 2007, Hassenzahl 2001, Hekkert 2006, Karapanos et al. 2009, Marcus 2006, Roto 2006, Yamazaki & Furuta 2007). Also, most researchers supposed that UX comes from interactions between user and a product/service (Arhippainen & Tähti 2003, Hassenzahl & Tractinsky 2006, Hiltunen et al. 2002, Mäkelä & Fulton Suri 2001, Marcus 2006, Roto 2006, Yamazaki & Furuta), so this level of experience can be considered UX.

4.2 *Elements of UX*

There are three elements in the UX (those are performance, affect and user value). First, performance is the extent to which customers are able to use products/services effectively and efficiently (Hiltunen et al. 2002, Yamazaki & Furuta 2007). Since it focuses on the functionality of the product/service, it can easily be evaluated. The literature survey shows performance has usually been studied in ergonomics, design and information engineering.

Secondly, affect is user's subjective satisfaction or feelings about the product's aesthetic such as the appearance of interface and kindness of service (Agarwal & Meyer 2009). Many affective engineering researchers have studied it (Desmet & Hekkert 2007, Hassenzahl 2001).

Finally, user value is all kinds of values that users have when they use a product/service (Rust et al. 2004). Since the user value contains a lot of sub-elements, there is no systematic method to evaluate it yet. However, it is an important element as a key criterion to divide UX from the usability concept (Hoeffler & Keller 2002, Rust et al. 2004). Sub-elements of the user value include sense of superiority, trust on products and self-fulfillment (Hartmann 2006, Karat et al. 2003, Kim & Moon 1998, Swallow et al. 2005).

5 CONCLUSION

In this study, we collected relevant literature and analyzed it to define the concepts and elements of UX. At first, 247 articles were collected from various sources, and then 127 important articles were selected for detailed analysis. Many studies have been conducted in different institutions, regions and fields. We could integrate various aspects of UX and summarized key attributes considered by many researchers. In Particular, UX was composed of three specific elements: usability, affect and user value.

This study contributes toward conceptualizing UX definition specifically. It is expected that product/service designers will enhance UX level by considering the elements of UX in detail. For future work, the UX scope and elements should be verified, and a user survey would be a good way to do.

ACKNOWLEDGEMENT

This work was supported by Mid-career Researcher Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (MEST) (No. 2010-0000364).

REFERENCES

Aaker, D.A. 1994. Should you take your brand to where the action is? *Harvard Business Review* 135–143.

Agarwal, A. & Meyer, A. 2009. Beyond usability: evaluating emotional response as an integral part of the user experience. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*.

Arhippainen, L. & Tähti, M. 2003. Empirical evaluation of user experience in two adaptive mobile application prototypes. In *Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia MUM 2003*.

Davis, S. & Wiedenbeck, S. 2001. The mediating effects of intrinsic motivation ease of use and usefulness perceptions on performance in first-time and subsequent computer users. *Interacting with Computers* 13: 549–580.

Desmet, P. & Hekkert, P. 2007. Framework of product experience. *International Journal of Desig 1*(1): 57–66.

Farquhar, P.H., Han, J.Y., Herr, P.M. & Ijiri, Y. 1992. Strategies for leveraging master brands: How to bypass the risks of direct extensions. *Market Research 4*(3): 32–50.

Forlizzi, J. & Ford, S. 2000. The building blocks of experience: an early framework for interaction designers. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*.

Hartmann, J. 2006. Assessing the attractiveness of interactive systems. In *CHI '06 extended abstracts on Human factors in computing systems*.

Hassenzahl, M. 2001. The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interactio 13*(4): 481–499.

Hassenzahl, M. & Tractinsky, N. 2006. User experience-a research agenda. *Behaviour & Information Technology 25*(2): 91–97.

Hekkert, P. 2006. Design aesthetics: Principles of pleasure in design. *Psychology Science 48*(2): 157.

Hiltunen, M., Luomala, J. & Laukka, M. 2002. *Professional Mobile User Experience*: IT Press.

Hoeffler, S. & Keller, K. 2002. Building brand equity through corporate societal marketing. *Journal of Public Policy & Marketin 21*(1): 78–89.

ISO 2009. *Human-centred design for interactive systems: Ergonomics of human system interaction* (Vol. ISO FDIS 9241-210). Switzerland: International Organization for Standardization (ISO).

Karapanos, E., Zimmerman, J., Forlizzi, J. & Martens, J.-B. 2009. User experience over time: an initial framework. In *Proceedings of the 27th international conference on Human factors in computing systems.*

Karat, C., Brodie, C., Karat, J., Vergo, J. & Alpert, S. 2003. Personalizing the user experience on ibm. com. *IBM Systems Journal 42*(4): 686–701.

Keller, K. & Lehmann, D. 2006. Brands and branding: Research findings and future priorities. *Marketing Science 25*(6): 740.

Kim, J. & Moon, J.Y. 1998. Towards emotional usability in customer interfaces: Trustworthiness of cyber-banking system interfaces. *Interacting with computers 10*: 1–29.

Krishnan, B. & Hartline, M. 2001. Brand equity: Is it more important in services? *Journal of Services Marketing 15*(5): 328–342.

Law, E.L.-C., Roto, V., Hassenzahl, M., Vermeeren, A.P.O.S. & Kort, J. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the 27th international conference on Human factors in computing systems*.

Mäkelä, A. & Fulton Suri, J. 2001. Supporting Users' Creativity: Design to Induce Pleasurable Experiences. In *International Conference on Affective Human Factors Design*, Singapore.

Marcus, A. 2006. Cross-cultural user-experience design. *Diagrammatic Representation and Inference*, 16–24.

Norman, D., Miller, J. & Henderson, A. 1995. What you see, some of what's in the future, and how we go about doing it: HI at Apple computer. In *Proceedings of CHI 1995*, Denver, Colorado, USA.

Roto, V. 2006. *Web browsing on mobile phones-characteristics of user experience.* Unpublished Dictoral Disseration, Helsinki University of Technology, Espoo, Finland.

Rust, R., Zeithaml, V. & Lemon, K. 2004. Customer-centered brand management. *Harvard Business Review 82*(9): 110–120.

Swallow, D., Blythe, M. & Wright, P. 2005. Grounding experience: relating theory and method to evaluate the user experience of smartphones. In *Proceedings of the 2005 annual conference on European association of cognitive ergonomics*.

Wikipedia. 2010a. Experience. retreived from *http://en.wikipedia.org/wiki/Experience*

Wikipedia. 2010b. User experience. retreived from *http://en.wikipedia.org/wiki/User_experience#cite_note-2*

Yamazaki, K. & Furuta, K. 2007. Design tools for user experience design. In *Human-Computer Interaction. Interaction Design and Usability*, 298–307.

Yoo, B. & Donthu, N. 2001. Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research 52*(1): 1–14.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

An in-depth interview for conceptualizing user experience in the Korean mobile phone industry

Hyun K. Kim, Sung H. Han, Jaehyun Park, Youngseok Cho & Yong S. Park
Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea

ABSTRACT: Many studies related to User Experience (UX) have been conducted over the past few years. However, a consensual definition of UX has yet to be developed. In this study, an in-depth interview was conducted with experts in the Korean mobile phone industry in an effort to conceptualize the typical user experience in this industry. The study gathered various opinions regarding the UX concept from industry experts and integrated different UX definitions from previous research to reflect the views of the industry. Korean mobile manufacturing companies, mobile communications companies, and internet service companies were the targets of the interview in this study. Additionally, several questions which were controversial in the previous studies were posed so that the interviews would be more in-depth. As a result of this study, UX is conceptualized by collecting opinions of experts and by integrating the perspectives of industry and academia.

Keywords: user experience, in-depth interview, conceptualization of ux

1 INTRODUCTION

User experience refers to the various experiences that a person can have when he/she interacts with a product (Alben 1996, Arhippainen & Tähti 2003, Desmet & Hekkerrt 2007). Over the last decade, UX concepts have been widely studied in diverse domains such as human-computer interaction, industrial design and computer science engineering (Law et al. 2009). However, the concept of UX remains a controversial issue, with different opinions regarding its definition and viewpoints (Law et al. 2008).

Many recent UX studies have focused on defining and understanding the UX concept. Mäkelä & Fulton Suri (2001) focused on the time aspect of UX. They posited that previous experience affects current experience and that current experience affects future experience again. Hassenzahl & Tractinsky (2006) classify UX concept in terms of specific perspectives and discuss an integrated UX concept. UX is the result of a user's internal state and characteristics of the system in the context where interaction occurs. Law et al. (2009) survey several researchers and practitioners to analyze the UX concept. In this process, various definitions of UX from earlier studies were classified into five categories and the types of preferences were surveyed.

The UX definition was analyzed from integrated views in previous research, but few studies that deal with a concrete scope, object and element of UX have taken place. Moreover, academia and industry have different perspectives on UX (Väänänen-Vainio-Mattila et al. 2008). The academic research focuses on UX theories, models and frameworks while the industrial research focuses on how UX can be applied practically in the product development stage. However, many UX studies have been conducted from the perspective of academia, while relatively few come from industry.

This study conducted in-depth interviews with UX experts who work for the Korean mobile phone industry. The interview results were used to integrate views of academia and industry regarding UX concepts and to investigate applications of UX in the existing Korean mobile companies.

2 THE IN-DEPTH INTERVIEW PROCEDURE

2.1 *Selecting the interviewee*

The in-depth interviews were conducted with UX experts from the Korean mobile phone industry.

The mobile phone industry area was selected in that it is the most active industry of UX research. It is also considered as an important factor that organizations related to UX exist in most Korean mobile phone companies and that the companies actually carry out tasks relevant to UX through these organizations. By collecting expert opinions at the actual work-site, this research focuses on supplementing the previous body of research as regards the definition of UX through a literature survey and attempts to understand the practical use of the UX concept in industry

2.2 *Organization of the questionnaire*

A detailed questionnaire was prepared to conduct in-depth interviews with UX experts at their actual work-site. It consisted of three parts: an explanation of the UX concept, a collection of expert opinions, and the application and perception of UX in the company. In the explanation section, various definitions of UX concepts were introduced by summarizing the findings of previous related research. In the expert opinion section, fifteen questions which are controversial topics in the current UX research were asked. The questions were related to the definitions, scopes, elements and measurements of UX (shown in Table 1). Lastly, in the application and perception part, information about the companies' perception and application of UX was collected in the form of seven open-ended questions.

Additionally, in the expert opinion section, thirteen out of fifteen questions were implemented as a quantitative evaluation. The interviewees were asked to indicate a degree of agreement to the hypotheses on a five-point scale (1: totally disagree, 2: partly disagree, 3: neither agree nor disagree, 4: partly agree, 5: totally agree). For example, with the first hypothesis that 'service experience as well as product experience belongs to UX', an interviewee should give 1 point if he or she strongly disagreed with this hypothesis. They were also given the opportunity to present reasons why they had each opinion.

3 RESULTS AND DISCUSSIONS OF IN-DEPTH INTERVIEW

3.1 *Participants*

Fourteen workers participated in the interviews. The interviewees worked for eight different companies, including cell-phone manufacturing, mobile telecommunication and internet service companies. They were industry experts who dealt with tasks related to the user interface or to UX in the company. The various opinions from industry experts were reflected in the result by the casting of one to two experts from each company.

On average, the interviewees worked for six years (SD = 2.93) in their UX-related field. They had a diverse education background. Six interviewees were originally educated in the field of industrial engineering; four had an education in design, one in computer science, one in cognitive engineering, and two in other areas. Moreover, most had a master's degree or higher. Three interviewees

Table 1. Interview questions and quantified results.

Interview questions (13 hypotheses and 2 open-ended questions)	Avg.*	Std.*	95 CI*
UX definitions and scopes (Q1)			
Q1-1. Service experience as well as product experience belongs to UX	4.79	0.43	4.56~5
Q1-2. The user can gain UX without interacting with a product	3.21	1.42	2.47~3.96
Q1-3. A nonprofit product brings UX	4.86	0.36	4.67~5
Q1-4. UX occurs due to the user's relationship with a corporation	4.21	0.89	3.75~4.68
Q1-5. UX exists before the actual use of a product	4.14	1.03	3.60~4.68
Q1-6. Promotional activities before a product launch bring UX	3.79	1.37	3.07~4.50
UX elements (Q2)			
Q2-1. Usability of a product and user's affect influence UX	4.79	0.43	4.56~5
Q2-2. The user's subjective value toward a product influences UX	4.50	0.65	4.16~4.84
Q2-3. Information given by other users influences UX	3.93	1.00	3.41~4.45
Q2-4. Previous experience with similar products influences UX	4.79	0.43	4.56~5
Q2-5. Culture background of the user influences UX	4.86	0.36	4.67~5
Q2-6. The three most important factors influencing UX (Open-ended)	–	–	–
UX measurements (Q3)			
Q3-1. UX is possible to be quantified numerically	2.64	0.93	2.16~3.13
Q3-2. The long term UX is more important than temporary feelings	2.86	1.70	1.96~3.75
Q3-3. Measurement method for 'Q2-6' (Open-ended)	–	–	–

* Avg., Std., and 95 CI mean average, standard deviation, and 95% confidence interval.

had a bachelor's degree, nine had a master's degree, and two held a doctorate.

3.2 *The definition on UX*

UX remains a controversial topic, and many researchers hold different opinions about the scope and elements of UX. In-depth interviews were conducted here to gather views on UX from an industry perspective. Specifically, the results were analyzed in terms of five aspects: service experience, brand experience, UX lifecycle, UX elements, and a quantification of UX. Table 1 shows the fifteen hypotheses for the quantitative and the qualitative analyses.

First, in terms of the service experience, most experts considered that UX can be formed using both services and products (Q1-1 in Table 1). Several studies which consider services as the object of UX support this result (Roto 2006, Shedroff online, UPA 2006).

Second, many experts agreed on the fact that UX can occur with the product brand (Q1-4 in Table 1). From the interview, some experts reported that factors such as brand image and brand philosophy can have a direct impact on UX or can build a desire for a purchase. On the other hand, there was an opinion that experience that does not come from products or services are out of the UX scope. This was also mentioned in prior studies (Karapanos et al. 2009, Law et al. 2009). By putting all of the ideas together here, it is proposed that the UX scope should be limited to the product/service experience and that the impact of the brand should be considered while analyzing and evaluating UX.

Third, the interviewees stressed that UX can occur throughout the entire product/service lifecycle and that the product lifecycle covers the awareness of the product as well (Q1-5 in Table 1). Moreover, there was a lack of consensus regarding the hypothesis that UX can occur before the launch of a new product (Q1-6 in Table 1). Interviewees reported that a UX evaluation before the actual use of a product is not reliable because the company has a tendency to show what they want to emphasize. Hence, consideration of the product lifecycle aspect of UX is necessary in a future study. Many previous researchers remarked on the temporal aspect of UX as well (Karapanos et al. 2009, Yamazaki & Furuta 2007).

Fourth, many researchers had a skeptical response to the quantification of UX (Q31 in Table 1). They reported it is difficult to integrate many dimensions of UX and it is also difficult to interpret the result from such an integration process. However, if the problem of how to give weight to each dimension can be solved and if a concrete meaning of a quantified UX value can be defined, quantification of UX is possible in the future. Moreover, some participants emphasized the importance of a qualified evaluation as well as a quantified evaluation.

Lastly, the experts selected usability, affect, user value and previous experience as important UX elements. In the results, usability was selected by fourteen interviewees, affect by ten, user value by eight and previous experience by eight interviewees. Usability and affect were also regarded as important criteria for a comparison and evaluation of a product and its related service in several previous studies (Desmet & Hekkert 2007, Hassenzahl 2001, Karapanos et al. 2009). In the case of user value, there was an opinion that it is related to brand value and is an important factor in determining a user's subjective satisfaction. Moreover, some participants commented that previous experience is also an essential factor when considering an actual case. For example, users who become accustomed to a specific text input method on their mobile device prefer this familiar input method, even when a better typing method is developed. These four elements should be considered carefully when developing UX evaluation methods.

3.3 *UX applications in a company*

As UX is becoming a significant issue, attention to UX is also increasing rapidly in the industry. This research investigated the perception and application of UX and the companies' organizations related to UX in the Korean mobile phone industry. Seven questions were asked to participants. Table 2 shows these questions.

First, UX was not defined concretely in the companies. Only a few companies had a definition of UX, but they did not distinguish UX from

Table 2. Seven questions about UX applications in companies.

Questions
Q1. How is UX defined in your company?
Q2. How important is UX in your company?
Q3. Are there any UX guidelines for product/service development as references?
Q4. Do they evaluate the overall UX level after the development of new products or services?
Q5. What is the approximate size of the UX group or department in your company?
Q6. What is the relationship between the UX group/department and the product/service development group in your company?
Q7. What are the educational backgrounds and previous career paths of the UX designers in your company? (Ex. HCI, Industrial design)

UI concept clearly. However, most companies surveyed here had a separate department for UX, and UX was considered as a more significant concept than the current usability concept.

Second, the majority of the companies did not have a UX guideline. Therefore the companies usually did not have any standardized procedure to evaluate UX after the development of products or services. They had a tendency to depend on current usability evaluation methods such as usability testing. However, all participants said they were willing to use new UX evaluation methods if the evaluation methods were reliable.

Lastly, the power of UX designers was increasing in the companies surveyed here, but most of them did not have a UX designer who took full charge of UX-related works. In addition, the UX designers at these companies generally had diverse educational backgrounds.

4 CONCLUSION

In this study, an in-depth interview was conducted with fourteen experts at cell phone manufacturing, mobile telecommunication, and internet service companies. The purpose of the interview was to conceptualize UX from an expert perspective by assessing the opinions of those who undertook tasks related to user interfaces or UX at their respective worksites. Additionally, information pertaining to UX applications in the Korean mobile company was gathered. In conclusion, experience related only to the use of products or services was defined as UX, and usability, affect, user value, and previous experience were suggested as UX elements.

As follow-up research, an ethnographic study can be conducted to collect end-users perspectives on UX. This type of work can supplement and verify the elements that affect UX significantly. In addition, UX evaluation methods can be defined by analyzing relationships among UX elements and by investigating measurement methods for UX elements.

ACKNOWLEDGEMENT

This work was supported by Mid-career Researcher Program through NRF grant funded by the MEST (No. 2010-0000364).

REFERENCES

Alben, L. 1996. Quality of experience: defining the criteria for effective interaction design. *Interactions* 3(3): 11–15.

Arhippainen, L. & Tähti, M. 2003. Empirical evaluation of user experience in two adaptive mobile application prototypes. *MUM 2003* 27–34.

Desmet, P. & Hekkert, P. 2007. Framework for product experience. *International Journal of Design* 1(1): 57–66.

Hassenzahl, M. 2001. The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction* 13: 479–497.

Hassenzahl, M. & Tractinsky, N. 2006. User experience - a research agenda. *Behaviour and Information Technology* 25(2): 91–97.

Karapanos, E., Zimmerman, J., Forlizzi, J. & Martens, J-B. 2009. User experience over time: An initial framework. *CHI2009* 729–738.

Law, E.L.-C., Roto, V., Vermeeren, A.P.O.S., Kort, J. & Hassenzahl, M. 2008. Towards a shared definition of user experience. *CHI2008* 2395–2398.

Law, E.L.-C., Roto, V., Hassenzahl, M., Vermeeren, A.P.O.S & Kort, J. 2009. Understanding, scoping and defining user experience: A survey approach. *CHI2009* 719–728.

Mäkelä, A. & Fulton Suri, J. 2001. Supporting users' creativity: design to induce pleasurable experiences. In M.G. Helander, H.M. Khalid, T. Ming Po (Eds.), *Proceedings of International Conference on Affective Human Factors Design, Singapore 2001.* ASEAN Academic Press, 387–394.

Roto, V. 2006. *Web browsing on mobile phones - characteristics of user experience*. Doctoral dissertation at Helsinki University of Technology, Finland.

Shedroff, N. 2006. An evolving glossary of experience design, online glossary from http://www.nathan.com/ed/glossary

UPA 2006. Usability body of knowledge. Usability Professionals' Association online glossary: *http://www.usabilitybok.org/glossary*

Väänänen-Vainio-Mattila, K., Roto, V. & Hassenzahl, M. 2008. Towards practical user experience evaluation methods. *VUUM*.

Yamazaki, K. & Furuta, K. 2007. Design tools for user experience design. *HCII2007*, 298–307.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Identifying and comparing user experience quantification models

Jaehyun Park, Sung H. Han, Hyun K. Kim, Seunghwan Oh, Heekyung Moon, Youngseok Cho & Chun Jaemin
Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea

ABSTRACT: As interest in the concept of User Experience (UX) grows, studies on improving UX have been widely conducted with various academic and practical approaches. Therefore, quantification methods for UX are regarded as important. This paper uses a case study to identify and compare UX quantification models. Tested models included linear, polynomial, conjunctive, and disjunctive. The case study was carried out with 26 participants using a commercial tablet computer in a laboratory environment. Each participant was asked to evaluate a total of 22 measures using a modified magnitude estimation method. The measures, organized hierarchically, included usability, affect, and user value as the main elements of UX. The results of this study show that polynomial models have better performance than other models in terms of adjusted R^2 values. The results are expected to help product or service designers develop an overall UX index.

Keywords: user experience, quantification methods, mathematical models, decision strategies

1 INTRODUCTION

In the last decade, the concept of user experience (UX) has received special attention from HCI researchers. UX reflects all aspects of the interaction between a user and a product (Alben 1996). Thus, the UX concept includes various aspects that have been overlooked from the perspectives of usability and affect engineering (Hassenzahl & Tractinsky 2006). The main factors that influence UX have been investigated in numerous studies. For example, context of use (Arhippainen & Tähti 2003), temporality (Karapanos et al. 2009), and the user's subjective value (Kim et al. 2010) can be important elements of UX.

As many studies defining UX have been conducted, improving UX has been regarded as an important challenge. Finding ways to improve UX begins with measuring and evaluating the present state of UX. However, relatively few studies on quantifying UX have been carried out, while many studies have measured brand equity, service quality, and perceived ease of use as a numerical value. The quantification methods for such constructs as brand equity might be applied to UX.

This paper identified quantification models for UX and compared them by conducting a case study. First, quantification models including linear, polynomial, conjunctive and disjunctive models were elicited by considering user decision strategies and mathematical models. Then, a hierarchical structure of UX measures, their evaluation method, and controlled factors were determined for the laboratory experiment. Participants were asked to assume that they just purchased the experimental device so that we could investigate an early stage of the customer lifecycle.

2 QUANTIFICATION METHODS

2.1 *Literature review*

Articles concerning quantification methods were collected from various academic domains such as human factors, software engineering, quality engineering, psychology, business and marketing science, and information science. The keywords used in this interdisciplinary survey included quantification and evaluation method, information integration theory, decision making process, customer judgment, and customer choice prediction. A total of 62 articles were collected. Two criteria for classifying the quantification methods were developed: type of mathematical model and user decision strategy.

First, the mathematical models can be classified in two ways, according to whether the objective functions and constraints are represented by a linear relationship (linear model vs. nonlinear model) or whether variable states are described by probability distributions (deterministic model vs. probabilistic model). Second, user decision strategies for evaluating alternatives, including compensatory and noncompensatory decision rules, should be considered while quantifying UX. Compensatory rules imply that a product's evaluation on any dimension may be offset by evaluations on other dimensions. On the contrary, noncompensatory rules mean that a decision may be determined by a product's score on a single dimension, irrespective of scores on other dimensions (Elrod et al. 2004).

2.2 *Quantification models*

Based on the user's decision strategies and the mathematical criteria stated above, four quantification models were identified: linear, polynomial, conjunctive, and disjunctive. Linear and polynomial models are classified as compensatory strategies, while the others are classified as noncompensatory strategies. The linear model is equivalent to a simple weighted sum. The polynomial model consists of a quadratic formula. Conjunctive and disjunctive models which were approximated by a log transformation (Einhorn 1970) were used. In the conjunctive model, the user adopts a certain product if every attribute of the product satisfies his or her minimum requirements. On the other hand, in the disjunctive model, the user may choose a product if just one attribute of the product is remarkable. Many formulae reflect conjunctive and disjunctive models. Among them, the log transformation has been widely used; it is easy to use because parameters about customer requirement levels are not needed. The quantification models used in this study are as follows:

Linear model:

$$Y=\sum_{i=1}^{n}\omega_i X_i \tag{1}$$

Second-order polynomial model:

$$Y=\sum_{i=1}^{n}a_i X_i+\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}X_iX_j+\sum_{i=1}^{n}c_iX_i^2 \tag{2}$$

Conjunctive model:

$$Y=\Pi_{i=1}^{n}X_i^{\alpha_i} \tag{3}$$

Disjunctive model:

$$Y=\Pi_{i=1}^{n}\left(\frac{1}{c_o-X_i}\right)^{\beta_i} \tag{4}$$

where Y = dependent variable (e.g., overall UX, usability, affect, and user value); X_i = ith independent variable of the model; n = number of independent variables included in the model; and c_0 = a constant variable above the largest X_i to ensure Y will not be infinite.

3 CASE STUDY

3.1 *Concept of UX quantification*

As many literatures mention, UX itself is an ambiguous and abstract construct. Therefore careful considerations should be given to quantifying UX. First, the context of use, user group, and device should be defined. For example, UX changes over time (Karapanos et al. 2009). A user who was attracted by the external appearance of a certain smart phone just after purchase can become disappointed because its interface is complicated at the peak of learning. In this study, the point of time was assumed as just after purchase.

Second, measures for quantifying UX should be defined comprehensively. UX can be decomposed into a hierarchical structure (Fig. 1). In this study, a total of 22 measures, including overall UX, were developed by modifying the structure of a previous study (Kim et al. 2009). Among them, usability, affect, and user value are regarded as major elements influencing UX. Each element has sub-elements, which are attributes of the elements. For example, usability consists of seven sub-elements: simplicity, directness, efficiency, informativeness, flexibility, learnability, and user support. Similarly, affect and user value have six (color, delicacy, texture, luxuriousness, and attractiveness) and five sub-elements (self-satisfaction, pleasure, customer need, sociability, and attachment).

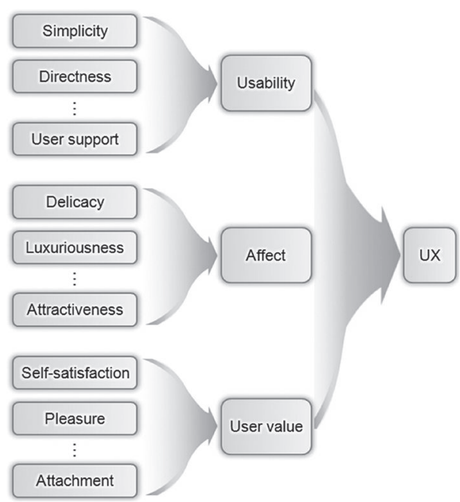


Figure 1. Hierarchical structure of UX measures.

3.2 *Experimental methods*

3.2.1 *Apparatus*

As a commercial mobile tablet computer, iPad without 3G was used for this experiment. As an experimental prototype, the iPad can minimize the effect of the previous participant experience, because the iPad had not been released in South Korea until the end of the experimental period.

3.2.2 *Participants*

A total of 26 qualified participants, consisting of an equal number of males and females, took part in the experiment. Their ages ranged from 18 to 28 years old (mean = 22.4, SD = 2.6). Participants who had not had previous experience with the iPad, iPhone or iPod were recruited to set the experimental situation as if they had just made a purchase.

3.2.3 *Tasks and evaluations*

Tasks were developed with the default applications of the iPad because participants were assumed to have just purchased the device. There are thirteen built-in applications: calendar, Safari, mail, iPod, settings, maps, videos, YouTube, photos, contacts, notes, iTunes, and App Store make up thirteen tasks. Each task has two to four sub-tasks. For example, the mail task includes three sub-tasks: e-mail registration, reading an e-mail, and sending an e-mail.

A modified magnitude estimation method was used for the evaluation. Each participant was asked to evaluate each of 22 attributes with a zero-to-hundred rating scale. Zero and hundred indicated two extremes, and fifty means unbiased point. Participants could perform these tasks without any strain because all measures were reported subjectively.

3.2.4 *Procedure*

Each participant was instructed about the tasks and the evaluating measures at the beginning of the main experiment. Then they completed their tasks. The presentation order of tasks was determined by the Latin square balancing technique. After all tasks were done, the participant was allowed to freely operate the device for ten minutes as he or she wanted. A total of five evaluations were repeatedly conducted to analyze a time effect. The first evaluation was conducted after the first task, with the other evaluations after the fifth, ninth, and thirteenth tasks and after the free session.

3.3 *Results*

3.3.1 *User experience over time*

An analysis of variance (ANOVA) was used to analyze whether each measure varied over time. The effect of time was found to be significant at $\alpha = 0.05$ for 16 of the 22 measures. A total of six measures turned out to be insignificant. Five of them were affect dimensions (affect, color, texture, luxuriousness, and attractiveness), and one of them was a usability dimension (learnability). Post-hoc analyses, the Student Newman-Keuls (SNK) test were conducted on significant main effects at $\alpha = 0.05$. The results show that early ratings tended to be lower than later ratings (Fig. 2).

3.3.2 *Model building*

A total of 80 models were developed with combinations of four types of model, four dependent measures, and five time dimensions. The four types of model included linear, polynomial, conjunctive, and disjunctive models, as mentioned above. The four dependent measures consisted of overall UX, usability, affect, and user value. Their modeling formations depended on the hierarchical structure of the UX concept. For example, if overall UX is selected as a dependent variable, then usability, affect, and user value are considered as independent variables. Similarly, when the usability measure is a dependent variable, simplicity, directness, efficiency, informativeness, flexibility, learnability and user support are regarded as alternative independent variables.

Every analysis was conducted with all possible regression methods. However, as an exception, when usability, affect, and user value were dependent variables and their models were assumed as the polynomial form, stepwise regression was applied because of the large number of candidates for the independent variables. The main criterion used to select the best model was the adjusted R^2 value. In addition, if the maximum Variance Inflation Factor (VIF) value was higher than ten, the model was not selected to reduce the multi-collinearity effect (Kutner 2005).

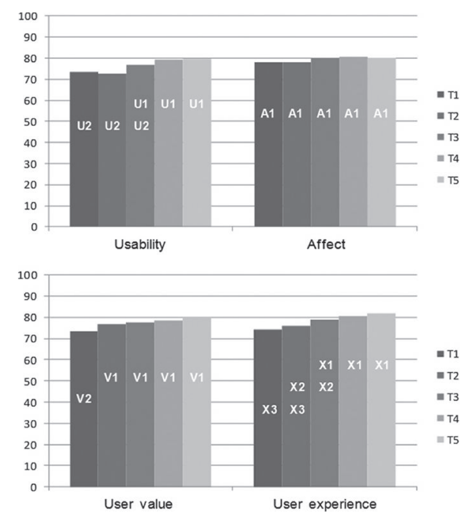


Figure 2. Scores of measures over time (T1–T5) (The same code have no statistical significance at $\alpha = 0.05$).

Table 1. Selected model of UX for each time dimension.

Time	Model	Adj. R^2	Regression equations (UX =)*
1st	P	0.82	$0.52 \times V + 0.003 \times U^2$
2nd	P	0.86	$0.426 \times U - 0.005 \times A^2 + 0.01 \times A \times V$
3rd	C	0.92	$A^{0.29} \times V^{0.616}$
4th	P	0.83	$0.479 \times V + 0.003 \times U \times A$
5th	P	0.84	$0.002 \times V^2 + 0.004 \times U \times A$

* 'P' and 'C' denote for polynomial and conjunctive.
* 'U', 'A', and 'V' denote for usability, affect, and user value.

The results reveal that the polynomial models tend to have better performance than other models in terms of adjusted R^2. For overall UX, the polynomial models turned out to be the best models except for the third evaluation (Table 1). On the other hand, disjunctive models tended to have lower performance than other models.

3.4 *Discussion*

As the results of the ANOVA, the effect of time was not significant for five affect dimensions. Affect toward a product or service may not vary over time. Affect involves the degree to which a product's appearance or image appeals to users. Thus, unless users' standards of beauty change, subjective satisfaction from the outward appearance of a product would not change. On the other hand, a user's performance and perceived usability are likely to vary, because the user is learning about the product over time. User value is also apt to vary over time for psychological or contextual reasons.

Adjusted R^2 values of the models were more than 0.7, which implies that the models have high predictability. Even if the performance of polynomial models turned out to be better than other models in terms of adjusted R^2, conjunctive models also had high adjusted R^2 values. Thus, user strategy can be explained by both the compensatory and noncompensatory rules. Compensatory models such as the polynomial model have the merit of being parsimonious and powerful, while noncompensatory models such as the conjunctive model provide more insight into user judgment (Sethi & King 1999).

4 CONCLUSION

The starting assumption of this study was that the degree of overall UX can be integrated into a single value. Skeptics may argue about the representativeness of a single numerical value for everyday experience. However, most HCI specialists from academia and industry emphasize the necessity of quantifying UX (Law et al. 2009).

This study identified quantification models for UX through a literature review. The four identified quantification models consisted of two compensatory models (linear and polynomial) and two noncompensatory models (conjunctive and disjunctive). These models were compared by a case study using a commercial tablet computer. Participants evaluated the products using a modified magnitude estimation method on 22 hierarchical measures. A total of 80 models were developed with high values of adjusted R^2. These results provide evidence that an overall UX index can be explained by the combination of usability, affect, and user value.

ACKNOWLEDGEMENT

This work was supported by Mid-career Researcher Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (MEST) (No. 2010-0000364)

REFERENCES

Alben, L. 1996. Quality of Experience, *Interactions* May/June.

Arhippainen, L. & Tähti, M. 2003. Empirical evaluation of user experience in two adaptive mobile application prototypes. In: *Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia (MUM 2003)*: 27–34.

Einhorn, H. 1970. The use of nonlinear, noncompensatory models in decision-making. *Psychological Bulletin* 77.

Elrod, T., Johnson, R.D. & White, J. 2004. A new integrated model of noncompensatory and compensatory decision strategies. *Organizational Behavior and Human Decision Processes* 95(1): 1–19.

Hassenzahl, M. & Tractinsky, N. 2006. User experience - a research agenda. *Behaviour and Information Technology* 25(2): 91–97.

Karapanos, E., Zimmerman, J., Forlizzi, J. & Martens, J-B. 2009. User experience over time: an initial framework. In: *CHI2009*.

Kim, H.K., Han, S.H., Park, J. & Park, Y.S. 2010. The development of components of user value through an ethnographic study. In: *HCI Korea 2010*.

Kim, H.K., Han, S.H., Park, J., Park, W., Park, Y.S., Cho, Y., Chun, J. & Oh, S. 2009. The definition of user experience through a literature survey. In: *Proceedings of the 2009 Fall Conference of the Korean Institute of Industrial Engineers*.

Kutner, M.H., Christopher, J.N., Neter, J. & William, L. 2005. *Applied Linear Statistical Models*. fifth ed. McGraw-Hill, New York.

Law, E.L.-C., Roto, V., Hassenzahl, M., Vermeeren, A.P.O.S. & Kort, J. 2009. Understanding, scoping and defining user experience: a survey approach. In: *CHI2000*.

Sethi, V. & King, R.C. 1999. Nonlinear and noncompensatory models in user information satisfaction measurement. *Information Systems Research* 10(1): 87–96.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

From design to experience: Towards a process model of user experience

Uwe Drewitz & Stefan Brandenburg
Berlin Institute of Technology, Berlin, Germany

ABSTRACT: Nowadays, interactive products, such as computers, mobile devices and automatic orders are widely used to manage tasks in both, working environmental and everyday life contexts. Up to now, there is no model available, which draws a line between objective design features and subjective user experience. The presented approach *towards a process model of user experience* (MUX) aims to link the two worlds. Therefore it utilizes existing concepts of ecological and cognitive psychology as well as user experience research. Experimental data is presented, which supports the theoretical assumptions. Based on these empirical results, first very promising steps have been made to fill the long lasting gap between design and user experience.

Keywords: user experience, design, immediate interactive behavior, affordance

1 INTRODUCTION

Over the last years the research focus of human-computer interaction has shifted from a technology-centered to a user-centered perspective. Modern HCI research focuses no more on limitations and capabilities that can be described objectively. Instead researchers, as well as HCI designers and engineers, try to capture user experience (UX). Thereby experience is exclusively subjective and hard to explain. Nevertheless, researchers and large companies keep focusing on this concept mainly for two reasons. First, there is a scientific interest in explaining the emergence of UX due to the gap between the design of a technical environment (i.e. interface) and the users experience while working with that technical artifact. Second, companies are interested in the long-term engagement of their customers, which presumably can be achieved by positive UX. Therefore, it is crucial to explain the process, which starts at the design of technical artifacts and ends up at user experience while working with the artifact itself.

We suppose, that research on UX does not have to start from scratch. Dealing with human experience it is possible to endeavor well-known concepts of different domains such as cognitive science to fill the gap between design and experience. The proposed process model of user experience aims to deploy such concepts. This framework uses approaches from the psychology of perception (Gibson, 1979; Norman, 1988) to define task related *affordances*, *constraints* and *mapping* in a *well-structured environment*. Building upon this environmental structure the interdisciplinary approach of *Immediate Interactive Behavior* (IIB, Neth et al., 2007) comes into play. Following IIB, basic inductive learning processes occur, which are captured by the *Ease of Induction concept* (EOI, Brandenburg et al., 2009). Given the appearance of all these concepts, positive user experience emerges, that is captured by the *Ease*- and *Joy of Use* (EOU & JOU) concepts of Davis (1989).

However, the question is: How do these single concepts connect to a theoretical chain attaching the design of technical devices with user experience? The proposed process model of user experience (MUX) accomplishes that task. The basic mechanism of this process model of user experience can be described like the following: when an environment is pre-structured in terms of the task, constraints help to enforce the correct action. In this case *Immediate Interactive Behavior* (IIB, Neth et al., 2007) is likely to occur. IIB refers to the trait of cognitive agents, to adapt and structure their environments in service of their goals. The routinely and dynamically use of their embodied and environmentally embedded nature augments the agents cognitive processes.

In the process model of user experience IIB plays the key role in human machine interaction mainly for two reasons. First, the occurrence of

IIB represents the expansion of cognition. This will impact both, users emotion and motivation. In result, users experience feelings of competence and ability while making progress using the artifact. Second, if present IIB facilitates further learning processes, implicit as well as explicit. Especially the easy and fast acquisition of operating knowledge, due to inductive learning processes, benefits. We call this process *Ease of Induction* (Brandenburg et al., 2009). Both, IIB as well as EOI lead to ease of use and positive user experience, since an unostentatious interaction style and therefore the unobtrusive acquisition of operating rules helps to reach interaction goals directly and efficiently.

2 THE PROCESS MODEL

A theoretical exploration of user experience does not need start from scratch. For example, Thüring and Mahlke (2007) investigated determinants of user experience in human-technology interaction and developed the CUE-Model (*C*omponents of *U*ser *E*xperience), which is displayed in Fig. 1. The CUE-Model states that user experience is a compound of emotions and perceptions of instrumental (e.g. learnability) as well as non-instrumental (e.g. visual aesthetics) qualities. System properties, user characteristics and the task context are factors of the interaction characteristics, which guide the perception of instrumental and non-instrumental qualities.

In accordance with the CUE-Model, the presented model of user experience (MUX) assigns a high importance to the task, which should be accomplished. First of all, interactive devices and its software are designed not for any but some concrete tasks. We think of users as knowing which task they want to or which tasks can be accomplished with the device they are interacting with. Second, the task becomes very important for setting constraints and therefore promoting certain affordances when designing a technical artifact, which results into a high ease of use and positive user experience based on IIB. Another commonality between the theoretical models is their explanatory scope. Both aim to account for differences in usage behavior/ease of use (CUE/MUX) and user experience/ joy of use (CUE/MUX).

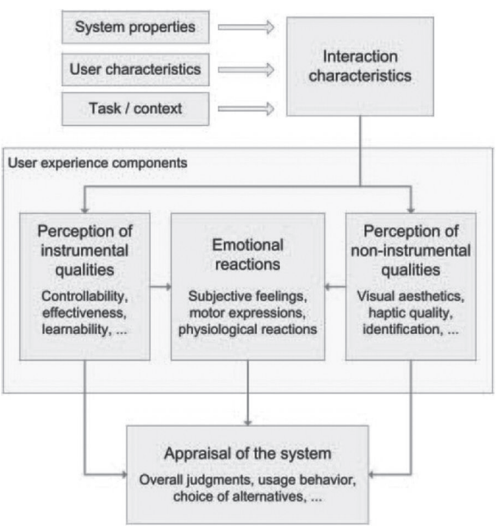


Figure 1. Components of user experience: The CUE-Model (Thüring & Mahlke, 2007).

In contrast to the CUE-Model our MUX-model only focuses on system properties and the perception of instrumental qualities. We do not state that the user experience and its qualities can be reduced to those factors. Instead we assume that some central aspects of user experience, especially the *Ease of Use* and the *Joy of Use* can be described depending on processes that evolve due to system properties alone. In fact we believe that there are features of the technical artifact which can be captured objectively and which trigger cognitive processes yielding user experience of high quality. How *Ease of Use* and *Joy of Use* emerge in the MUX-model is displayed in Fig. 2. The model distinguishes two locations of implementation, the environment (the user interface provided by the technical artifact) and the person. As mentioned above, there are a number of theoretical concepts used by the MUX-model. For each of these concepts the location of implementation (environment or person), its content (i.e. routines, skills or experience) and a measure (i.e. reaction times) are specified.

Within the bottom line of Fig. 2, the structured environment, e.g. an interface, is located. A structured environment contains action possibilities (affordances), constraints that determine which

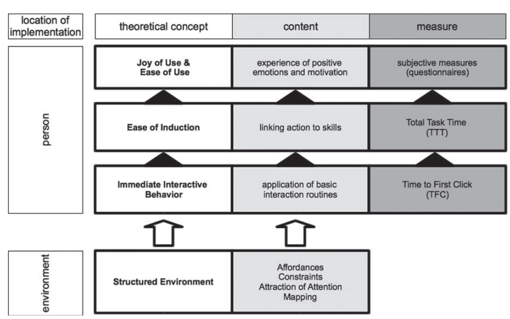


Figure 2. Process model (MUX) of the emergence of Ease- and Joy of Use.

interactive behavior is possible and cues that guide the attention of the user. Additionally the expectations of users should be satisfied to some degree, which can be ensured if the users common knowledge maps (cf. Norman, 1988) to the environment that is created by the artifact. Whereas Norman (1988) distinguishes *affordances* and *"perceived affordances"* and *conventions*, we think of affordances as pointed out by James Greeno (1994): *"In any interaction involving an agent with some other system, conditions that enable that interaction include some properties of the agent along with some properties of the other system. [...]The Term affordance refers to whatever is about the environment that contributes to the kind of interaction that occurs." (pp. 338)*.

However, having a lot of competing affordances it is hard for the user to choose the right action. Therefore the design of an artifact has to be constrained. Constraints are limitations within the interface or the operational concept of the artifact. Limiting the number of actions that can be done at one time leads the user to choose the right action. Moreover it is important that *affordances* can be learned (cf. Gibson, 2000). On the part of the user, we have the capabilities of embodied cognition or *abilites* (Greeno, 1994), which is the complement to affordances. The abilities of the user together with the affordances of the environment (artifact) produce *immediate interactive behavior*. An Example is the perception of an icon on a touch screen and pressing that icon. That can be done very easily and does not require higher-level cognition as for example deductive reasoning.

However, even if constraints, cues and mappings support the perception of affordances, they should be perceivable for their own to facilitate the occurrence of IIB. The central concept of the process model is that of immediate interactive behavior. As an interdisciplinary research approach *Immediate Interactive Behavior* is the successor and extension of Allan Newells (1990) *Immediate Behavior*. IIB refers to behavior that is routinely and dynamically applied and which connects perception and action without so called higher cognitive processes. The occurrence of IIB takes place on a 300 ms to 1 sec. (3 sec. maximum) time scale. There is no deliberate behavior that can occur on a time scale smaller than that of IIB. Actions that are shown at this time scale are deliberate but not higher cognitive, in the sense that people do not have to think or reason as much which action would be appropriate. It is simply acting appropriate in response to events in the environment. People show IIB when they want to structure or adapt the environment to serve their goals. As mentioned before, a necessary condition for the occurrence of IIB is the existence of perceivable offers for manipulation or transformation on the part of the environment. If the skill acquisition is facilitated due to the appearance of IIB, operating knowledge (especially procedural knowledge) can be acquired easily by induction (Brandenburg et al., 2009). The reasoning behind is, that when then the right actions appear in sequence rapidly the probability increases that these actions will be linked up with each other and built into new (procedural) knowledge. In consequence people will experience *Ease of Use* (EOU). Examples for such mechanisms of skill acquisition are provided by recent theories of cognition, for instance the production compilation mechanism in ACT-R (Anderson et al., 2004).

Closely linked with the EOU is the *Joy of Use.* When people are able to show IIB that, due to the well-structured environment, will produce the right action sequences to accomplish a goal, EOI as well as positive emotional and motivational reactions occur. First people become aware of the progress they make in processing a task. Not getting stuck in an impasse but approaching the target when using a new or unknown artifact is accompanied by the feeling of being successful. Second, making progress will affect the motivation of users. Experiencing ones self-efficacy will improve the motivation to proceed with the interaction or to start another interaction at a later date (Minge & Thüring, 2009). That is why interactions, that affect the experience of users as described above, should yield a positive appraisal of an artifact.

Integrating the specified concepts into the process model, we set up a theoretical framework that can be used as basis for research as well as engineering and design. Starting at the environment, the environmental factors are the independent variables of the model, which can be manipulated to influence the processes on part of the user. Researchers, engineers and designers can and should manipulate these factors along the dimensions that are relevant. For example, a user interface can provide affordances (action possibilities) or not. An interface can be high or low constrained, which means there are a lot or fewer possibilities to act (cf. Gray et al., (2006) for a distinction of hard and soft constraints of interactive behavior). Cues that guide the attention of users can be present and they can have different degrees of attracting the attention of the user. An artifact can be designed with respect to the mental models, which users have for instance due to prior knowledge. It is obvious that there are a lot of manipulations and much more combinations of manipulations possible. But not all of them should yield Ease- and Joy of Use. We claim that artifacts should provide *learned affordances*, should have *constraints,* which limit the number of actions possible at a time, should make use of cues to guide the attention of the user and should give respect to user expectations.

The allocation of measures to every concept on the persons' side will give support to proof the assumptions and the conception of the model. On the other hand, having measures that are linked directly to concepts should also give support to evaluate and revise designs for interaction artifacts based on theoretical consideration. And despite of the subjective nature of user experience, our model proposes testable assumptions concerning connections between cognitive processes that depend on the structure of the environment. The environment, which is the technical artifact itself, can be described objectively. Nevertheless such descriptions always have to take into account the (assumed) users of an artifact. The consideration of both users and artifact will bridge the gab between the subjective and the objective perspective, as claimed by Gibson (1979). In a first empirical study we tested our model assumptions with respect to constraints. We designed a mobile device and manipulated the number or interaction elements provided by the interface to test the hypothesis that a more constrained interface will result in higher IIB, EOI as well as increased ratings for EOU and the JOU compared to the less constrained interface.

3 EMPIRICAL STUDY

3.1 *Subjects*

A total of N = 25 participants (age: M = 28.4, SD = 5.3, 5 female/20 male) participated in the present experiment. Subjects participated voluntarily and did not receive any incentives.

3.2 *Material*

For the purpose of theory testing four instantiations of an interactive HTML-mockup were realized (see Fig. 3).

Within Figure 3 versions 3a and 3b display two interface types with indirect control elements. They were thought to be indirect because users had to move the rectangular cursor with the left and right arrow-buttons. After placing the cursor at the desired icon, participants had to press the enter-button (white squared button). In contrary, versions 3c and 3d visualize versions with direct control elements. Here, subjects had a direct mapping of buttons to icons (3c) or the opportunity to touch the icon directly (3d). Within each instantiation of the mockups the icons on the screen as well as the structure of the menu was identical. The experimental variation only changed the appearance of the type of interaction. As experimental variation one factor was manipulated, the type of

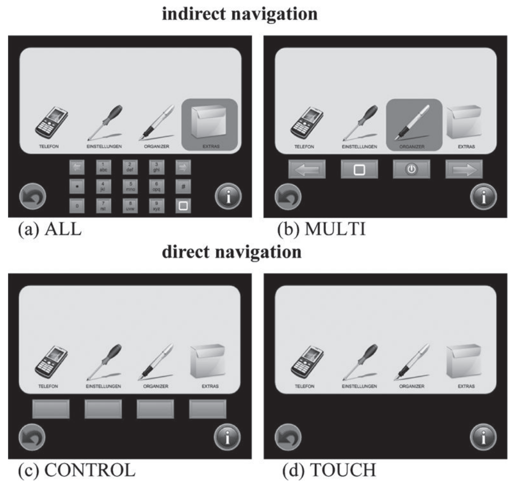


Figure 3. Stimulus material.

navigation (within subjects) was either direct (3c & 3d) or indirect (3a & 3b).

3.3 *Procedure*

Every subject worked itself through three prototypical tasks with each of the four interfaces at three points in time. The order of tasks and interfaces was randomized. With that experimental design a total number of N = 36 measures were captured from each subject. For convenience three prototypical tasks, which were repeated three times were summarized to three blocks. After completion of each of the three blocks with one interface, participants received three questionnaires. These questionnaires assessed perceived usefulness and utility (Davis, 1989), stress and strain (Hart & Staveland, 1988) as well as an emotional product rating (Mano, 1991). As objective measures Total Task Time (TTT) and the time until the first click occurred (TFC) were measured for each block.

3.4 *Results*

For the test of the model assumptions ANOVAs with type of navigation (4 systems), and blocks (3 for each interface) as repeated measures were calculated. Since the model building process is in focus in this paper, only results with tight connection to the model are reported in the upcoming paragraphs.

Figure 4 shows the results with respect to the occurrence of IIB. IIB was defined as time to the first click (TFC) within a block of tasks.

As shown in Figure 4, all of the interface variations facilitated IIB since TFC was below 3 sec.

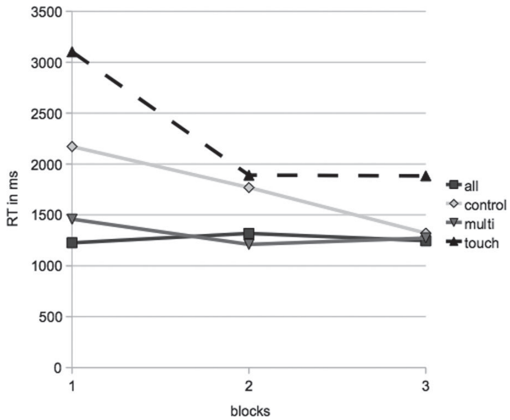


Figure 4. Time to First Click (TFC) for the three blocks and four interfaces.

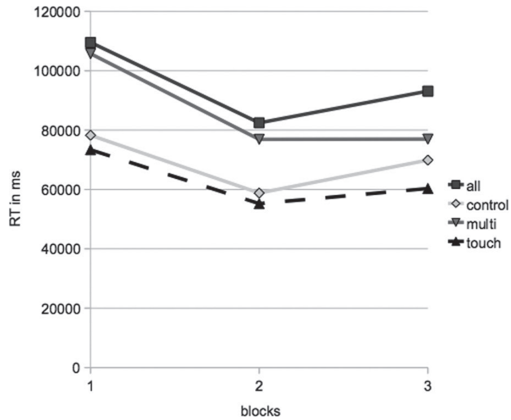


Figure 5. Total Task Time (TTT) for the three blocks and the four interfaces.

For the ANOVA a significant main effect for blocks was obtained, $F(2,44) = 6.63$, $p = 0.003$, eta = 0.23. Further, the four systems differed significantly, $F(3,66) = 10.41$, $p < 0.001$, eta = 0.32. No interaction was observed, $F(6,132) = 1.35$, $p = 0.24$, eta = 0.05. A post hoc Scheffé test revealed significant differences at time at block 1. There, the pure touch screen (touch) differed significantly from the simple indirect (multi) navigation mockup ($p = 0.001$) and the complex (all) indirect navigation mockup ($p < 0.001$). The comparison between the pure touch screen (touch) and the control condition (control) almost reached marginal significance ($p = 0.14$), which becomes important for the interpretation of the effects. No significant differences between the systems were obtained for the blocks 2 and 3.

As practical measure for EOI the Total Task Time (TTT) was proposed. Figure 5 shows the variations in TTT with respect to the experimental variation. The ANOVA revealed two main effects. First, the three blocks differed in TTT, $F(2,40) = 14.16$, $p < 0.001$, eta, 0.42. Second, subjects needed significantly more time using some of the four systems, $F(3,60) = 2.88$, $p = 0.04$, eta = 0.13. Again, no interaction was observed, $F(6,120) = 0.56$, $p = 0.75$, eta = 0.03. A post hoc Scheffé test revealed a significant difference between the pure touch screen (touch) and the complex indirect navigation mockup ($p = 0.05$).

Regarding the subjective measures, the direct controllable interfaces appeared to be significantly better in terms of usability, utility, elation, palatableness and serenity compared to the indirect controlled mock-ups (Fig. 6). The smallest significant difference for these effects was for serenity, $F(3,60) = 4.22$, $p = 0.01$, eta = 0.17. With respect to all other subjective measures, the effects were larger.

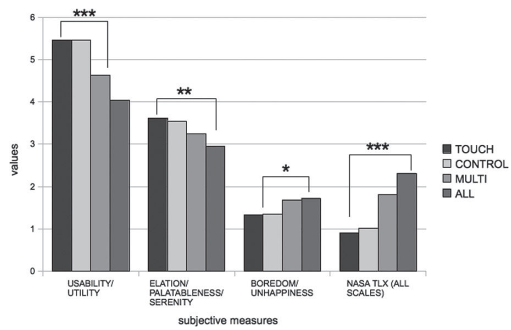


Figure 6. Subjective measures for the four interfaces; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

On the other hand, the direct versions (touch, control) of the interface appeared to impose less stress and strain (NASA-TLX scales; $F(3,60) = 5.11$, $p = 0.003$, eta = 0.2) and less feelings of boredom ($F(3,60) = 3.30$, $p = 0.03$, eta = 0.14) and unhappiness ($F(3,60) = 3.46$, $p = 0.02$, eta = 0.15) than the indirect versions (multi, all) of the mockup.

4 DISCUSSION

The first experimental results support the process model of user experience (MUX). On the very basic level of interaction IIB occurred with all four interfaces. These occurrences of IIB lead to EOI, which was demonstrated by the main effects for the blocks in the TTT. Finally, on the user experience level, mockups with a very direct interaction (touch, control) appeared to be better in usability, utility, elation, palatableness and serenity as well as less demanding.

Despite this general support, certain facets of the results still need attention. First, the significant difference in IIB at the very first interaction in block one. At first sight, this result seems to contradict our hypothesis, that a pure touch screen should support IIB more than other types of navigation. But examining the properties of the mockups in detail, one can see, that touch and all differed with respect to more than the direct interaction. They also differed in the presence or absence of buttons. Taking the marginal significant difference between the control and the touch condition into account, the buttons in the control interface seem to have a higher affordance to be pressed compared to the pure icons on the screen in the touch condition. This interpretation is undermined by another analysis (the Ryan procedure; Toothaker, 1993), which grouped three (all, multi, control) of the four interfaces and marked them significantly different from the fourth interface (touch). Second, the pure number of buttons (versions all and multi) does not seem to generate differences neither on the behavioral, nor on the experience level. One possible explanation might lie in the ability of users to cognitively block out the presence of additional buttons, which are not task relevant.

Further research should target this issue by constructing tasks that utilize all available buttons in the interface. Additionally, more concepts of the design level of the MUX-model need to be tested. Present research only considered constraints. Subsequent work needs to focus on the role of attention or mapping and their single and collaborate contribution to IIB, EOI and user experience.

REFERENCES

Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C. & Qin, Y. 2004. An integrated theory of the mind. *Psychological Review* 111(4): 036–1060.

Brandenburg, S., Drewitz, U., Thüring, M., Minge, M. & Brune, T. 2009. Touch me if you can - immediate interactive behavior als bedienkonzept für multimediageräte. In *Der Mensch im Mittelpunkt technischer Systeme. 8.Berliner Werkstatt Mensch-Maschine-Systeme*. 22(29): 105–106. Düsseldorf: VDI Verlag GmbH.

Davis, E. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly* 13(3): 319–340.

Gibson, E.J. 2000. Where is the information for affordances? *Ecological Psychology* 12(1): 53–56.

Gibson, J. 1979. *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Greeno, J.G. 1994. Gibson's affordances. *Psychological Review* 101(2): 336–342.

Gray, W.D., Sims, C.R., Fu, W.-T. & Schoelles, M.J. 2006. The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological Review* 113(3): 461–482.

Hart, S.G. & Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (Ed). *Human mental workload. Advances in psychology* 52: 139–183. Oxford, England: North-Holland.

Mano, H. 1991. The structure and intensity of emotional experiences: method and context convergence. *Multivariate Behavioral Research* 26(3): 389–411.

Minge, M. & Thüring, M. 2009. Erleben von bnutzbarkeit und athetik in der Mensch-Technik-Interaktion. In A.B. Eder, K. Rothermund, S.R. Schweinberger, M.C. Steffens & H. Wiese, (Hrsg.), *Beiträge zur 51. Tagung experimentell arbeitender Psychologen*, (S.38). Lengerich: Pabst Science Publishers.

Neth, H. Carlson, R.A., Gray, W.D. Kirlik, A. Kirsh, D. & Payne, S.J. 2007. Immediate interactive behavior: How embodied and embedded cognition uses and changes the world to achieve its goals. In D.S. McNamara & J.G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society*: 33–34.

Newell, A. 1990. *Unified teories of cgnition*. Harvard University Press. Cambridge, Massachusetts.

Norman, D.A. 1988. *The psychology of everyday things.* New York: *Basic Books*.

Thüring, M. & Mahlke, S. 2007. Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology* 42(4): 253–264.

Toothaker, L.E. 1993. *Multiple comparison procedures*. Sage Publications.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Investigating the basic components of user value through an ethnographic study

Youngseok Cho, Sung H. Han, Jaehyun Park & Hyun K. Kim
Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea

ABSTRACT: The aim of this study was to develop and verify the basic components of user value, an element of UX, by using an ethnographic method. An initial list of the components was developed by literature review. Daily experiences of interacting with mobile phones were collected in the ethnographic experiments and the list of the components was revised based on the data. The results of this study can be used to understand the user value of a product or a service.

Keywords: user value, user experience, ethnographic method, mobile phones

1 INTRODUCTION

Defining and evaluating User Experience (UX) have become important research topics since UX is critical to the market competitiveness of products and services. UX broadly describes all aspects of the interactions between the user and a product (Alben 1996, Arhippainen & Tähti 2003, Forlizzi & Ford 2000, Hassenzahl & Tractinsky 2006, Kuniavsky 2007). Many researchers, however, have proposed different definitions of UX depending on their research interests and have provided impractical evaluation methods reflecting only a small portion of UX.

A research funded by the Korean government, has been performed to define comprehensive UX and its evaluation model for mobile phones and services, and to develop systematic methods and practical tools for evaluating UX. This study suggests that usability, affect and user value are three elements of UX.

Usability is defined as the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments (ISO DIS 9241-11). Affect is considered as an emotion, which is a consequence of interaction with a product or service. User value is a subjective value that the user attaches to a product. The value may be related to how the user thinks the product is meaningful and significant in his or her life.

Numerous previous studies have been conducted to investigate components or measurements methods for usability and affect (Han & Hong 2003, Han et al. 2000). In addition, many researchers have attempted to define a concept of user value. However, to date, few studies have investigated components of user value from the user's standpoint.

The aim of this study was to develop and verify the basic components of user value. An initial list of the components was developed based on the related studies and it was revised by analyzing the results of an ethnographic study.

2 METHODS

2.1 *Ethnographic method for capturing experiences*

Ethnographic approaches were used to capture and understand the daily experiences of users for designing a product or a service. The Experience Sampling Method (ESM) and the Day Reconstruction Method (DRM) are widely used to capture the experiences of users. In ESM, experimenters ask participants to write down their experiences on a record form at random times (Hektner et al. 2007, Larson & Csikszentmihalyi 1983). This requires much effort by the experimenters and participants, so Kahneman et al. (2004) applied the Day Reconstruction Method (DRM). In this approach, participants are just asked to record their experiences once before going to sleep.

In this study, a modified DRM was used to capture the experiences of users when they interacted with their mobile phones. Mobile phones were frequently used all day, so it is difficult to recall all the experiences at a time. Participants were asked to write down their experiences three times a day, and the records formed as a series of episodes. The name of the episode, place, time, affair and participant's feeling were recorded in each episode.

2.2 *Participants*

To define user groups, the characteristics of mobile phone users related to UX were analyzed with regard to social, cultural and contextual aspects. As a result, various characteristics were defined, such as age, gender, region, occupation, period of ownership, previous experiences with other phones, etc. Considering the purpose and constraints of the experiment, three characteristics (i.e. occupation, period of ownership and gender) were selected among them. There were two user types (i.e. male and female) for gender, and occupation was classified as student (e.g. high school or university student) or non-student (e.g. office worker). The period of ownership was also categorized into two groups considering that the average lifecycle of a Korean mobile phone was about 1.5 years: within a week after purchase (short term UX) and 6 to 12 months after purchase (long term UX). Thus, eight different user groups were defined by combining the characteristics.

Eight participants were recruited to represent each user group. Their average age was 24.5 years old with a standard deviation of 3.25 years. Three of them used full-touch phones, two used folder type phones, two used slide type phones, and one used an unspecified type of phone.

3 RESULTS

3.1 *Literature review*

To investigate the basic components of user value, an initial list was developed based on previous studies. Fourteen papers (i.e. journal, proceedings, and technical reports) in the fields of human factors, computer science, industrial design, and business management were surveyed. As a result, a total of twenty components were defined as an initial list (see Table 1).

3.2 *Collecting user experiences*

Seven days are sufficient to capture users' experiences (Kuniavsky 2007), so the ethnographic experiment was conducted over a week for each participant. The users recorded episodes of daily experiences using the modified DRM, and in total 216 episodes were collected.

Collected episodes were assigned to one of more of the components in the initial list. For example, the following episode was assigned to 'Challenge'.

Table 1. An initial list of the components.

– Customer need	– Cost
– Eagerness	– Jealousy
– Usefulness/Utility	– Challenge
– Personality	– Sociability
– Self expression	– Control
– Unique	– Security
– Independence	– Trust
– Confidence	– Loyalty
– Relaxation	– Addiction
– Novelty	– Customizability

The title of episode: Sending SMS (Short Message Service)

Beginning and end of time: 5:05 am~6:20 am

Place: In the subway

Affair: We, I and my old friend, texted each other for nothing important.

Feelings related to this affair (including reasons for the feeling): When I just purchased this mobile phone, it was very inconvenient to text. So it was really slow. But, now, I can text fast. It makes me feel good and happy. I seem to be texting non-stop these days. I'd better get improving my SMS skill a little more.

(Author's comment: In Korea, several methods are used to input the Korean alphabet on mobile devices. The methods depend on manufacturer, and differ from each other in many ways. Accordingly, users who are accustomed to a certain method for inputting the Korean alphabet often have).

3.3 *Revising components of user value*

A new concept was generated if there was no component in the initial list that matched a certain episode. For example, fun and relaxation were added to the list. Finally, the components were organized and a definition of each component was developed. Table 2 shows the components of user value, their definitions, grouping results, and the frequency of matching to collected episodes.

4 DISCUSSION

There were differences in the frequency of each component depending on the user group, especially related to the period of ownership. The frequencies of 'Novelty', 'Preciousness' and 'Challenge' in the 'within a week after purchase' group were much higher than those in the '6 to 12 months after purchase' group. This shows that the types of experiences could change over time after the purchase of a product. Thus, the time factor is one of the important factors in UX (Karapanos et al. 2009).

Table 2. Components of user value.

Components	Frequency	Definition
Customer need	96	Degree to which functions or appearances of a product/service satisfy the user's needs
Usefulness/Utility	56	Degree to which a product/service has a beneficial, practical use
Eagerness	27	User's perception of having keen interest or intense desire
Customizability	7	Degree to which a product/service is changed or built easily to fit personal specifications or preferences
Expectation	6	User's act or state of looking forward or anticipating
Attachment	66	Ability for the user to attach subjective value to a product/service
Preciousness	40	Degree to which a product/service is valuable, precious to the user
Novelty	13	Degree to which a product/service is novel, new, or unique (Related term: curiosity)
Trustworthiness	13	Degree to which a product/service deserves of trust or confidence (Related terms: Belief, Trust)
Sociability	24	Degree to which a product/service satisfy the user's desire of being sociable
Social emotion	22	Degree to which a product/service set the stage where the user can feel, express or sharing their emotions socially
Social value	1	Degree to which a product/service provides the user values related to social issues, problems and reforms
Friendship	1	User's perception of having a friendly relation with other people
Pleasure	24	User's feeling of being pleased or gratified for interacting with a product/service
Fun	15	Degree to which a product/service gives the user enjoyment, amusement, or pleasure
Refresh	9	Degree to which a product/service provides the user new vigor and energy
Self-satisfaction	17	Degree to which a product/service gives the user a great satisfaction with oneself or one's achievements
Confidence	9	Belief in oneself and one's abilities reflected by a product/service (Related terms: Pride, Fullness)
Challenge	7	User's perception of achieving something new and difficult which requires great effort and determination (Related terms: Achievement)
Identity	1	User's perception of being the distinct personality of an individual (Related terms: Personality, Self expression, Unique)

Previous studies considered 'Identity' as an important component of user value, because the distinct personality of an individual can be seen when users interact with a product or a service. However, the frequency of 'Identity' was very low in this study. This shows a limitation of the DRM, the logs of experiences were collected wholly depending on the users' descriptions. Moreover, the frequency can't exactly correspond with the importance. This highlights that further research is needed to analyze the importance of each component.

5 CONCLUSIONS

The incorporation of user value differentiates this study from existing definitions of UX. In this study, a variety of behaviors and experiences were collected using an ethnographic method. A total of 216 episodes were collected and 20 components of user value were developed by analyzing the data. The results of this study can be used to understand the user value of a product or a service.

ACKNOWLEDGEMENT

This work was supported by Mid-career Researcher Program through NRF grant funded by the MEST (No. 2010-0000364).

REFERENCES

Alben, L. 1996. Quality of experience: defining the criteria for effective interaction design. *Interactions* 3(3): 11–15.

Arhippainen, L. & Tähti, M. 2003. Empirical evaluation of user experience in two adaptive mobile application prototypes. In *Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia (MUM 2003)*.

Forlizzi, J. & Ford, S. 2000. The building blocks of experience: an early framework for interaction designers. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*.

Han, S.H. & Hong, S.W. 2003. A systematic approach for coupling user satisfaction with product design. *Ergonomics* 46(13–14): 1441–1461.

Han, S.H., Yun, M.H., Kim, K.J. & Kwahk, J. 2000. Evaluation of product usability: Development and validation of usability dimensions and design elements based on empirical models. *International Journal of Industrial Ergonomics* 26(4): 477–488.

Hassenzahl, M. & Tractinsky, N. 2006. User experience-a research agenda. *Behaviour & Information Technology* 25(2): 91–97.

Hektner, J.M., Schmidt, J.A. & Csikszentmihalyi, M. 2007. Experience Sampling Method: Measuring the Quality of Everyday Life. Sage Publications Inc.

ISO 2009. Human-centred design for interactive systems, Ergonomics of human system interaction (Vol. ISO FDIS 9241-210). *Switzerland: International Organization for Standardization (ISO)*.

Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N. & Stone, A.A. 2004. A survey method for characterizing daily life experience: the day reconstruction method. *Science* 306: 1776–1780.

Karapanos, E., Zimmerman, J., Forlizzi, J. & Martens, J.-B. 2009. User experience over time: an initial framework. *Paper presented at the Proceedings of the 27th international conference on Human factors in computing systems*.

Kuniavsky, M. 2007. User Experience and HCI. In A. Sears, J.A. Jacko, (eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, 2nd Edition. Lawrence Erlbaum Associates Inc, New York.

Larson, R. & Csikszentmihalyi, M. 1983. The experience sampling method. *New Directions for Methodology of Social and Behavioral Science* 15: 41–56.

Part III: Man-machine interface & cognitive ergonomics

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The similarity index of medicine names to prevent confusion

Keita Nabeta, Takahiro Imai, Masaomi Kimura & Michiko Ohkura
Shibaura Institute of Technology, Tokyo, Japan

Fumito Tsuchiya
Tokyo Medical & Dental University, Tokyo, Japan

ABSTRACT: There are many reports of medical errors mainly caused by the similarity of drug names in Japanese such as 'アマリール' (Amaryl) and 'アルマール' (Almarl). In order to prevent such errors, one of we studied methods of measuring the similarity of drug names and implemented a search engine to find pairs of similar medicine names in Japan. However, the existing methods only evaluate sequences of common substrings. In addition, although some methods take the similarity of character (letter) shapes into consideration, they assume that the similarity of each character pair is given, usually by hand. In order to define the objective similarity index of medicine names, it is also necessary to measure the similarity of character shape objectively. In this study, we propose a method of evaluating character similarity by image processing and apply the method if defining medicine name similarity.

Keywords: medical safety, medicine names, similarity

1 INTRODUCTION

There are many reports of medical errors mainly caused by the similarity of drug names. In Japan, mix-up accidents occurred with the diabetes drug, 'アルマール' (Amaryl) being confused with 'アルマール' (Almarl), which is a drug for hypertension, and this has resulted in patient death. In order to prevent such accidents, The Ministry of Health, Labor and Welfare issued notices and raised awareness among medical experts.

In order to prevent errors caused by medicines that have names that are easily confused with others, it is necessary to avoid the approval of such medicines. For this purpose, methods to measure the similarity of medicine names have been studied.

Tsuchiya et al. (2001) proposed some similarity indexes of medicine names: *cos1* defined by the cosine value of the angle between vectors whose elements denote the existence of each letter, *htco* counting the common two head and tail letters, *edit* counting the number of times required to change the letters in one name to comply with another by insertion, deletion and replacement, *dlen* measuring the difference of name length, and *head* counting the number of continuously matched letters from the head. Based on these criteria, the 'Medicine similar name search engine' was developed and has been operated by The Ministry of Health, Labor and Welfare to avoid approval of drugs that have names that can be confused with existing ones. The system measures the similarity based on *head*, taking account of the existence of character pairs with a similar shape and the position of the prolonged sound sign (dash) and the letter for a nasal sound in Japanese.

Ohtani et al. (2006) proposed *Htfrag* (Head and Tail-weighted Fragmentary Pattern Based Measure) which is an extension of *htco*. It has two extended versions: *Vwhtfrag* (Visually Weighted Htfrag) taking account of the similarity of the letter shape, and *Awhtfrag* (Auditorily Weighted Htfrag) taking sound similarity into account. They evaluated these measures by applying them to the names in incident reports of mixing-up medicines and claimed that their methods were superior to Tsuchiya's indexes.

However, these existing methods only evaluate common sequences of substrings. In order to include the similarity contribution based on their positions, Tatsuno et al. (2006) proposed a measure based on *LSK* (Letter Sequence Kernel). However, this measure ignores similarity originating from letter shape in medicine names.

In addition, although some methods take the similarity of character (letter) shapes into consideration, they assume that the similarity of each character pair is given, usually by hand.

Yamade et al. (2008) reported subjective estimation regarding the similarity in the shape of

Katakana characters, which are mainly used in the brand names of medicines. In the experiment, subjects classified 71 Katakana letters into groups. The experimenter calculated the ratio for each character pair to belong to the same group and derived similar pairs such as 'シ' and 'ツ' (1.00), 'ソ' and 'ン' (0.94), 'コ' and 'ユ' (0.89).

However, this is not sufficient because there are many other characters used in Japanese medicine names such as Hiragana, Kanji and the Roman alphabet and it is also necessary to measure the similarity of their shape. It is not realistic to classify such a huge number of characters by hand. In addition, they have not discussed the way of calculating the extent of the similarity of names based on their character similarity indexes.

In order to measure the shape similarity of characters automatically, we utilized techniques originally used in image processing. In this study, we propose a method of evaluating the similarity of medicine names taking account of the shape similarity of characters calculated based on the template matching method. The proposed method extends the method based on LSK proposed by Tatsuno et al.

We should note that the effects of phonetic similarity are not included and that we compared name pairs whose length is equal, since we focus on the effect of the shape similarity of characters in this paper.

2 TARGET DATA

In this study, we targeted brand names of ethical drugs that are included in the 'Standard Drug Master' provided by the Medical and Devices Agency (MEDIS-DC) in Japan.

The brand names consist of brand, form and standard unit. In the case of 'アマリール1 mg '錠' (Amaryl 1 mg tablet), 'アマリール' (Amaryl) is the brand part, '1 mg' is the standard unit, which expresses active ingredient content, and '錠' (tablet) denotes its form. We considered that it is important to evaluate the similarity between brand parts, since pharmacists identify a medicine by them. We therefore focused on only the brand part.

The brand names of Japanese medicines are expressed in Hiragana, Katakana, Kanji characters, Roman alphabet, numerical characters and other symbols. In these character types, we concentrated on Hiragana, Katakana, the Roman alphabet and numerical characters since they are used to express many medicine names.

In order to calculate the shape similarity of characters, we apply the template matching method developed in the image processing area. As source data, we used character images (height: 200 px, width: 200 px) generated by the Japanese character font, 'MS Gothic' (150 points).

3 METHOD

3.1 *Shape similarity of characters*

In order to compute the shape similarity of characters, we applied a template matching method. Template matching is a general method that is used for image retrieval. By means of the algorithm, we can calculate the similarity value defined as the ratio of the number of the same colored pixels at the same location to the number of whole pixels. In this study, we digitalized font images, namely assigned 0 to a white pixel and 1 to a black pixel, and calculated the similarity of each combination between all pairs of the target characters.

After the calculation, we realized that the obtained similarity values were high for all pairs: even the minimum similarity was 0.68. In order to redefine the similarity index so as to allow the minimum value to be zero and the maximum value to be one, we normalized them using the following equation:

$$\omega_{a,b} = \frac{D(a,b) - \omega_{\min}}{\omega_{\max} - \omega_{\min}}, \tag{1}$$

where $D(a,b)$ is the similarity value between the characters a and b calculated by the template matching algorithm. The $\omega_{\max}$ and $\omega_{\min}$ denote the maximum and minimum value of similarity, respectively.

As it is desirable that less similar character pairs should not contribute to the value of the similarity index, we again redefined it to set 0 to the value that is less than the threshold t.

Figure 1 shows the values of the similarity index sorted in descending order. The values in the graph

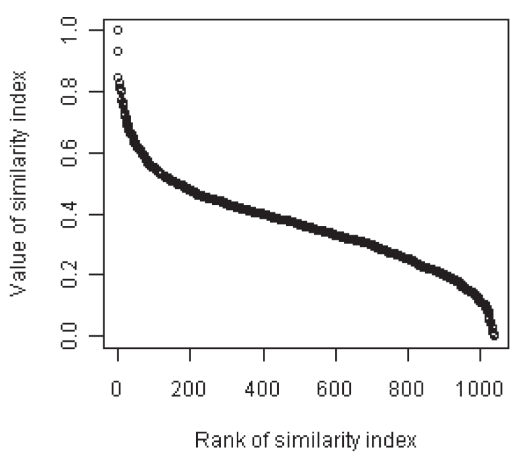


Figure 1. The values of similarity index of Katakana pairs in descending order.

are calculated only for Katakana character pairs, since most brand names are spelled with Katakana. Since we observed a steep slope in the graph within the range where the degree of similarity is between 0.65 and 1, we set $t = 0.65$.

3.2 *Method of evaluating the similarity of drug names*

In order to measure the medicine name similarity taking account of the character similarity index defined in 3.1, we extended the existing *LSK* and *htco*. Hereinafter, we denote the extended *LSK* and *htco* as *ELSK* and *ehtco*.

3.2.1 *Extended LSK (ELSK)*

The existing *LSK* is a method of evaluating similarity between strings based on common substrings and defined as the inner products of values returned by the function that counts the number of appearances of the character subsequence *u* in the target string *S* as follows:

$$S \mapsto \Phi_u(S) = \sum_{u=S[i_1,i_2,\ldots,i_n]} \lambda^{i_n - i_1 + 1 - n}, \tag{2}$$

$$K_{LSK}(S,T) = \sum_{u \in \sigma(n)} \Phi_u(S) \cdot \Phi_u(T), \tag{3}$$

Let $S[i_1, i_2, \ldots, i_n]$ be a substring constructed with characters located at $i_1, i_2, \ldots, i_n$ in *S*. Each term in the left side of Equation 2 decays by multiplying parameter λ, if the segments of the substrings exist separately. Their exponent value is the number of skips. λ is defined as follows:

$$\lambda = 1 \frac{2}{Len(S) + Len(T)'} \tag{4}$$

where $Len(S)$ and $Len(T)$ respectively express the string length of *S* and *T*.

We illustrate this method by calculating *LSK* in the case where *S* is 'アレロツク' (Allelock) and *T* is 'アテレツク (Atelec). First, substrings *u* are generated such as 'ア', 'レ', 'アレ', …, 'アレツク'. Next, $\Phi_u(S)$ and $\Phi_u(T)$ are calculated for each *u*. In the case of 'アレ', $\Phi_u(S)$ equals 1 (λ^0) because 'ア' and 'レ' exist next to each other. On the other hand, $\Phi_u(T)$ equals λ because there is one letter, 'テ' between 'ア' and 'レ'. After multiplying them, we sum them up for all substrings and obtain the value of $K_{LSK}(S,T)$. (Table 1).

This method assumes that each pair of characters is the same or not. However, in this study, we compare not only common sequences, but also all other substrings, which may contain similar characters. Assuming shape similarity between letter $S[i_j]$ and $u[j]$ as a weight, we can obtain a similarity index of strings including the character shape similarity.

Table 1. Example of calculation.

u	$\Phi_u(S)$	$\Phi_u(T)$	$\Phi_u(S)\Phi_u(T)$
ア	1	1	1
レ	1	1	1
アレ	1	λ	λ
⋮	⋮	⋮	⋮
アツク	λ^2	λ^2	λ^4
レツク	λ	1	λ
アレツク	λ	λ	λ^2

$$S \mapsto \Phi_u'(S) = \sum_{S[i_1,i_2,\ldots,i_n]} \lambda^{i_n - i_1 + 1 - n} \prod_{j=1}^{n} \omega_{u[j],S[i_j]} \tag{5}$$

Let ω be an index of character shape similarity. This value is 1 if the compared letters are the same, 0 if the degree of similarity is under the threshold t and takes a number between 0 and 1 if the compared letters are a similar character. In brief, existing LSK is a special case whose threshold is 1.

Finally, we normalized similarity to express $K_{ELSK}(S,T)$ as follows:

$$ELSK(S,T) = \frac{K_{ELSK}(S,T)}{\sqrt{K_{ELSK}(S,S)}\sqrt{K_{ELSK}(T,T)}} \tag{6}$$

3.2.2 *Extended htco (ehtco)*

The existing *htco* is the degree of similarity measuring the coincident of the head two characters or the tail two characters of the strings on which pharmacists focus. Let $\delta(a,b)$ be the delta function:

$$\delta(a,b) = \begin{cases} 0\,(a \neq b) \\ 1\,(a = b) \end{cases} \tag{7}$$

And, $S = s_1, s_2, \ldots, s_m$ and $T = t_1, t_2, \ldots, t_n$ are comparing strings.

$$htco = \frac{1}{4} \sum_{i=1,2} (\delta(s_i,t_i) + \delta(s_{m+1-i}, t_{n+1-i})) \tag{8}$$

In this study, we propose an extended *htco* (*ehtco*) that takes account of the shape similarity of characters. This index is computed as the sum of the similarity between each letter located at each two letters at the head and the tail, given as the following equation:

$$ehtco = \frac{1}{4}\sum_{i=1,2}\left(\omega_{s_i,t_i} + \omega_{s_{m+1-i},t_{n+1-i}}\right) \tag{9}$$

3.2.3 *Integration of similarity*

Finally, the index for measuring name similarity is defined as the following equation:

$$Sim\,(S,T) = \alpha ELSK + (1-\alpha)\,ehtco \tag{10}$$

Let α be the contribution ratio between *ELSK* and *ehtco*. In this paper, we set $\alpha = 0.5$.

4 EXPERIMENT AND DISCUSSION

4.1 *Shape similarity of characters*

4.1.1 *Method*

We evaluated the method of measuring the shape similarity of characters based on template matching.

First, we compared the results obtained by our method to the results reported by Yamade.

4.1.2 *Result and discussion*

Table 2 shows the top 10 letter pairs with the similarity index sorted by the score in decreasing order. We excluded the old character ‘ヲ’ and small characters such as ‘ャ’ and ‘ッ’ since they were not included in the target characters of the experiment reported by Yamade.

In this result, we can see pairs whose values of both similarity indexes, ours and Yamade’s, are high, i.e. groups such as (‘シ’, ‘ソ’, ‘ツ’, ‘ン’), (‘ク’, ‘タ’), (‘ス’, ‘ヌ’), (‘コ’, ‘ユ’) and (‘チ’, ‘テ’). This result indicates that some outputs of our method coincide with subjective evaluation. On the other hand, there were other pairs that indicate a different result such as combinations like (‘ソ’, ‘フ’), (‘レ’, ‘ン’), (‘ス’, ‘ノ’). These pairs have a common characteristic: they have the loose curve that connects the upper right to the lower left. This suggests that people cognize similarity by not only finding common component lines, but also comparing the connectivity of the component lines or the way of curving.

Table 2. Similar pairs of Katakana letters.

Character A	Character B	Our method	Yamade
ノ	メ	0.94	0.53
ク	タ	0.92	0.79
ソ	フ	0.90	0.05
シ	ン	0.90	0.86
エ	ニ	0.90	0.58
ソ	ツ	0.89	0.86
レ	ン	0.89	0.09
フ	ワ	0.89	0.73
ス	ヌ	0.89	0.85
ス	ノ	0.88	0.03

In order to present our results visually, we classified the characters by the agglomerated hierarchical clustering algorithm. The clustering method is a method of classifying many data to groups called clusters, and the agglomerated hierarchical method merges data in ascendant order of distance. In this study, we defined the distance between character a and b as $1 - \omega_{a,b}$, and applied a single linkage method in which the distance between clusters is defined to be that of the nearest pair of the elements contained in each cluster, respectively. The result is usually visualized by a tree structure, called a dendrogram, the height of whose node expresses the distance between the two merged clusters. Figure 2 shows the resultant dendrogram. From this, we can observe

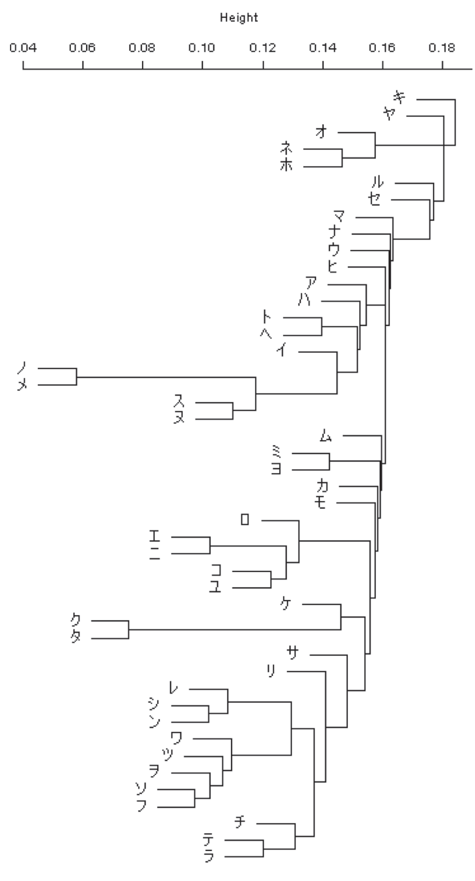


Figure 2. The result of clustering (Single linkage method).

some clusters such as ('シ', 'ソ', 'ツ', 'ン') and ('エ', 'コ', 'ニ', 'ユ') as we already discussed.

The above results indicate the validity of our method for measuring character shape similarity based on template matching. However, it is also important to handle characteristics such as edge points, branch points and folding points in order to improve our results, which should be compatible with subjectively measured similarity.

Yamade pointed out the difficulty of experiments for measuring the similarity that is experienced by people in realistic situations. In order to measure the similarity that comes purely from the shape of characters, it might be necessary to measure the accuracy of recognizing the characters instantaneously, and it will be difficult to conduct experiments since there are too many combinations of characters. In order to compare our results to the similarity that people experience, we should improve the method of measuring subjectively measured similarity.

4.2 *Similarity of drug names*

4.2.1 *Method*

In order to evaluate the proposed method, we compared the computed index values with the extent of similarity answered by pharmacists in a paper-based questionnaire.

In the questionnaire, we asked respondents to estimate the similarity of medicine name pairs on a five point scale, where '1' denotes least similar and '5' denotes most similar. The names are the stem part of existing drugs. The pairs are selected so that values of similarity distribute evenly.

The number of respondents was 129 (male: 82, female: 43, no answer: 4), the mean age was 47.4 (max: 65, min: 23) and the mean length of period that they were employed was 21.9 years (max: 38, min: 1).

4.2.2 *Results and discussion*

Table 4 shows the question numbers, pairs of drug names, mode and standard deviation of answered values and similarity index of our proposed method and the method based on the existing LSK. The column 'Delta' denotes the difference of similarity values between our method and the method based on the existing LSK. In this table, we show the top 10 Delta value pairs.

In the table, the pairs ('リピラート', 'ソピラート'), ('ペンフィルR', 'ペンフィルN'), ('タキソール', 'タキソテール') obtained a much higher value of both our index and the mean of the answered score than other pairs. This suggests that our estimation of similarity corresponds to the subjective type.

In order to estimate the superiority of our method to the existing one, we compared it with the method proposed by Tatsuno. For this purpose, we extracted pairs whose similar index is higher than the value obtained by the method based on existing LSK.

The pairs, ('エコラン', 'ユニコン'), ('グリンケン', 'ゲルソゲン'), ('ツルメリン', 'ソルノミン'), respectively have the three largest differences. Apparently we could obtain pairs that are similar because of similar shape characters.

Figure 3 shows the frequency distribution of the answered values for ten pairs whose Delta value is large. We classified the type of distribution of answers into three groups, the 'similar' group (No. 1, 19, 26), the 'dissimilar' group (No. 7, 12, 18) and the 'neutral' group (No. 6, 16, 24, 31). We discussed the correspondence between the characteristics of each group and the computed values of the similarity index.

It is clearly seen that we obtained a high score for the 'similar' group. As for No. 26, there is the tendency that the obtained value, 0.57, corresponds

Table 3. Experimental results.

No	Name A	Name B	Mode	Std dev	Our method	Tatsuno	Delta
31	エコラン	ユニコン	3	1.15	0.58	0.23	0.35
16	グリンケン	ゲルソゲン	3	1.13	0.43	0.17	0.27
6	ツルメリン	ソルノミン	3	1.01	0.57	0.30	0.27
1	リピラート	ソピラート	5	0.94	0.89	0.65	0.24
12	ノンソル	メゾカル	2	0.89	0.39	0.16	0.22
19	ペンフィルR	ペンフィルN	5	1.09	0.88	0.65	0.22
18	コートン	ユーパン	2	1.03	0.57	0.35	0.22
7	コートン	ユーワン	2	1.04	0.57	0.35	0.22
26	セスデン	ゼスラン	4	1.09	0.57	0.35	0.22
24	コータミン	ユーコシン	3	1.01	0.48	0.33	0.15

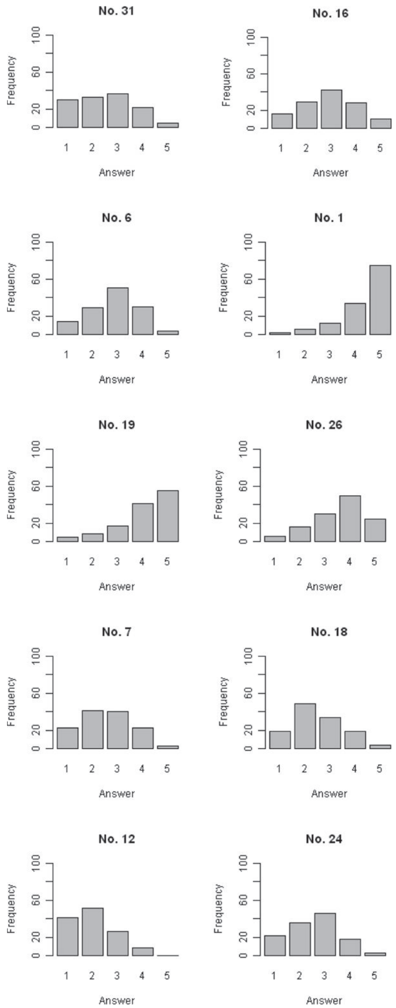


Figure 3. The frequency distribution of answered values for ten pairs whose Delta value is large.

to the answered value, because respondents tended to answer that the pair is only slightly similar.

As for the ‘dissimilar’ groups, the value of the similarity index was low in the pair of No. 12, though not in the case of No. 7 and No. 18. In the case of No. 7 and No. 18, respondents might focus on the difference of the third head characters, whose shapes are very different, though head characters’ shape, ‘コ’ and ‘ユ’, are very similar.

As for the ‘neutral’ group, the values of our similarity index were between 0.4 and 0.6. This result also suggests that our index is compatible with the subjective evaluation of similarity answered by pharmacists.

Not all respondents may judge the extent of similarity purely based on the similarity of characters. Namely, some respondents may estimate the similarity based on another viewpoint, such as phonetic similarity. As we discussed in the shape similarity of characters, it is therefore necessary to develop a method of measuring the similarity that rules out effects other than the characteristics of characters.

5 CONCLUSION

In this study, we proposed a method of measuring the similarity of Japanese drug names to prevent medical errors caused by drug confusion. We extended LSK and *htco* and proposed a similarity index taking account the shape similarity of characters. We compared our index to subjective estimation obtained by a questionnaire. The experimental results indicate that our index corresponds to the answered scores more than the existing method. In a future study, we will improve our method by taking account of the characteristics of characters’ shape other than the positions of component lines and also improve the questionnaire. We will evaluate our method with other fonts and will extend it for the application to handwritten letters.

REFERENCES

Gregorz, K. and Bonnie, D. 2006. Automatic identification of confusable drug names, Artificial Intelligence in Medicine 36: 29–42.

Hirao, T., Suzuki, J., Isozaki, H. and Maeda, E. 2005. Automatic Sentence Alignment for Monolingual Corpora, Information Processing Society of Japan 46(10), 2533–2545.

Japan Pharmaceutical Information Center. Medicine similar name search engine, https://www.ruijimeisho.jp/index.aspx

Kasuya, Y. and Yamana, H. 2006. Recognition of Similar Character Pairs with Two Types of SVMs for Online Mathematical Expression Recognition, IEICE Technical Report: 55–60.

Tatsuno, K., Kimura, M., Ohkura, M. and Tsuchiya, F. 2006. The study of Similarity Index of the Name of Drugs, TECHNICAL REPORT OF IEICE: 207–210.

Lambert, B.L., Lin, S.-J., Chang, K.-Y. and Gandhi, S.K. 1999. Similarity as a risk factor in drug-name confusion errors: The look-alike (orthographic) and sound-alike (phonetic) model. Med Care 37(12): 1214–1225.

Ohtani, H., Takeda, M., Imada, Y. and Sawada, Y. 2006. Development of the Measures to Evaluate the Similarity of Drug Brand Names, YAKUGAKU ZASSHI 126(5): 349–356.

The Medical Information System Development Center. http://www.medis.or.jp/

Tsuchiya, F., Kawamura, N., Oh, C. and Hara. A. 2001. Standardization and similarity deliberation of Drug-names, Jpn. J. Med. Informatics 21: 59–67.

Yamade, Y. and Haga, S. 2008. Subjective evaluation of similarity of appearance of Katakana characters in drug names, Rikkyo Psychological Research 50, 79–85.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Bicyclist's violation and risk perception in different situations

Pei Wang, Pei-Luen Patrick Rau, Birger van Treek, Hui Li & Jun Liu
Department of Industrial Engineering, Tsinghua University, Beijing, China

ABSTRACT: The aim of this study was to describe bicyclists' violating behavior and perception of risk in different situations in mixed traffic and to clarify whether or not bicyclists were familiar with the traffic rules and aware of their violating behavior in China. One observation and one on-line survey were done in this study. The focus of the observation was on the passing lane for bicyclists crossing. An on-line survey was conducted to explore bicyclers' risk perception and knowledge about traffic rules. Both Chinese and international participants who had the cycling experience in China responded to the survey. The results showed that bicyclists committed a lot of violations and often put themselves in dangerous situations. International participants perceived the potential risk of riding bicycle in China significantly higher than Chinese participants (p = 0.000). The knowledge of traffic rules and regulations was very well.

Keywords: bicyclist's violating behavior, risk perception, knowledge of traffic rules and regulations

1 INTRODUCTION

In China, there are enormous numbers of bicycles. Though more and more people can afford to buy a car, the bicycle still will be a very important way of transportation for Chinese. Riding bicycle is an ecological and economical means of transportation. It helps keeping the environment from becoming more polluted and also has very low acquisition and maintaining costs. As shown in Figure 1, the number of bicycles boomed worldwide in these years, especially in European countries.

Riding bicycle seems dangerous in China. Due to the rapid increase of vehicles in China, the traffic situation becomes worse. A large percentage of traffic accidents and fatalities involve bicyclists. In the past decade, bicyclist fatalities accounted for 10% to 15% of the total fatalities every year (CRTAS 2008). Although bicyclist involved accident is a serious problem of China's road safety, except for some description of bicyclist accidents (Wang et al., 2008), there are no scientific research on bicyclists street-crossing behavior, nor the research of vehicle drivers' action course during interaction with bicyclists.

It is crucial to find ways to improve the cycling situation and to protect bicyclers from traffic dangers. Therefore it is necessary to figure out more about the bicyclists' point of view and the current situation. The aim of this study is to describe bicyclists' violating behavior and perception of risk in different situations in mixed traffic. Another aim is to clarify whether or not bicyclists are familiar with the traffic rules and aware of their violating behavior.

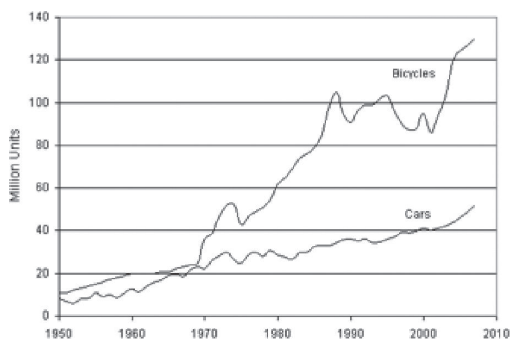


Figure 1. World bicycle and automobile production, 1950–2007 (Source: Earth Policy Institute from Worldwatch, BRIN).

2 LITERATURE REVIEW

Many researchers investigated the bicycles and motor vehicle collisions. Wang and Nihan (2004) studied the collisions between Bicycles and Motor Vehicles (BMV) occurred at intersections. The BMV accidents were classified into three categories as through motor vehicle related collisions, left-turn motor vehicle related collisions, and right-turn motor vehicle related collisions. It mentioned that most intersection BMV collisions involved turning movements of motor vehicles. Bicycle riders were at fault for most of the reviewed BMV collisions.

Moller and Hels (2008) studied the bicycle and car collisions in roundabout. The focus was on bicyclists' perceived risk in specific situations, factors influencing the perception of risk and bicyclists' knowledge about traffic rules regulating the interaction between road users in roundabouts of Danish. It was found that underestimation of risk and lack of knowledge about relevant traffic rules may contribute to bicycle-car collisions in roundabouts, and bicyclists preferred road designs with a clear regulation of road user behavior.

Most studies on perceived risk within the field of traffic safety have mainly focused on drivers but not on bicyclists. The influence of risk perception on bicyclists' behavior and accident is not well documented; the knowledge about risk perception in specific traffic situations is very rare. Detailed investigation of perceived risk in specific situations would be useful for understanding bicyclists' behaviors. First, such knowledge enables a comparison of perceived risk and actual risk of a certain behavior or action. This is valuable in the process of clarifying whether or not the bicyclists' perception of risk or the lack of knowledge about traffic rules is a contributing factor in car-bicycle accidents.

Rasanen and Summala (1998) studied the bicycle-car crashes, and the result suggested two main mechanisms producing bicycle-car collisions. The first is improper allocation of attention which is primarily related to visual search strategies, in which drivers may ignore an unexpected bicyclist. The other mechanism involves misplaced expectations of the behavior of the other party. Bicyclists supposed in many cases that the driver would give way as required by the law.

Daniels et al. (2008) studied the influence of road on bicycle accidents and did an observation of the effects of roundabouts on bicycle accidents. It was found that the bicycle accidents differ depending on whether these roundabouts are built inside or outside built-up areas. When it is inside built-up areas, the construction of roundabouts increased the number of injury accidents involving bicyclists by 48%. For accidents causing fatal or serious injuries inside built-up areas, an average increase of 77% was found.

There have been few studies about bicyclists in China and even less are formulated in English. Mixed traffic flows having different road users sharing the same roadway is an important characteristic of Chinese traffic. Bicycle lanes are always not enough because part of the bicycle lane is often used by motor vehicle. So bicycles share the sidewalks with the pedestrians in many cities. These problems have aroused intense debate among researchers and road users in China (Wang et al., 2007).

3 METHODOLOGY

3.1 *Observation*

In order to explore bicyclist's violating behavior in different situations and their risk perception and knowledge about traffic rules, two major parts of work were done in this study: one observation and one on-line survey.

Locations: Interactions are busy and dangerous for bicyclists. There are conflicts between bicycle flows and vehicle flows. A large interaction with high traffic volume, multiple lanes and regulating traffic lights for vehicles, bicyclists and pedestrians are target locations for observation. The interaction of Chengfu Road (Location A) and Zhongguancun East Road (Location B) in Beijing were selected.

Time: Considering the traffic volume, the observation was done both in rush hours and non rush hours. Rush hours occurs on a working day from 7:00 am till 9:00 am in the morning and from 5:00 pm till 7:00 pm in the evening. Non rush hours covered the rest of the time with no such heavy traffic volume. Finally, the rush-hour observation was done from 5:00 pm till 7:00 pm. For better comparison, an equal long period of 12:00 am to 2:00 pm was selected for non rush hours.

Apparatus: A digital camera was used to record the video. The camera was mounted on a tripod. Only one side of the interaction was considered for the observation. The focus was on the passing lane for bicyclists crossing. The traffic lights for bicyclers were also recorded. While recording the video the numbers of bicyclists running red lights and cycling in the wrong direction were counted.

3.2 *Survey*

An on-line survey was conducted to explore bicyclers' risk perception and knowledge about traffic rules.

Questionnaire construction: Beginning with an introduction of the objectives of the survey, five parts of questions were surveyed. Part A was general questions about participants' demographic profile. Part B was about participants' general bicycling experience and accident experience in China. Part C described 12 different violating behaviors of bicyclists, which were developed based on the previous observation. Participants were asked for their risk perception of each behavior with a seven-point scale from 1 "not risky at all" to 7 "extremely risky". Part D had eight statements of right and wrong facts of traffic laws and road rules. These facts were developed according to laws and regulations of China's road safety. Participants were asked to evaluate whether they think these statements were right or wrong. The last part E consisted of one question to ask participants' overall evaluation of riskiness of riding bicycle in China.

Participants: Both Chinese and international participants who had the cycling experience in China responded to the survey, with similar number of participants in each group. Two languages were used. Chinese participants used the Chinese version, and the international participants used the English one.

4 RESULTS

4.1 *Observation*

A total of 441 minutes observation was conducted. Due to the moving focus in the beginning, 60 minutes of each observation were used for further analysis at both locations and rush hour and non-rush hour. Generally the four kinds of violations were very salient in the observations: (1) driving across the interaction when the traffic light was red; (2) driving in the wrong direction; (3) carrying passengers or carrying too many goods on the bicycle; (4) driving on the walkways or on the car lanes.

As shown in Table 1, at location A during non-rush hour, a total of 335 bicycles per hour were counted. Of them, 60.9% run the red light, 25.1% drove in the wrong direction, and 3.6% carried passengers on the bicycle. Dangerous interactions were defined as getting close to a moving vehicle in the range of one meter. A total of 23.3% of the bicyclists were involved in possible dangerous interactions with motor vehicles, and 17.9% of them were caused by bicyclists' running red light. At location A during rush hour, a total of 530 bicycles per hour were counted. Of them 40.6% run red light, 35.1% drove in the wrong direction and 4.7% carried passengers on the bicycles. A total of 20.0% of the bicyclists were involved in possible dangerous interactions. The bicyclists caused 12.8% of them.

At location B during non-rush hour, a total of 267 bicycles per hour were counted. Of them, 44.6% run the red light, 30.0% drove in the wrong direction, and 3.0% carried passengers on the bicycle. A total of 25.1% of the bicyclists were involved in possible dangerous interactions with motor vehicles, and 13.9% of these were caused by bicyclists' running red light. At location B during rush hour, a total of 542 bicycles per hour were counted. Of them 55.7% run red light, 32.8% drove in the wrong direction and 5.4% carried passengers on the bicycles. A total of 21.0% of the bicyclists were involved in possible dangerous interactions. The bicyclists caused 7.4% of them.

4.2 *Survey*

A total of 249 participants responded to the questionnaire. Seven of them were deleted because of incomplete data, and 242 valid responses remained. Among the 242 responses, 115 (48%) were Chinese and 127 (52%) were international participants.

Of the Chinese participants, 67% of them were in the age range of 18–25 years old, and 53% were females. 62% of international participants were in the age range of 18–25 years old, and 67% were females. Most international participants were Germany and French. Other nationalities included American, Canadian, Norwegian, Mexican, and Italian. 70.1% of them currently lived in China and the others had lived in China before. Only 35.7% of Chinese participants had driver license, while 90.6% international participants had driver license.

Chinese participants had much longer experience of riding a bicycle. 60% of them had riding a bicycle for more than 10 years, 34.8% had the experience of 3 to 10 years. For international participants, most of them had the experience of less than one year. In the past six months, 23.5% of Chinese participants collided with another road user and 3.5% had ever been involved in a severe bicycle accident. However, it was much higher for international participants. 33.1% had been in a collision with other road users and 9.4% of them were severe accidents.

In part C of the questionnaire, participants were asked for their risk perception of behaviors under different situations. Most participants perceived "Running the red light for bicycles while automobiles are crossing" and "Making a left or right turn without looking over the shoulder" as extremely

Table 1. Number of violations during the observation.

Location/Time	1	2	3	4	5	6	7
A/Non rush hour	335	204 (60.9%)	84 (25.1%)	12 (3.6%)	78 (23.3%)	60 (17.9%)	12 (3.6%)
A/Rush hour	530	215 (40.6%)	186 (35.1%)	25 (4.7%)	106 (20.0%)	68 (12.8%)	25 (4.7%)
B/Non rush hour	267	119 (44.6%)	80 (30.0%)	8 (3.0%)	67 (25.1%)	37 (13.9%)	8 (3.0%)
B/Rush hour	542	302 (55.7%)	99 (32.8%)	29 (5.4%)	114 (21.0%)	40 (7.38%)	29 (5.4%)

risky. Least points were given to the following three situations "Running the red light for bicycles when there are no cars", "Running the red light when other bicyclists or pedestrians do so", and "Crossing an interaction during green light for bicycles".

The knowledge of traffic rules and regulations was very well. Chinese participants had more correct answers (up to 15%) than the international participants for all statements. For the last statement, 65% of the participants had never heard of the regulation that for even minor violating behaviors the authorities may punish you with a fee of up to 50RMB. If cannot pay it, the police may even keep the vehicle as pledge. For the final question, international participants perceived the potential risk of riding bicycle in China significantly higher than Chinese participants (p = 0.000).

5 DISCUSSION

The results showed that bicyclists committed a lot of violating behaviors and often put themselves in dangerous situations. The awareness of the risk and knowledge of traffic regulations were well. Based on the study, there are some suggestions for improvement. First, the road design should be adjusted for bicyclists' favor. The waiting zones at intersections should be larger and clearly marked. Most bicyclists did not stop at the waiting lines. More traffic signs for bicyclists should be put up. Bicycle lanes should be marked more clearly and continued at interactions. A possible short phase of red light for bicyclists would reduce waiting time and less people will run the red light. In order to have bicyclist riding in a good manner, free cycling education could be provided for children, regarding traffic laws and regulations, hand signals and shoulder look before turning.

REFERENCES

CRTAS 2008. *China road traffic accidents statistics*. Traffic Administration Bureau of China State Security Ministry, Beijing, China.

Daniels, S., Nuyts, E. & Wets, G. 2008. The effects of roundabouts on traffic safety for bicyclists: An observational study. *Accident Analysis and Prevention* 40(2): 518–526.

Møller, M. & Hels, T. 2008. Cyclists' perception of risk in roundabouts. *Accident Analysis and Prevention* 40(3): 1055–1062.

Rasanen, M. & Summala, H. 1998. Attention and expectation problems in bicycle-car collisions: An in-depth study. *Accident Analysis and Prevention* 30(5): 657–666.

Wang, D., Liang, C., Jing, C. & Wang, G. 2007. Bicycle conversion factor calibration at two-phase intersections in mixed traffic flows. *Tsinghua Science and Technology* 12(3): 318–323.

Wang, Y. & Nihan, N.L. 2004. Estimating the risk of collisions between bicycles and motor vehicles at signalized intersections. *Accident Analysis and Prevention* 36(3): 313–321.

Wang, H, Rong, Z. & Zhu, X. 2008. Investigation in characters and injury of road traffic accident for pedestrian or non-motorized vehicle. *Road Traffic & Safety* 8: 28–33.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Human factor on driver and human cognitive reliability by driving simulator

Hidetoshi Nakayasu
Konan University, Kobe, Japan

Masao Nakagawa
Shiga University, Hikone, Japan

Tetsuya Miyoshi
Toyohashi Sozo University, Toyohashi, Japan

Hirokazu Aoki
Konan University, Kobe, Japan

ABSTRACT: For the purpose of evaluation the human factors and human error in driving, a response process on perception, cognition and action are investigated by human cognitive reliability model. In order to evaluate human factors in driver's response process of perception, the response time that a subject perceives visual stimuli presented to his field in driving forward is obtained by cognitive experiment with driving simulator. The duration time that a subject stops a car by appearing obstacles on the road is obtained by braking experiment with driving simulator. It is investigated by human cognitive reliability model how to evaluate the effect of response time on several kinds of performance shaping factors such as driver's experience, stress level and quality of interface.

Keywords: human factor, driver, driving simulator, human cognitive reliability model, perception, cognition, action

1 INTRODUCTION

It is well known that the process of perception of driver is so effective to prevent hazard such as traffic accident in a timely fashion (Lalley 1982, Brown & Groger 1988). The function of hazard perception is to process information at earliest possible stage (Soliday 1974, 1975, Renge 1998). Therefore it is useful to analyze the relationship among the factors concerning human perception and cognition. To address this issue, this study has been performed concerned with the reliability and safety in a variety of driving situations through laboratory experiments using a Driving Simulator (DS).

In this paper, a new experimental paradigm is proposed for the measuring human perception and cognition during the simulated driving of a road vehicle. The experimental system proposed in this paper is useful for the simulating driving safety and the analysis of human factors related to the awareness of hazard or risk. Another aim of this work is to evaluate hazard information relevance that exists for several different traffic situations, to forecast a human behavior during driving. It has been shown that the relation between cognition time including perception and action time when one meets with the hazard situation of driving by the simulator.

From the viewpoints of reliability and safety, a method is proposed to evaluate the response behavior of a driver by using the performance ability of the driver including human factors such as driver's experience, stress level and quality of interface by perception, cognition and action process in driving.

2 EXPERIMENTS

2.1 *Experimental paradigm*

As shown in Figure 1, the human behavior is identified by a HIP model (Card et al. 1983) in which human behavior is divided into the three processes of response times (T_p, T_c and T_m). To measure human performance, two experiments to evaluate human perception and response are developed. The first experiment is concerned with cognitive process, therefore the response time T_{pc} including T_p, T_c and T_m are measured as shown in Figure 2(a). The second experiment is a braking experiment. In this experiment the duration time until the termination of vehicle (T_2) is measured as shown in Figure 2(b).

The automobile driving simulator (DA-1102: HONDA Motor Co. Ltd.) is an advanced 6-axis motion base simulator (sway-motion device) that closely approximates vehicle dynamics in real world situations to providing driver trainees with instruction in safe driving techniques. Specifications of the DA-1102 are shown in Table 1. In the simulator driving courses can be created to incorporate many kinds of driving events.

Eye movements were recorded at 250 Hz with a spatial resolution of 0.022 degree within a range of ±30 degree in the horizontal axis and ±20 degree in the vertical axis by an eye tracking system (SR Research Ltd. EyeLink II) controlled by a personal computer, which managed the timing of the experiment and collected data. Table 2 shows the specifications of the EyeLink II.

2.2 *Cognition experiment*

As seen in Figure 3, visual stimuli appeared randomly during the driving scenario. The subject must push a button located beside the steering wheel, whenever they recognized the appearance of stimuli. In this experiment, there are two kinds of visual stimuli that are target samples and standard samples. To avoid having the subject respond by the use of habituation and forecasting, two kinds of stimuli samples were developed. If the stimuli appeared is target sample, the subject must push the button, otherwise subjects should not push the button. The ratio of the appearance of target sample to standard sample is 50%. Since the number of trials is 68, the largest number of

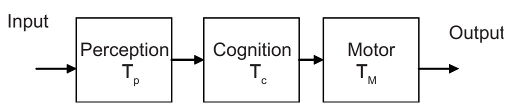
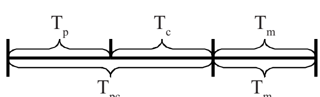
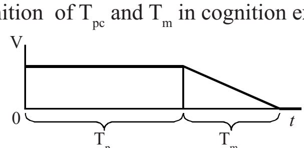


Figure 1. Human information processing model.



(a) Definition of T_{pc} and T_m in cognition experiment



(b) Definition of T_p and T_m in braking experiment

Figure 2. Response time T_{pc}, T_p and T_m.

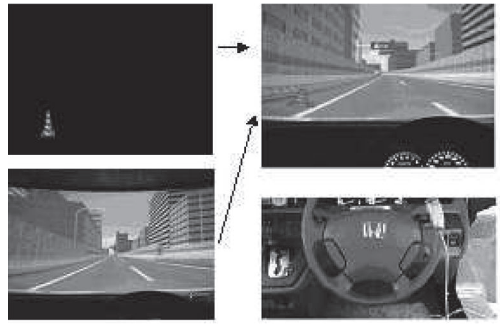


Figure 3. Visual stimuli for cognition experiment.

Table 1. Specifications of the DA-1102.

Elements of DS	Specification
Front view	Wide field (138 deg.) screen projection type
Rear view	3-mirror independent LCD display
CG	Redraw speed: 30 to 60 frame's
Mechanism	Six axis motion base system using G cylinders control
Frame	Lightweight space frame structure with aluminum extended mechanism
Body	Rear open structure fixed with FRP mold
Operation system	Steering device with reactive force control, accelerator, clutch, brake
Mission	AT/MT switch mechanism
CG computer	Cuamum3D Alchemy (1.5 million polygons)
Control system	6-axis servo amp
Dimensions	2,440 mm(D) × 2,280 mm(W) × 1,855 mm(H)

Table 2. Specifications of the EyeLink II.

Sampling rate	500 Hz (Pupil One), 250 Hz (pupil and corneal reflection)
Error of fixation	less 0.5°
Accuracy of pupil size	for diameter 0.1%
Range of trace	Horizontal ± 30° Vertical ± 20°
Data file	EDF
Rate of data transfer	3 mSec~14 mSec
Marker of infrared rays	900 nm
Eye camera	925 nm 1.2 W/cm^2
Weight	420 g

response to target sample is 34 per an experiment. The simulated highway course developed for the cognition experiment had a total length of 6.0 km where vehicle speed is limited to be less than 80 or 100 km/h. Thus the experiments consisted of the combination of control factors such as vehicle speed, day and night time, and two coefficients (1.0 for fair weather, 0.75 for rainy weather) of friction between vehicle and ground. In this experiment, T_{pc} was measured in Figure 2(a).

2.3 *Braking experiment*

To measure T_p and T_m, the braking experiment was performed on a simple straight-ahead course. In this experiment, subject must halt the vehicle by the brake pedal when he or she recognizes the appearances of a truck in a transverse as Figure 4. This experiment was also performed using the same control factors as the cognition experiment. To keep the instructions consistent, the following were given to the subjects.

1. Subject must step on the brake pedal as soon as possible when he or she recognized the appearances of a truck in a transverse temporal.
2. Subject must make an effort to stop vehicle in as small of a distance as possible, continuing to step on the pedal without releasing.
3. The appearance of a truck will be random.
4. 15 iterations will be performed per an experiment.

2.4 *Subjects*

In this experiment, 26 Japanese male students from 21 to 24 years old and 11 Japanese aged male persons from 60 to 64 years old were selected as subjects for DS experiments. There are two kinds of student subjects such as trained and untrained subjects. The numbers of former and latter subjects are 12 and 14. The former drove DS for 30 minutes a week over 6 months, and the latter had no experiences in driving DS. All subjects had regular class automobile licenses in Japan and had normal vision or corrected normal vision, and received explanation concerned with the aim and method of the experiment and provided informed consent.

Figure 4. Appearance of truck in braking experiment.

3 EXPERIMENTAL RESULTS

3.1 *Distribution of individual data and HCR model*

All the response data of T_{pc}, T_p and T_m for three categories of subjects on their driving experience such as trained, untrained and aged subjects were analyzed in order to estimate the Weibull distribution parameters by the linear regression analysis of double logarithm transformation of the equation

$$P(t) = \exp\left[-\left(\frac{t-B}{A}\right)^C\right] \quad (1)$$

In this estimation, the location parameter B was determined when the absolute value of correlation coefficient is largest. It is shown in Figure 5 as an example of the estimation of response time $t = T_{pc}$ of untrained subject at 80 km/h. It is noted that the 34 experimental observation probabilities obtained by mean rank method are plotted in Figure 5. It is seen that the observation probabilities by experiment are good fit to the estimated regression line, because of coefficient of correlation $\rho = -0.9874$. Thus, it is seen that the response time T_{pc} was good fit to Weibull distribution. It was also confirmed that T_p and T_m were good fit to Weibull distribution at all the combination of trained, untrained and aged subjects.

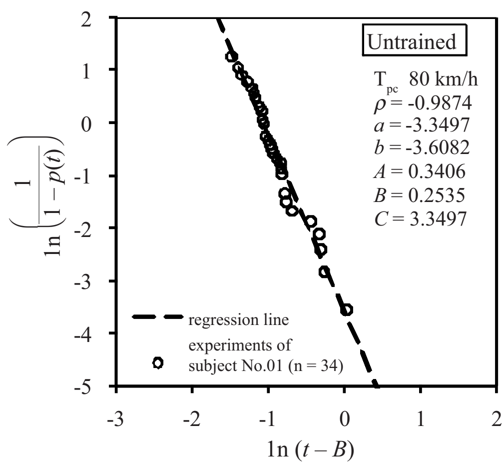


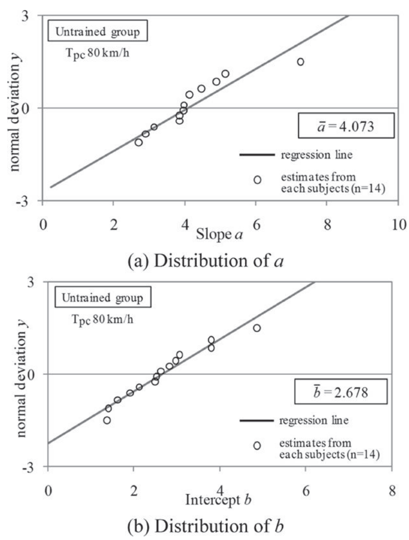
Figure 5. Relation between non-response probability and T_{pc} in Weibull probability paper (80 km/h, untrained subject).

3.2 *Parameter estimation for categorized groups*

After obtaining the estimation from individual response time as T_{pc}, T_p and T_m or normalized response time by double logarithm transformation of Eq. (1), the slope a and intercept b of the regression model estimated from each subjects were plotted in the normal probability paper as a linear line as shown in Figure 6(a) and (b). As for three categorized groups such as trained, untrained and aged drivers, it is confirmed that all the estimates of a and b from response time as T_{pc}, T_p and T_m obtained by the same method mentioned the above are applied to normal probability paper. From these results, it is seen that these estimates of a and b are good fit to the normal distribution. Therefore, the mean values of a and b are used to estimate the parameters A, B and C of Weibull distribution for categorized groups.

From these points of views, all the estimates of Weibull parameters A, B and C were determined by the mean value of regression coefficients a and b for each categories group of subject such as trained, untrained and aged subject groups. These summaries of the estimated Weibull parameters for response time t = T_{pc}, T_p and T_m are listed in Table 3. In the table, the mean and variance are calculated by

$$\text{mean} = B + A\Gamma\left(1+\frac{1}{C}\right)$$
$$\text{var} = A^2\left\{\Gamma\left(1+\frac{2}{C}\right) - \Gamma^2\left(1+\frac{1}{C}\right)\right\} \tag{2}$$



(a) Distribution of *a*

(b) Distribution of *b*

Figure 6. Distribution of estimated *a* and *b* of T_{pc}, on Normal Probability Paper. (80 km/h, untrained subject).

Table 3. Statistics of estimates of T_{pc}, T_p and T_m on Weibull distribution (80 km/h).

(a) Trained

Trained	T_m	T_p	T_{pc}
median	2.404	0.473	0.586
A	1.497	0.286	0.401
B	1.180	0.208	0.233
C	60.376	11.315	4.605
a	60.376	11.315	4.605
b	-24.340	14.145	4.209
mean	2.663	0.481	0.599
var	0.001	0.001	0.008

(b) Untrained

Untrained	T_m	T_p	T_{pc}
median	2.440	0.512	0.729
A	1.174	0.301	0.503
B	1.225	0.238	0.261
C	31.233	9.254	4.073
a	31.233	9.254	4.073
b	-5.016	11.102	2.800
mean	2.379	0.523	0.717
var	0.002	0.001	0.016

(c) Aged

Aged	T_m	T_p	T_{pc}
median	2.547	0.616	0.870
A	1.291	0.361	0.621
B	1.240	0.260	0.310
C	25.209	7.325	3.623
a	25.209	7.325	3.623
b	-6.444	7.462	1.728
mean	2.504	0.599	0.869
var	0.004	0.003	0.029

where $\Gamma(\bullet)$ means gamma function. From Table 1, it is interesting that the order of the values of C is

$$T_{pc} < T_p < T_m$$

for any category groups.

3.3 *PSF and HCR model*

In order to formulate of human performance and human error, the relationship between nonresponse probability of human and normalized time by Eq. (2)

$$P(t) = \exp\left[-\left(\frac{t/T_{1/2} - B}{A}\right)^C\right] \tag{3}$$

is proposed and this relationship is called as HCR model (Hannaman & Worledge, 1988, Kumamoto & Henly, 1996). $T_{1/2}$ in Eq. (3) is a normalized factor obtained by the following translation with K_1, K_2 and K_3

$$T_{1/2} = (1 + K_1)(1 + K_2)(1 + K_3)\overline{T}_{1/2} \tag{4}$$

K_1, K_2 and K_3 are called Performance Shaping Factors (PSFs) that mean the coefficients on human error derived from experience, stress level and quality of interface of device. $\overline{T}_{1/2}$ in Eq. (4) is the median of response time such as T_{pc}, T_p and T_m. The HCR curves for normalized response times of T_{pc}, T_p and T_m for rained subject on 80 km/h are shown in Figure 7. The factor $T_{1/2}$ are used to evaluate the normalized response time as shown in Table 5(a) and (b), where the PSF is necessary to calculate in Eq. (4). In the determination of PSFs, it is derived from PSFs Table(Hannaman & Worledge, 1988, Kumamoto & Henly, 1996) that $K_1 = -0.22$ and 0.44 for trained and untrained subject respectively at the case of T_{pc} and T_p. On the other hand, it is assumed by the proportional relation on response time, $K_1 = 2.21$ and 1.09 are extrapolated from the experimental data of aged group at the case of T_{pc} and T_p. By the way, it is assumed at the case of T_m that $K_1 = 0$, 0 and 0.25 for trained, untrained and aged groups. In all the case, it is determined that K_2 and K_3 are set up as zero. These conditions of PSF's and the estimated values for HCR model are listed in Table 4. This means that the kinds of driving work on T_{pc}, T_p and T_m are approximated to the kinds knowledge, rule, and skill based operations.

For the comparison study by HCR model (Hannaman & Worledge, 1988, Kumamoto & Henly, 1996), that estimated values by the present study on perception, cognition and action on driving work are listed along with the results for kinds of cognitive types by the previous study. From the Figure 7 and Table 5, it is verified that the tendencies of estimated parameters are similar to the results by the previous work (Hannaman & Worledge, 1988, Kumamoto & Henly, 1996). This means the features of human error on driver are similar to that of operators of nuclear power plant.

Table 4. Parameter of PSFs.

		T_m	T_p	T_{pc}
Trained	K_1	0.00	–0.22	–0.22
	K_2	0.00	0.00	0.00
	K_3	0.00	0.00	0.00
Untrained	K_1	0.00	0.44	0.44
	K_2	0.00	0.00	0.00
	K_3	0.00	0.00	0.00
Aged	K_1	0.25	2.21	1.09
	K_2	0.00	0.00	0.00
	K_3	0.00	0.00	0.00

Table 5. Estimated values on HCR model, (80 km/h, trained subject).

Trained	T_m	T_p	T_{pc}
median	2.404	0.473	0.586
K_1	0.00	-0.22	-0.22
$T_{1/2}$	2.404	0.369	0.457
A	0.630	0.775	0.880
B	0.629	0.572	0.503
C	58.868	11.315	4.605
mean	1.107	1.313	1.307
var	0.000	0.006	0.039

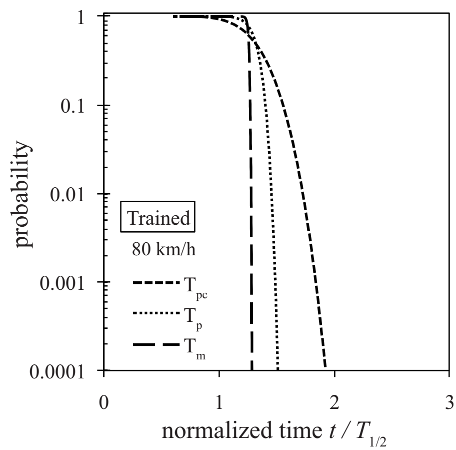


Figure 7. Relation between non-response probability and normalized time for Trained; 80 km/h, T_m, T_p and T_{pc}.

4 DISCUSSIONS

Figure 7 shows the differences of HCR model among T_{pc}, T_p and T_m that are concerned with perception, cognition and action at driving work. The values of PSF's in Table 4 are used to describe the HCR curves. From these results, it is suggested that the differences among T_{pc}, T_p and T_m are corresponding to perception with knowledge, rule and skill based behaviors.

Figure 8(a) and (b) show the relation between reliability that means probably of response and T_{pc}, and T_m for aged, untrained and trained groups.

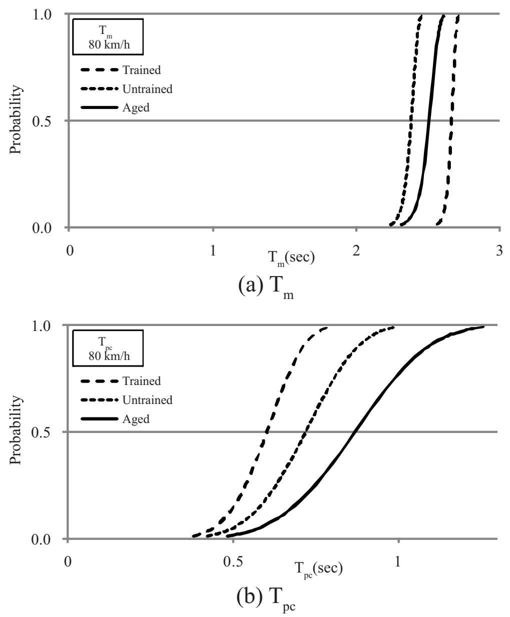


Figure 8. Relation between reliability and T_m and T_{pc} (80 km/h).

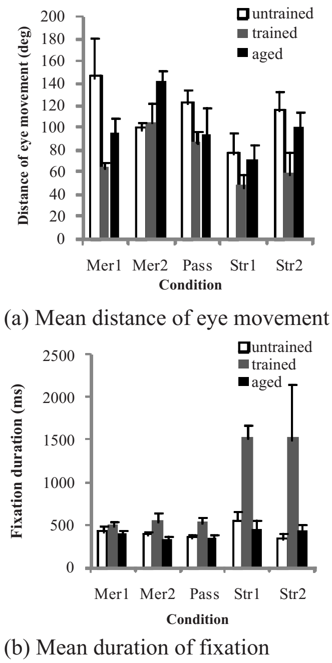


Figure 9. Eye movements at event sections.

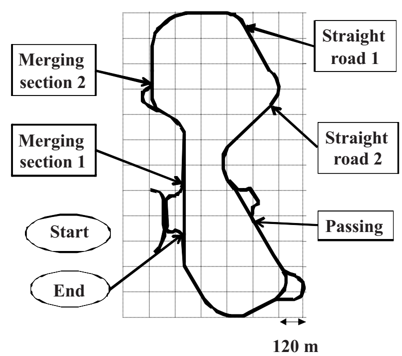


Figure 10. Event section in highway.

These figures are derived from the HCR model. From these results, it is seen that the differences between trained and untrained stems from the experiences or learning on driving technique. From these results, it is seen from Figure 8(a) that the response T_{pc}, of trained subject is faster than untrained and aged drivers. On the other hand, it is not seen from Figure 8(b) that the response T_m of untrained is faster than that of trained subject.

It is also seen that the response in perception and cognition process of aged driver is slower than trained and untrained drivers. From the study on eye movement of drivers, it is known by Figures 9(a) and (b) that distances of eye movement of trained driver at highway are smaller than those of untrained and aged drivers, while the fixation time of trained driver is longer than those of untrained and aged drivers. In the Figure 9, the abbreviations of traffic situation are defined by Figure 10. These facts suggest that the trained driver who has more experiences and well skill is able to catch the useful information during minimum opportunities of eye movement.

Finally, as an application of estimating the hazard, Figure 11 shows the possible distance for halting a vehicle with the reliability level when the trained person drives with 80 km/h. This figure can be obtained from HCR model in Figure 8(a) and (b) with the ordinal psychological case of K_1, K_2 and K_3 in Table 5. On the other hands, it can be

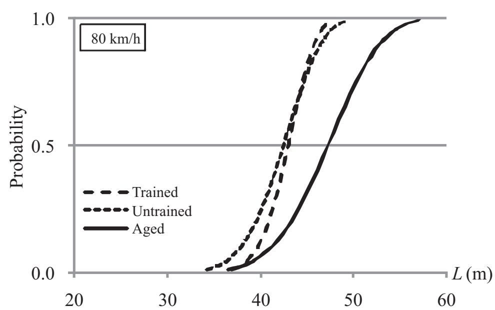


Figure 11. Distribution of possible distance $L = L_{pc} + L_m$ (80 km/h).

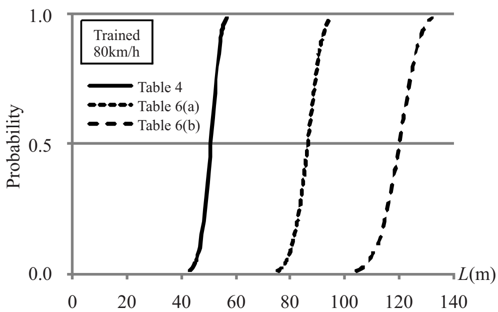


Figure 12. Distribution of possible distance $L = L_{pc} + L_m$ at several conditions (80 km/h).

Table 6. Parameter of PSFs on each condition.

(a) High workload and fair interface

		T_m	T_{pc}
Trained	K_1	0.00	-0.22
High workload	K_2	0.28	0.28
Fair	K_3	0.44	0.44

(b) Grave emergency and poor interface

		T_m	T_{pc}
Trained	K_1	0.00	-0.22
Grave emergency	K_2	0.44	0.44
Poor	K_3	0.78	0.78

estimated from Figure 12 by the possible distance for emergency stopping a vehicle will be longer than the ordinary distance, if the degree of hazard is changed by the stress level and quality of interface of device as Table 6.

5 CONCLUSIONS

In this paper it is verified by HCR model that the experience of a driver is an important factor to perception with cognition behavior such as decreasing the degree of risk when one must stop a vehicle unexpectedly. It can be possible by the proposed model that how longer distance will be necessary to stop the vehicle when the degree of the emergency and hazard increases. Statistical analysis found significances in factors with PSF, such driving experience and aging. These factors will be useful for determining reliability and safety for driving under a variety of conditions and will serve as the basis for future human factors experiments to uncover the causes of accidents.

REFERENCES

Card, S.K., Moran, T.P & Newell, A. 1983. *The psychology of human-computer Interaction*. Lawrence Erbium Associates Inc.: 23–97.

Brown, I.D. & Groger, J.A. 1988. Risk perception and decision taking during the transition between novice and experience driver status. *Ergonomics* 31: 585–597.

Hannaman, G.W. & Worledge, D.H. 1988. Some developments in human reliability analysis approaches and tools. *Reliability Engineering and System Safety* 22: 235–256.

Kumamoto, H. & Henly, E.J. 1996. *Probabilistic Risk Assessment and Management for Engineers and Scientist*, *Second Edition,* IEEE Press: 471–534.

Lalley, E.P. 1982. *Corporate uncertainty and risk management.* Risk Management Society Publishing.

Renge, K. 1998. Drivers' hazard and risk perception, confidence in safe driving, and choice of speed. *IATSS RESEARCH* 22: 103–110.

Soliday, S.M. 1974. Relationship between age and hazard perception in automobile drivers. *Perceptual & Motor Skills* 39: 335–338.

Soliday, S.M. 1975. Development and preliminary testing of a driving hazard questionnaire, *Perceptual & Motor Skills* 41: 763–770.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The effects of various naturalistic conversations on driving performances during simulated driving

Hung J. Chen
Department of Industrial Engineering and management, Ching Yun University, Chung Li, Taiwan

Chiuhsiang Joe Lin
Department of Industrial and Systems Engineering, Chung-Yuan Christian University, Taoyuan, Taiwan

ABSTRACT: There is evidence that the cell phone while driving caused the adverse effects of driving safety. However, it was hardly conducted naturalistic conversation in past driving-phone research. Hence, the actual effects of cell phone use remains unclear. The objective of this research was to assess the effect of naturalistic conversation on driving behaviour and provide a direct comparison of the driving performance of hands-free phone drivers among naturalistic conversations. The experiment was conducted on a high—fidelity driving simulator, with 24 participants divided into driver and conversation partner. The conversation was classified into theme, elicitation, and negotiation conversation by referring to the dialogue of the public in Taiwan. The results showed the purposive task (elicitation and negotiation) led to driver's workload increased and caused adverse effects in driver behavior compared to normal driving (no conversation), including: poor speed management, increased reaction time, and impaired lateral position maintenance. In addition, drivers might adopt compensation strategies to cope with the additional expenditure of attention to the demand of conversation, such as longer headway and lower driving speed. On the other hands, there were no significant difference between theme conversation and normal driving in terms of lateral position, speed management, and subjective mental load and the extent of impairment on some aspects of driving performance was less in theme than that in purposive task. These findings indicate the impairment associated with using naturalistic conversation were partially similar with that in artificial conversation from prior studies and show the deteriorations of driving performance caused by different types of naturalistic conversation were not same.

Keywords: driving distraction, naturalistic conversation, driving simulator

1 INTRODUCTION

The use of cell phones while driving is widespread growth as well as cell phone popularity. The cell phone not only produces the convenience but also the traffic safety hazards from crash reports.

On the basis of the prior research, legislations regarding the usage of cell phone while driving were established in as many as 40 countries. It was widely believed that the manual manipulation of cell phone had negative effects on driving (Brookhuis, de Vries, & de Waard, 1991; Briem & Hedman, 1995). Hence, in most cases, the law mainly focuses on peripheral factors such as dialing or holding. However, recent studies also reported that hands-free phone doesn't provide more safer than hand-held phone while driving (Patten, Kircher, Östlund, & Nilsson, 2004; McEvoy et al., 2005; Svenson & Patten 2005; Horrey & Wickens, 2006).

Strayer, Drews, & Johnston (2003) pointed out that the duration of phone conversation can be up to two orders of magnitude greater than the duration of dialing or receiving a call. It seemed to imply that the distracting effects (cognitive distraction) of the content of phone conversation are considerable. The concern issue focuses on a driver has to divert part of his attention from the primary task of vehicle operation to phone conversation. While the distraction reached a certain degree, the driving behavior would be affected significantly. These adverse changes had been reported, including increase in drivers' reaction time to driving—related events and the number of missed signals (Consiglio, Driscoll, Witte, & Berg, 2003; Strayer et al., 2003; Treffner & Barrett, 2004), poor speed management (Rakauskas, Gugerty, & Ward, 2004), increase in the likelihood of getting into an accident (Lin & Chen, 2006), and impairment of vehicle control (Treffner & Barrett, 2004). Although there has been a large of research concerned the impact of hands-free phone use on driving task since the mid 1990s, the conflicting findings has provoked a controversy (Horrey & Wickens, 2006). It is worthy of attention that the conversation tasks among prior studies vary in their nature, duration and difficulty. The amounts of attention diverted to those conversations at the wheel should be not identical because of the amount and complexity of a task decides the extent of cognitive processing required to invest in (Blanco, Biever, Gallagher, & Dingus, 2006). Certainly, the conversation is one of potential sources of variation in effects of cell phone conversation imposed on driving performance (Dragutinovic & Twisk, 2005).

Reviewing the existing literature, the simulated conversations can be divided into two types: artificial conversation and naturalistic conversation. Artificial conversation has two common forms. First, the majority of studies had used the mathematical tests (Hancock, Lesch, & Simmons, 2003; Patten et al., 2004; Shinar et al., 2005; Liu & Lee, 2005; Lin & Chen, 2006), the driver was usually asked for computations or recognition of presented digits, such as double-digit additions. Second, some of the previous studies conducted verbal tests included shadowing technique and word generation task. However, the research did not give clear comments on what difference on the effect of driving performance between artificial conversation and naturalistic conversation in real life. Rakauskas et al. (2004) pointed out that artificial conversations are easier to quantify and implement than naturalistic conversations but they represented neither typical conversations nor the demands of drivers in conversing by a cell phone. Actually, this conversation is much less structured and lacking of some elements of naturalistic conversation such as emotion and memory engagement. Usually, the subject of naturalistic conversation was often determined as interesting to participants and it requested the participant to offer answers to simple or complex question. In fact, there is a gap between the conversations in existing driving-phone research and real conversation. However,the conversation were experimenter-paced and didn't get rid of one-way communication. In other words, the driver passively engaged in conversation and did not experience leading the direction and contents of communication. The conversation in the study of Ferlazzo et al. (2008) were composed of 120 questions concerning their daily life (e.g., "What kind of music do you appreciate and why?"). Beede and Kass (2006) conducted analogous conversation which was primarily visuo-spatial in nature. ("How do I get to school from train station?" or " How many windows do you have in your home"). The two past studies made same disadvantage on the conversation flow with Rakauskas's study. In the study of Shinar et al. (2005), math computations task (a sequence of numbers and operations were presented) and an emotionally involving conversation were conducted. The emotional conversation was composed of a series of questions interested by the driver would be emotionally challenging (e.g., if the driver was a sports fan of a specific team, the question would be around the team such as the reason of poor records).

To gain a more comprehensive understanding of the effect caused by cell phone conversation on driving performance, it is necessary to explore impairments from more naturalistic and various conversations while driving. Thus, the current study incorporated three types of naturalistic conversation with realistic driving simulator. There were two hypotheses that the negative effect on driving performance would occur regardless of the type of naturalistic conversation and the degree of impairment would be different among those conversations.

2 METHODS

2.1 *Participants*

There were 24 healthy volunteers who were paid for their participation. Half of them serviced as driver in the experiment, and the other half was the conversation partner. Those driver (10 males, 2 females), between 20 and 35 ($M = 24.1$, S.D. = 3.8) years of age had held a driving license at least one years and driven at least 1000 km. Conversation partner group (6 males, 6 females; mean age = 25.2, standard deviation of age = 3.7) was conducted to converse with driver.

2.2 *Apparatus*

A fixed-based driving simulator used in the experiment creates the realistic sensations in a laboratory environment and measures numerous driving performances.

2.3 *Experimental design and procedure*

Upon arriving at the experimental facility, the both participants were told the purpose of the study briefly and signed an informed consent prior to the experiment. To help participants understand the experiment fast and clear, they were presented a pre-recorded video explaining the method, procedure and precautions of the experiment. The driver were then conducted a practice drive to familiarize the driving simulator and driving scenario and they were also informed that the experimenter would sit aside to driver at all time in order to assist the process of the trial. In the primary task, the drivers must follow a lead vehicle in low traffic density of highway. The experimenter emphasized that the drivers should drive as in real traffic. Meanwhile, they were instructed to step on the brake pedal in response to the activation of the lead car's rear-lamp. The secondary task was conversing with a partner who was at other room through a simulated hands-free phone. Each driver would perform three kinds of conversation in the experiment. Those types were selected and classified by collecting the common conversation contents in vehicles of Taiwan citizen. In theme, the driver was told that it was casual talking which focused on a specific topic that was identified in the questionnaire as being of interest to both sides. In contrast, the elicitation, the driver was not only communicating around a specific theme but also trying to make the partner say a group of preset keywords relative to the topic. The elicitation simulated a dialogue that we wanted to acquire an answer or promise to a concerned thing from conversation partner. To keep the conversation smooth and natural, we informed the driver of not just asking straight forward questions. In brief, the simple mechanics of this conversation, the driver would embed their attempts to elicit the keywords so that the partner would be conscious of that's being a casual talking. The negotiation involved the commercial deal and competitive price. The driver actively bargained with the partner over the price for a high price product. Obviously, the work results of buyer and selling party is contrary. Both parties all made an effort to fight for self-interest which had relation with cash award. As mentioned above, the participants were expected to go all out to engage in each condition, they were offered extra reward depending on their performance. These conversations were highly expected to be more natural in form and content and show up its characteristic fully. Except for the three conditions, no conversation condition was also conducted, in which the driver concentrated on the driving task alone.

As to dependent variable, driving speed, lateral position, and headway were logged at a frequency of 2 Hz. Lateral position was measured as the distance between the car and the road central line. Headway was bumper to bumper distance between the lead car and the following car. In addition, brake reaction time, missed brake response, and number of collisions with lead car were recorded by event-oriented. Brake reaction time recorded in millisecond was the time interval between activation of the rear lamp of lead car and the onset of the brake pedal. The number of missed brake response divided by total number of the lead car braking was the ratio of missed brake response.

3 RESULTS

The thorough empirical evaluation of the research's hypothesis is carried out in the context of realistic driving simulation. This research compared the effects caused by the theme, elicitation, negotiation to baseline conditions on driving performance. Table 1 depicts the conversation type yielded a main effect for driving speed. The multiple comparison showed that each other of all have significant difference, the highest and the lowest speeds were in no conversation and negotiation respectively and the speed were higher in theme compared with the elicitation.

Refer to Table 1, the main effects of the conversation tasks on driving maintenance (i.e., standard deviation of driving speed and standard deviation of lateral position) were found. Tukey's test showed that the variation of speed for four conditions could be divided into two groups, i.e., elicitation, negotiation, and theme and theme and no conversation. Meanwhile, the highest and lowest variations appeared in the elicitation and the no conversation respectively. In other words, negotiation and elicitation were significantly higher than no conversation. In regard to standard deviation of lateral position, the difference between no conversation and theme was not statistically significant, but both elicitation and negotiation increased the

Table 1. Means and standard deviations for driving performance and maintenance.

Performance measures	None	Theme	Elicitation	Negotiation	p-Value
Mean speed (km/h)	94.03^{a}	91.78^{b}	88.03^{c}	85.61^{d}	0.000
Standard deviation of speed (km/h)	5.44^{a}	6.10^{a,b}	6.95^{b}	6.82^{b}	0.012
Headway (m)	8.15^{a}	9.64^{b}	11.35^{c}	12.07^{c}	0.000
Standard deviation of lateral position (m)	0.60^{a}	0.66^{a,b}	0.76^{b,c}	0.82^{c}	0.000
Brake reaction time (sec)	0.91^{a}	1.15^{b}	1.17^{b}	1.39^{c}	0.000
Ratio of missed brake response	0.02	0.14	0.24	0.27	0.000
Number of collision	0.08^{a}	0.17^{a}	0.42^{a}	0.83^{b}	0.014

Note.- In rows ^{a,b,c,d} indicates differences between conditions.

variability of driver's lane maintenance behavior noticeably compared with no conversation. Seeing the result in detail, there was a significant difference between theme and negotiation, but two differences, one is between theme and elicitation, the other is between elicitation and negotiation, were no statistically significant.

The results illustrated that the conversation was also found to be significant for headway (see Table 1). Obviously, the headway in theme was significant longer than no conversation and shorter than elicitation and negotiation, no significant difference between elicitation and negotiation. In terms of the brake reaction time for detecting the lead car's rear-lamp, the negotiation was significant longer than other conversation conditions (shown in Table 1). Theme and elicitation had similar effects in brake reaction time. It is clear that when the participants concentrated on driving task alone, their reaction time were significantly shorter than driving with the three conversations.

Table 1 also indicates that there was significant difference in the number of collision among the four driving conditions. In addition, no significant difference among no conversation, theme, and elicitation and between elicitation and negotiation. The number of collision in no conversation and theme were significant lower than that in negotiation.

4 DISCUSSION

The main findings are that the naturalistic conversation indeed has a negative effect on driving behavior and the degree of degradation in whichever measures were marked discrepancy among the three conversations. The theme had greatly part of communication in our daily life. The drivers in the present study had large experiences of engaging in casual talking while driving, the talk topic was chosen by the driver prior to the experiment, and thus he was interested and good at the content of theme. Meanwhile, no goal was required to achieve in theme and thus the driver conducted it in haphazard and relaxation. In addition, the conversation could be interrupted at any moment without memorizing the progress by the driver while getting into high demand of traffic because it could be restored at new talking point. The architecture of the conversation was loose and the content was shallow. In principle, the driver's involvement of mental resource in theme might be relatively low. In terms of elicitation the driver needed to rehearse the keywords in his mind constantly. The driver should organize the conversation flow to keep it smooth and induce the specific words spoken by the partner successfully. The driver would make the partner say the words as many as possible in restricted conversation time. Therefore, the driver would play a leading role in the conversation and time constraint increased the stress levels (Hill and Boyle, 2007). Negotiation, the driver bargained continuously with conversation partner over the price that needed to concentrate on the status and progress of the conversation, involving decision-making, problem solving and memory. For the same reason, both driver and conversation partner desired a higher performance of the secondary task. Consequently, the driver felt depressed and stressed easily because of the strong resistance of the conversation partner during negotiation. Even if elicitation and negotiation fell into the category of purposive tasks and produced internal stress in driver's mental, there were apparent differences between them; the driver needed to copy with and respond the declaration of the opponent passively and continuously in negotiation. In other words, the driver required the more attention to apply appropriate strategies to contend with an opponent. Therefore, we speculated the driver's involvement of mental resource in negotiation was higher than that in elicitation.

Törnros and Bolling (2006) found driving speed was reduced by phone conversation; the authors interpreted it as the compensatory effort. Haigney et al. (2000) also claimed that the driver engage in compensatory behavior which enables mental resource availability to be maintained at a safety level when using a cell phone. In other words, drivers apparently intented to resolve the conflict of attentional demands by reducing driving speed allowing more time for both work. From the point of view, there was no doubt that the driving speed in other conversation conditions should be slower than that in no conversations. Further viewing on the significant difference of driving speed among three conversations implied the cognitive workload resulted from various conversations really had dissimilarity that supported our main hypothesis.

On the basis of the same reason, increase of headway providing more time to react is another way of reducing task demand (Cnossen et al., 2004). It was reasonable that those conversations revealed similar pattern of results for both measures of performance. From the results of driving speed and headway, the driving workload involved in theme markedly higher than that in no conversation significantly differed that while the two purposive conversations involved in. Moreover, no significant difference of headway between elicitation and negotiation but it was close to significance. On the aspect of the driving speed, the negotiation was substantially slower than elicitation. It was inferred from the findings that the load caused by negotiation should be higher that by elicitation which was in agreement with the inference about the comparison of the two conversations in last paragraph.

Strayer and Johnston (2007) suggested that a driver involved in a cell phone conversation would lead to withdrawal of attention from driving environment. The viewpoint supported the result in the study; the reaction time toward lead vehicle braking was prolonged compared to no conversation driving regardless of the conversation type. The finding appeared to be consistent with the relevant literatures (i.e., Strayer et al., 2003; Consiglio et al., 2003). As noted above, the driver's involvement of mental resource in negotiation, including memory, decision-making, thinking, etc., was higher than that in theme. The more involvement of mental resource invested in secondary task, the more attention diverted from driving activity. It was widely accepted that the degree of impairment on the capability of perceptual signal detection was in accordance with the amount of attention being diverted to utilize in secondary task. Therefore, the braking reaction time in negotiation was slower than that in theme.

From the result of number of collision, the risk of traffic accident significantly was increased dramatically in whichever naturalistic conversation, elicitation and negotiation especially. Because of the feature of phone conversation; immediate response, the celluar phone conversation seemed to be urgent and important, the driver was probably to give the conversation task higher priority compared to driving task in a short time, especially in a smooth and steady traffic condition (i.e., following the lead car in the low traffic density). At this moment, the driver seemed to get into the conversational situation entirely and remain few mental resource to maintain the essential demands of driving task. We had experience of not remembering what action taken (i.e., braking, overtaking) or scenary passed through in a moment ago while simultaneously conversing. Hence, we considered the cell phone conversation will yield a form of inattention blindness, is the phenomenon of not being able to see things that are actually there and increases the probabilty of car accident certainly. Furthermore, inattention blindness would occur more frequently in the purposive conversation because of a conversational goal that emphasized the importance and essentiality of conversation.

The result of variation of lateral position demenstrated that the conversation seem to impair the ability to maintain the steady lateral position, was inconsistent with the claim of Rakauskas et al .(2004) and Lin and Chen (2006). We suspected that the discrepancy could be attributed to difference of the reality of conversation. The naturalistic conversation usually involved fluctuations of emotions and the intense or extreme (e.g., anxious, excited, angry) emotion aroused a slight vibration of hands that directly affected the control of steering. Owing to care about the goal of conversation, the driver had larger variety of emotion in elicitation and negotiation than that in theme. The conversations in the experiment of Rakauskas et al. (2004) and Lin and Chen (2006) were a form of artificial test which did not cause the emotional fluctuations and thus did not influence the maintenance of driving position.

5 CONCLUSION

The subjective evaluation and objective driving measurement lead us to conclude that the different contents will not result in the same mental load and negative effect.

The finding interprets why the results from different studies seemed contradictory and supports the notation that the naturalistic conversation should be adopted to explore the precise results about the effect on driving behavior caused by cell phone.

Although casual talking is superior to purposive conversations in terms of some driving performances and subjective workload, we don't agree that it is entirely safe to conduct. First, it is impossible for the driver to ensure that none of other types of conversations will occur in the whole conversation process. The second explanation, the findings of this study are applicable for highway driving, it is lacking of evaluation on other traffic condition.

The range of circumstances investigated in this study was limited and the more different traffic densities are needed to examine the driving behavior imposed by the naturalistic conversation such as urban traffic. Therefore, the further research would assist generalizations to be made.

REFERENCES

Alm, H. & Nilsson, L. 1994. Changes in driver behaviour as a function of handsfree mobile telephones: A simulator study. *Accident Analysis and Prevention* 26(4): 441–451.

Beede, K.E. & Kass, S.J. 2006. Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis and Prevention* 38: 415–421.

Blanco, M., Biever, W.J., Gallagher, J.P. & Dingus, T.A. 2006. The impact of secondary task cognitive processing demand on driving performance. *Accident Analysis & Prevention* 38: 895–906.

Briem, V. & Hedman, L.R. 1995. Behavioural effects of mobile telephone use during simulated driving. *Ergonomics* 38: 2536–2562.

Brookhuis, K.A., de Vries, G. & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*, 23(4): 309–316.

Cnossen, F., Meijman, T. & Rothengatter, T. 2004. Adaptive strategy changes as a function of task demands: a study of car drivers. *Ergonomics* 47(2): 218–36.

Consiglio, W., Driscoll, P., Witte, M. & Berg, W.P. 2003. Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. *Accident Analysis and Prevention* 35(4): 495–500.

Dragutinovic, N. & Twisk, D. 2005. *Use of mobile phones while driving—effects on road safety: A literature review.* R-2005-12. SWOV, Leidschendam.

Ferlazzo, F., Fagioli, S., Nocera, F.D. & Sdoiaa, S. 2008. Shifting attention across near and far spaces: Implications for the use of hands-free cell phones while driving. *Accident Analysis & Prevention* 40: 1859–1864.

Hancock, P.A., Lesch, M. & Simmons, L. 2003. The distraction effects of phone use during a crucial driving maneuver. *Accident analysis and prevention* 35(4): 501–504.

Haigney, D., Taylor, R.G. & Westerman, S.J. 2000. Concurrent mobile (cellular) phone use and driving performance: Task demand characteristics and compensatory processes. *Transportation Research Part F* 3: 113–121.

Hill, J.D. & Boyle, L.Ng. 2007. Driver stress as influenced by driving maneuvers and roadway conditions. T*ransportation Research Part F* 10:177–186.

Horrey, W.J. & Wickens, C.D. 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors* 48(1): 196–205.

Liu, B.S. & Lee, Y.H. 2005. Effects of car-phone use and aggressive disposition while approaching signalized intersections, *Transportation Research Part F* 8:4–5, 369–382.

Lin, C.J. & Chen, H.J. (2006). Verbal and cognitive distractors in driving performance while using hands-free phones. *Perceptual and Motor Skills, 103*, 803–10.

Matthews, R., Legg, S. & Charlton, S. (2003). The effect of cell phone type on drivers subjective workload during concurrent driving and conversing. *Accident Analysis and Prevention, 35*, 451–457.

McEvoy, S.P., Stevenson, M.R., McCartt, A.T., Woodward, M., Haworth, C., Palamura, P. & Cercarelli, R. 2005. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: A case-crossover study. *British Medical Journal* 331: 428–430.

Parkes, A.M., Fairclough, S.H. & Ashby, M.C. 1993. Carphone use and motorway driving. In T. Lovesy (ed.) *Contemporary Ergonomics.* Taylor and Francis, London. 403–408.

Patten, C.J.D., Kircher, A., Östlund, J. & Nilsson, L. 2004. Using mobile telephones: cognitive workload and attention resource allocation. *Accident analysis and prevention* 36: 341–350.

Rakauskas, M., Gugerty, L. & Ward, N.J. 2004. Effects of cell phone conversations on driving performance with naturalistic conversations. *Journal of Safety Research* 35: 435–464.

Shinar, D., Tractinsky, N. & Compton, R. 2005. Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident analysis and prevention* 37(2): 315–26.

Strayer, D.L. & Drews, F.A. 2007. Cell-phone-induced driver distraction. *Current Directions In Psychological Science* 16: 128–131.

Strayer, D.L., Drews, F.A. & Johnston, W.A. 2003. Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied* 9: 23–32.

Svenson, O. & Patten, C.J.D. 2005. Mobile phones and driving: A review of contemporary research. *Cognition, Technology and Work* 7: 182–197.

Törnros, J. & Bolling, A. 2006. Mobile phone use—Effects of conversation on mental workload and driving speed in rural and urban environments. *Transportation Research Part F* 9: 298–306.

Treffner, P.J. & Barrett, R. 2004. Hands-free mobile phone speech while driving degrades coordination and control. *Transportation Research Part F* 7: 229–246.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Cognitive process of evacuee with map on personal digital assistance

Daiji Kobayashi
Chitose Institute of Science and Technology, Hokkaido, Japan

Sakae Yamamoto
Tokyo University of Science, Tokyo, Japan

ABSTRACT: Japanese government orders the inhabitants to evacuate by self-help or by mutual assistance in urban disaster situation. When urban earthquake hazards are occurred, the bumps and obstacles on the street could block the evacuation behavior. More, the ever-changing blaze along the evacuation route jeopardizes the evacuee's life. Although the systems distributing information about the evacuation such as broadcasting network have existed, the appropriate information for each evacuee's situation should be available. The personal navigation system which informs the appropriate evacuation route in accordance with the user's physical and cognitive characteristics is a way to deal with the issues. To design the human interface of disaster information system, the evacuee's cognitive characteristics should be considered. Therefore, the evacuee's cognitive behavior was investigated through simulation experiments. As the results, an evacuee's cognitive model was proposed and the information required on the evacuation route map on PDA was revealed.

Keywords: cognitive process, cognitive model, urban disaster, evacuation, personal digital assistance

1 INTRODUCTION

1.1 *Background*

The government of Japan orders the inhabitants to evacuate by self-help or by mutual assistance in urban disaster situation. Although the systems distributing information about the evacuation such as broadcasting network have existed, the appropriate information for each evacuee should be available. The personal navigation system which informs the appropriate evacuation route in accordance with the situations of the evacuation route is a way to deal with the issues. Thus, the existence of representing the route and the user interface of the navigation device should be considered based on the evacuee's cognitive characteristics. For example, if the evacuees move to the evacuation center with the evacuation route map, we have not known that what could be clues for them to choose the route. Further, if the evacuee move with the information device providing the evacuation route information, the existence of the user interface of the information device should be considered based on the evacuee's view point.

Recently, Personal Digital Assistance (PDA) with touch panel is in widespread use in Japan. Touch panel operation is low-skilled interface for young individuals. However, various factors relating to the map on PDA could influence the evacuation behavior of out-of-towner. From these viewpoints, we conducted two experiments in order to reveal the relation between the evacuee's performance and cognitive behavior which relates to the map on PDA.

1.2 *Cognitive process of evacuee*

When an out-of-towner evacuee has a portable device such as the PDA or a mobile phone, the evacuee could refer to the map in order to proceed to an evacuation center. However, the evacuee has to decide or judge or understand the present location and the direction to the goal again and again. Whenever the evacuee thinks about the details, the cognition relating to the thinking could occur.

It was pointed that the plant operator's cognitive process consist some subtle thinking step and the thinking step includes a 'cognitive state' and a 'strategy' (Kobayashi et al., 2007). The cognitive state is a state of cognition affected by the operator's emotion while the plant operator identifies the malfunctions. The strategy is the direction to the identification of the malfunctions. Although these concepts of cognition are considered for plant operators, we assume that the concept could be applied to the evacuee. Because the evacuee's thinking process could be a sequence of problem solving such as route selection. Nevertheless, the evacuee's cognitive process should be cleared. Then, the evacuee's cognitive process should be investigated and clarified in order to consider the design of the navigation system for evacuation.

1.3 *Objective of this study*

Although some methods or systems of providing disaster information has proposed concretely, existence of the network for providing disaster information should be discussed. Especially, the systems should be based on a simulation experiment from the view of evacuee (Takahashi et al., 2009). Therefore, considering the user interface of disaster information device is useful for developing the disaster information service system. For addressing these points, we investigated the evacuee's cognitive process through experiment.

2 METHOD

2.1 *Overview*

In order to reveal the characteristics of the evacuee's cognitive process, we conducted the evacuation simulation. The participants were 25 students with ages ranging from 20 to 23 years, and the every participant was walked from the specified starting location to an evacuation center in Shinjuku, Tokyo with a PDA in hand. They were not familiar with the area. The PDA shows the map of the region and the participants operate the map by touch operation.

2.2 *Districts for experiment*

The evacuation simulation was conducted in 2 regions in Shinjuku, Tokyo. Region 1 had 200 × 150 square meters residential area and included a few apartments and small alleys. Thus, we assumed that many clogs or accidents could prevent the evacuee's behavior. On the other hand, Region 2 has an area of 400 × 300 square meters residential area and crowded city blocks of wooden dwellings and temples. Thus, many fires could break out after a quake occurs, and evacuees could be the fire victims. The residents in these regions consists many elderly people.

2.3 *Apparatus and materials*

In the evacuation simulation, the participant's think aloud protocol and the performance were recorded using 2 video cameras (SONY Handycam HDR-SR12) and a voice recorder (OLYMPUS Voice Trek V-61). One of the video cameras recorded the manipulation of PDA and another video camera recorded the participant's evacuation behavior from some distance by an experimenter.

The PDAs we used were EMOBILE S11HT and NTT DoCoMo T-01 A. S11HT has a 2.8 inches display with touch panel and T-01 A has a 4.1 inches wide VGA display with touch panel. These PDA was running Microsoft Windows Mobile 6.1 Professional Japanese edition and the custom software displaying map. The custom software was running on Adobe Flash Player. The map was made based on blank map which was published by Ministry of Land, Infrastructure, Transport and Tourism of Japan. In this regard, the map of Region 1 was almost blank excepting the name of the elementary school for the purpose of evacuation center and the town name. The map of Region 2 includes the name of the temples in addition.

The PDA displayed two semi-transparent buttons on the map for zooming and the map position was moved by rubbing the touch panel with a finger. The map was fully-zoomable by touch panel operation. Further, the map showed some impracticable ways/roads with icons. The icons represented a fire or a traffic accident after a certain period of time.

3 INVESTIGATION OF COGNITIVE PROCESS

3.1 *Experiment*

The aim of the evacuation simulation was to investigate the evacuee's cognitive process while they moved. The simulation was conducted in the 2 regions. The participants for Region 1 were 5 students and walked with the PDA (S11HT) in hand. Further, the participants for Region 2 were 10 students and walked with T-01 A in hand.

Before the participants started to proceed to an evacuation center one by one, we instructed all participants to think aloud the contents of their thinking about the evacuation.

While the respective participants moved, 3 experimenters walked with the participant, and recorded their behavior using the video cameras and asked to the participant the reason of behavior such as route choice.

After the participants reached the goal (the evacuation center), the participant asked the questionnaire including the question items about the usability of the map on PDA and etc.

3.2 *Result*

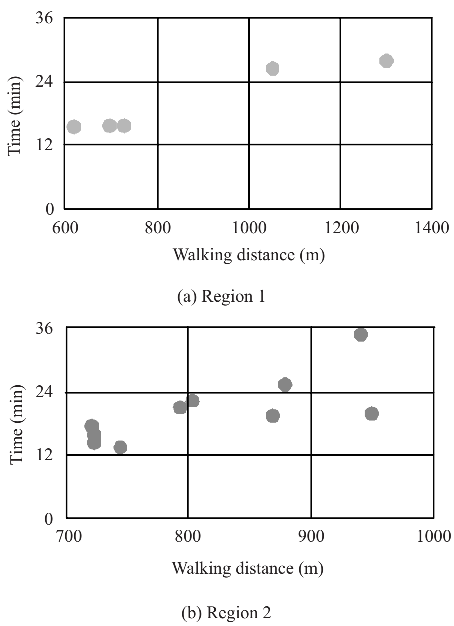
From the evacuation simulation's results, all participants could get to the goal using the map in spite of the difference of experimental regions. However, there are no participants who walked the shortest path of Region 1. The distance of the shortest path of Region 1 was about 320 meters; however, Figure 1(a) shows that every participant walked over 600 meters. In other words, all the participants who walked in the Region 1 detour or lost their way. On the other hand, the shortest path of Region 2 was about 710 meters. Figure 1(b) shows that 4 participants could walk by shortest path without backtracking. It is important that whether the evacuee is able to walk the shortest path for evacuation in disaster situation. Therefore, these results suggest that the quality of evacuation of Region 1 was worse than the evacuation of Region 2.

Based on consideration of the difference of the results between Region 1 and Region 2, we assumed that the participants who were not able to get to the goal by the shortest path were not able to get the appropriate information for evacuation from the map on PDA. In these regards, the difference of the evacuation route or distance was suggested that their cognitive process differed. Therefore, we investigated the participants' think aloud protocols and behaviors recorded by the video cameras.

3.3 *Discussion*

The think aloud protocols and their behaviors which were recorded by the video camera showed that all participants of Region 1 intended to walk wider road and avoided walking through the narrow lane. Further, the participants could not understand a narrow lane was just the shortest path because the narrow path could not tell just by looking the evacuation environment and/or the map on PDA. Thus, we assumed that these factors prevent the evacuee from getting clues for choosing preferable path to the evacuation center. On the other hand, the results also showed that the map on PDA could not tell any clue, and the participants went back and forth, and walked long distance. Those participants' think aloud protocol represented that they misunderstood their present location or direction to the goal, and repeated trial and errors until they backtracked to the familiar point.

Although the participants' data of Region 2 showed that the participants' cognitive characteristic was as well as the case of Region 1, the name of temple which indicated on the map of Region 2 and the gatepost of the temple became clues for path preference. This means the participant could match information from the map and the land mark in the evacuation environment, and they understand the present location and direction to the goal.



(a) Region 1

(b) Region 2

Figure 1. Relation between necessary time for evacuation and walking distance of respective regions.

3.4 *Evacuee's cognitive process*

According to the above results and consideration, we could estimate the evacuee's cognitive process as shown in Figure 2.

Figure 2 shows a phase of cognitive process while the evacuee proceeds to the evacuation center with a map on PDA in hand. When an evacuee has to decide the behavior such as path choice, the various factors we found in the experimental data could affect the cognition or thinking. These factors fall into 2 categories: one is external factors and the other is internal factors. The external factors exist in the evacuee's environment and the map such as evacuation route and the map on PDA. The internal factors relate the evacuee's mental state and/or decision error. For example, if the narrow path and confusing map do not tell the evacuee appropriate information, the evacuee could backtrack or trials and errors. In these regards, we assumed that the cognitive phase consist 3 steps: a cognitive state that is the evacuee is having the clue or not; strategy that is the evacuee's principle of behavior; and the evacuee's behavior e.g. their route choice and proceeding. Figure 2 show that the behavior 'proceeding based on the plan' or 'proceeding to the goal direction' is preferable because the evacuee is able to proceed to the next phase. However, the behavior we observed such as 'getting lost' or 'backtracking' are not preferable because the evacuee has no keys to deal with the situation.

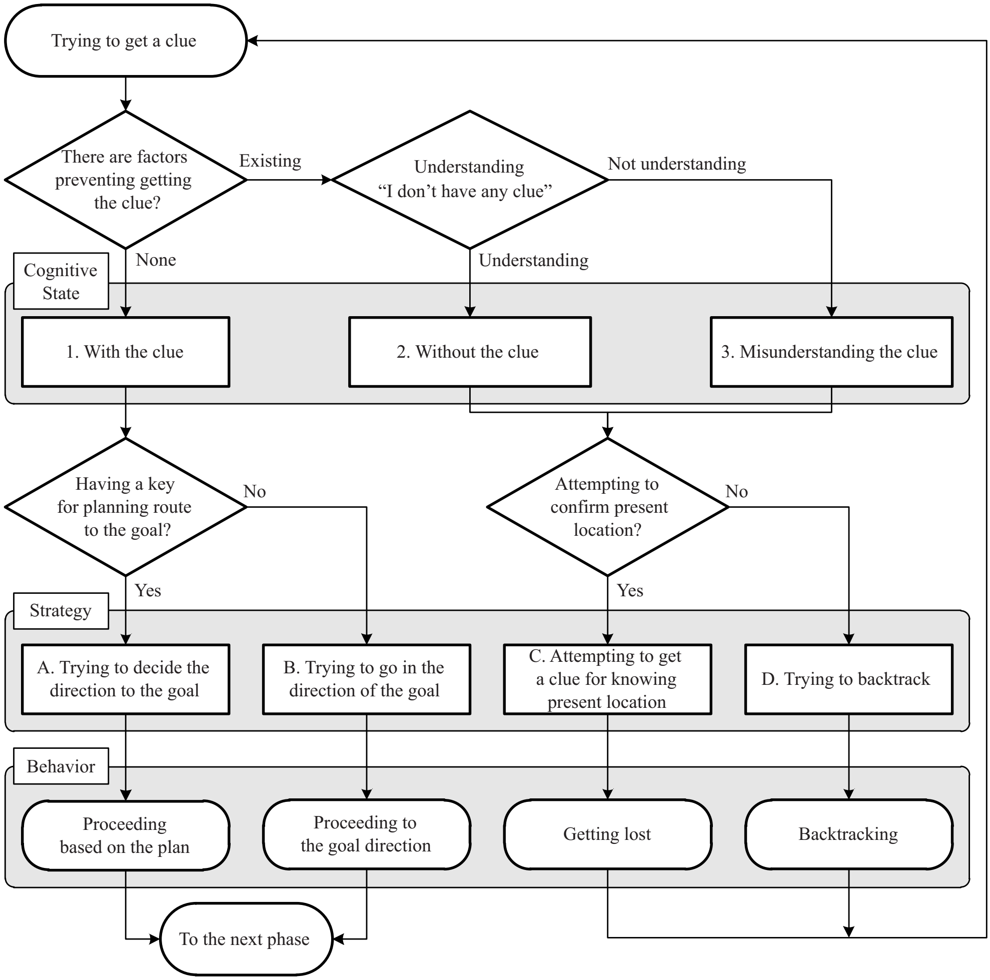


Figure 2. Cognitive process of a phase in evacuation.

Figure 2 also indicates the types of cognitive state and strategy. For example, 'With the clue' is preferable cognitive state and 'Trying to decide the direction to the goal' is better strategy than the other strategies because these cognitive state and strategy lead to the above preferable behaviors. Therefore, we assumed that the types of cognitive state and strategy represent the quality of evacuee's cognition in the phase. In other words, the types of cognitive state and strategy could be a measure of the quality of cognition.

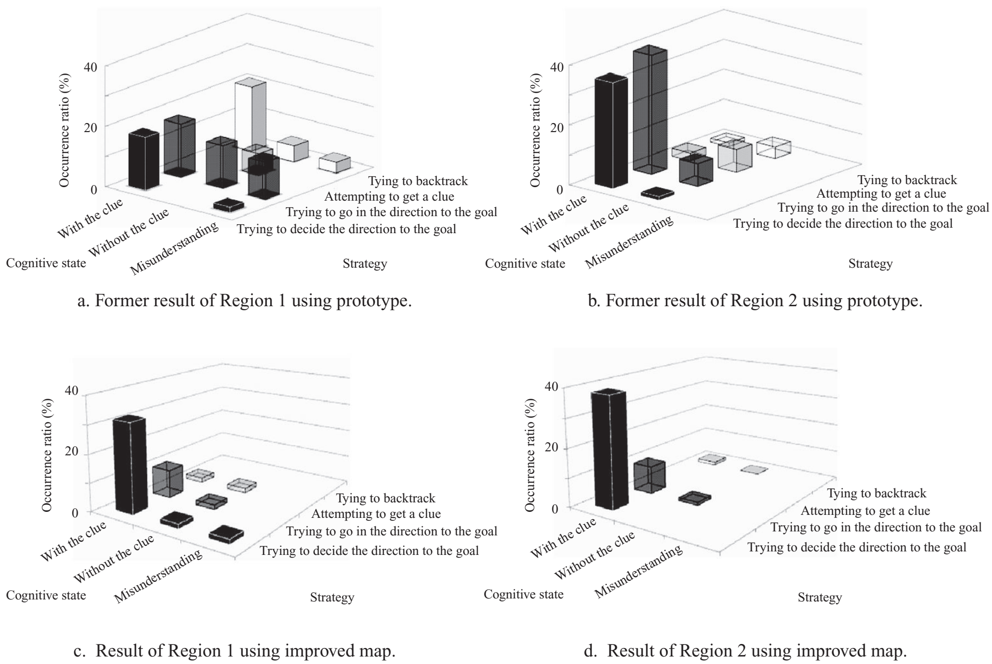
3.5 *Summary*

From the participants' think aloud protocol and behavior observed, the evacuee's cognitive process in a phase was represented as a model. Further, we assumed that the types of cognitive state and strategy could be a metrics. Therefore, the concept of cognitive process could be applicable to evaluation of the evacuee's cognition. In order to clarify these points, we executed further research.

4 VERIFICATION OF COGNITIVE MODEL

4.1 *Experiment*

In order to clarify the validity of the cognitive model, we evaluated the newly designed map on PDA using two measures, and compared with the prototype. The measures are participants' performance and quality of cognition. The participants' performance measured using the distance they walk and the quality of cognition measured using the occurrence of each type of cognitive state and strategy.



a. Former result of Region 1 using prototype.

b. Former result of Region 2 using prototype.

c. Result of Region 1 using improved map.

d. Result of Region 2 using improved map.

Figure 3. Comparison among the results of occurrence ratio of cognitive state and strategy.

The newly designed map we used in this verification experiment is based on the knowledge from above experimental results. For example, the new map included the name of land mark and the name of building located around corner. In addition, the new map was color-coded in order to let the evacuee know the narrow pathway. Further, the triangle dotted lines were drawn on the way toward the direction to the goal for providing a sense of distance.

The participants were the other 10 students and divided into 2 groups including 5 of each. The respective groups were tried to evacuation simulation in each region using the improved map on the PDA (NTT DoCoMo T-01 A). However, the respective participants were tried to evacuate one by one. The other procedure of the verification experiment and instruction to the participants was as well as before.

4.2 *Performance result*

Experimental results showed that 2 of 5 participants were able to walk the shortest path of Region 1. Further, the other 3 participants of Region 1 could walk short distance around 400 meters.

In addition, 4 of 5 participants of Region 2 could walk through the shortest path and the average distance was shorter than before. Therefore, we decided the improved map on PDA was more valid than before from the performance data.

4.3 *Quality of cognition*

The quality of cognition was decided using the cognitive model as shown in Figure 2. To say more concretely, the participant's cognitive state and strategy were judged using the model and the occurrence ratio of cognitive state's and strategy's type was calculated.

Figure 3a–3d shows the comparison of occurrence ratio. Figure 3a, 3b shows the results of former evacuation simulation using the map of prototype. Figure 3a indicates that undesirable cognitive state and strategy were observed in many phases. Concretely, the ratio such as 'Misunderstanding' and 'Trying to backtrack' were relatively high in the case of Region 1 (see Figure 3a).

Meanwhile, Figure 3c, 3d shows that the results using improved map lead to preferable types of cognitive state and strategy. By comparing Figure 3a with Figure 3c, we know the occurrence ratio of undesirable type of cognitive state and strategy is lower in the case of using improved map (see Figure 3c). As explained above, we concluded that the improved map made an effect on not only for evacuee's performance but also their cognition.

Above results shows that the participant's quality of cognition is in accordance with the performance data; therefore we are able to decide that the occurrence ratio of cognitive state's and strategy's type reflected the quality of cognitive process of evacuee. Thus, we found that the occurrence ratio of cognitive state's and strategy's type could be a metrics for evaluation of evacuee's mental state, environment, and user experience of the evacuation route map on PDA.

4.4 *Summary*

From the viewpoint of cognitive state and strategy, the quality of cognition was estimated. As the results, the evacuee's quality of cognitive process was evaluated and the validity of evaluation was clarified.

5 CONCLUSION

From the evacuee's view, the evacuee's cognitive process was discussed through experiments, and the cognitive model of evacuee was proposed. The validity of the model was clarified experimentally. Further, from the result of verification experiment, it was revealed that the information included in the improved map on PDA was the required information for the evacuee.

REFERENCES

Kobayashi, D., Murata, H. & Yamamoto, S. 2007. Study on Training Method for Preventing Plant Operator's Erroneous Thinking in Malfunction Diagnosis, In D. Harris (ed.), *Enginering Psychology and Cognitive Ergonomics*; *Proc. HCII2007,* Las Vegas, 2007. Springer.

Takahashi, Y., Kobayashi, D. & Yamamoto, S.A. 2009. Development of Information System for Disaster Victims with Autonomous Wireless Network, *Proc. HCII2009,* San Diego, 2009. Springer.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A research on high school students' capacity of information processing and information searching

Chuang Li-Wen
Graduate School of Management, National Kaohsiung First University of Science and Technology, Taiwan

Chu Wei-Cheng
Fashion Design Department of Shu-Te University, Taiwan

Chiu Shu-Ping
Graduate School of Design, National Yunlin University of Science and Technology, Taiwan

ABSTRACT: The capacity of Information Processing and information searching are the basic ability of the modern citizen in the 21st century. Nowadays, the rapid development of information technology has motivated developed countries to establish schools and organizations that are good at using information. In terms of present enterprises, they also begin to rethink about their working process and how to integrate all their staffs' knowledge and capacity through the assistance of information technology. Only people who have enough the ability of Information Processing and information searching are able to access, evaluate, organize and integrate information effectively with the appropriate methods, while following the exigencies of information ethics. Thereby, they can think critically, conduct self-oriented learning, and increase their own ability to solve their information needs. This research mainly focuses on high school students and does research on the features of students' various background, information environment, the ability of Information Processing and information searching and their relationship. Firstly, this thesis begins with the literature reviews of information literacy related theories and researches. The survey was done using questionnaires with validity and reliability on 1150 high school students. The survey was extracted in the way of stratified random sampling, and was analyzed via descriptive statistics, percentage, mean, standard deviation, chi-square test, t test, and one-way ANOVA. The research produced the following findings that students' backgrounds exhibit significant difference in information environment, the capacity of information processing and information searching. (1) Students whose parents have higher educational degree are significantly different in information processing capacity. (2) Types of public and private schools as well as districts of schools cause significant differences in "information searching capacity". (3) The information environment for the students is well designed and organized. Those students with better information environment are more competent in the capacity of Information Processing and information searching.

Keywords: information literacy, information processing, information searching, capacity

1 INSTRUCTIONS

Today, an age of information technology, the society, organizations and individuals develop their respective features more and more depending on computer and Internet technology. As Rogerson & Bynum say, "The revolution of information and technology is gradually changing not only the locations and ways of our working and learning, but also the modes of shopping, leisure, diet, election, medical service, war, making friends, etc. The wave of information revolution has a great, potential effect on human's values. Therefore, it is shortsighted and wrong to simply consider revolution of information and technology as an issue of 'technology' (Rogerson & Bynum 1996).

Living in this 21st century of informationization, we have difficulties not in a lack of information or an anxious problem of information explosion, but in how to confront information, cognize the value of information, acquire the skills of utilizing information as well as form healthy and correct attitudes such as converting and integrating useful information to generate knowledge in need (i.e., what is the information in need and how to acquire it).

This study is theoretically based on summarizing, discussing and analyzing literatures about information processing, information searching and information literacy. It also aims at analyzing statistics data through questionnaire scale, so as to know senior high school students' present information processing and searching capacity, demands as well as other relevant factors, and offer a reference for chief educational offices and schools to teach or implement software and hardware construction of information education in the future. In this way, it hopes to enhance senior high school students' information processing capacity, information searching capacity and learning efficacy. To achieve the goals of this study, the following is summarized: (1) Discuss and generalize differences in senior high school students' capacity of information processing and searching. (2) Discuss and generalize whether senior high school students have significant differences in information processing and searching capacity due to their various backgrounds. (3) Discuss and generalize whether senior high school students have significant differences in information processing and searching capacity due to their various information environment.

2 THE CAPACITY OF INFORMATION PROCESSING AND INFORMATION SEARCHING

2.1 *The capacity of information processing*

During the last several decades with the globalization of business and eLearning, enhanced market competition, and the need to respond quickly to customer requirements. These systems facilitate the sharing and flow of information, allowing members to access large volumes of information within a short period of time. Collaboration technology can boost the productivity of virtual groups provided that the members can effectively process the acquired information (Paul & Nazareth 2010).

The productivity of individuals and groups depends strongly on their information-processing characteristics (Driver & Streufert 1969). Although information processing is a well established area of research in cognitive science and psychology. In cognitive psychology, there are multiple views about how humans process information. The generalists adopt a simple view, and posit that all humans are more or less similar in processing information (Miller 1956, Streufert. S. et al. 1965). While processing information, individuals or groups sort the input information into different dimensions or points of view, which is referred to as differentiation. Differentiation represents the perception of different dimensions within a stimulus domain, and the adoption of different perspectives when considering the domain (Smith 1992). Groups responding to emergency conditions often experience overload, primarily due to threat rigidity (Turoff et al. 2004), wherein planned responses are effective for predefined threats, but novel threats prove unsettling (Plotnick et al. 2009). Likewise, increases in irrelevant information and loss of cognitive attention are posited to contribute to increased information overload (Plotnick et al. 2009).

2.2 *The capacity of information searching*

Students use a very limited range of information resources to meet their learning needs in the school or in the house. Recent studies demonstrate that undergraduate students are dependent on the internet with the vast majority of students (73–91%) reporting they use search engines as their primary search method (Callinan 2005, Griffiths & Brophy 2005, McHarg et al. 2006, Urquhart et al. 2005). Brophy and Bawden (2005) investigated whether searching the internet using Google™ could replace database searching for typical undergraduate student research queries. Their study examined in detail the information retrieved from Google and database searches in terms of both quality of information retrieved and coverage of the topic area. Brophy and Bawden's (2005) study also showed that database searching yielded a higher proportion of good quality documents (84%) than Google™ (52%). Evaluating the quality of rerieved documents is therefore an important component of successful information searching when using the internet.

Defining learning as the development of new knowledge, Marchionini (2006) elaborates on the learning elements inherent in information searching by describing searching in order to learn as increasingly viable in a content rich, online environment.

Marchionini (2006) conjectures that searching can be a learning process requiring multiple iterations and cognitive evaluation of retrieved results. The researcher specifically employs terminology from Bloom's categorization to describe attributes of learning while searching for information. Tsai and Tsai (2003) explored the relationship among students' information searching strategies in Web-based science learning activities and the influence of students' Internet self-efficacy. The researchers reported that students with high self-efficacy employed better searching strategies and learned better relative to those students with low self-efficacy.

Charles McClure (1994) expressed this in an early model of information literacy which relates

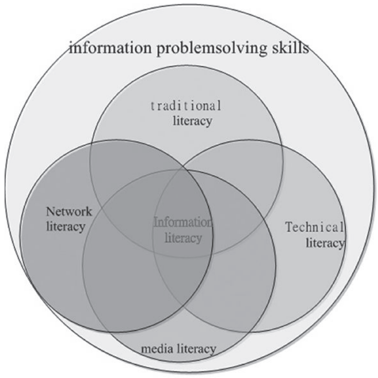


Figure 1. Charles McClure (1994) model of information literacy.

information literacy to other literacies: (see Figure 1) (1) the traditional literacy —an individual must be able to read and write; (2) the technical literacy—the person must be technically literate, e.g., be able to operate computer, telecommunication, and related information technologies; (3) media literacy—be able to use post—print media, e.g., Electronic media, to interpretation, evaluation, analysis, production, dissemination of information capacity; (4) network literacy—defined as the ability to identify, access, and use electronic information from the information network. All of these types of literacies can be cast in the context of information problem solving skills.

Above scholars' opinions have various emphases. The capacity of information processing can be generalized as facing various problems to be resolved, properly dealing with and defining the nature and category of these problems at the same time, judging the urgency and importance of problems processing, and confirming the information scope of problems solving. The capacity of information searching includes finding out the value, benefit and significance of information through information cognizing, searching, choosing, integrating and applying during the problem solving process, as well as evaluating the appropriateness and information ethics compliance of problem solving results.

3 RESEARCH METHOD

According to above research motivations and purposes, this paper is generated through referring to and analyzing relevant research literatures. It aims at discussing and knowing the influences of individual background variable and information environment variable on current senior high school students' capacity of information processing and information searching, so as to provide a reference for relevant educational units to plan information education activities and compile teaching materials about information capacity.

Firstly, it identifies the research topics and subjects, i.e., collecting, studying and analyzing relevant literatures as the theoretical basis of this research. Then, it establishes research purposes, frameworks and hypotheses. Thirdly, the design and investigation of quantitative research questionnaire is carried out, questionnaires are sent to research subjects, and analysis is implemented by SPSS 12 software after questionnaire recycling. Finally, it verifies research purposes and hypotheses, generalizes research results, and proposes conclusions and suggestions.

For analyzing and summarizing the research goals and literature discussion, this study investigates the following aspects among senior high school students (as shown in Figure 2, research model):

(1) Individual background: gender, average family revenue per month, educational degree of parents, parents' occupations, whether the parents use Internet, parents' attitudes towards your being on Internet, public or private school, the district of your school. (2) Information environment situations: individual situation of possessing computer and network, daily situation of using various tool ware and 3C multimedia devices, the time of going to library and using library resources per week, situation of utilizing computer and network during learning, study (activities) hold by the school, information study (activities) inside and outside school. (3) Classified into two parts: A: information processing capacity, B: information searching capacity.

This study analyzes above questionnaire survey data by means of descriptive statistics, chi square test, independent samples t test, one-way ANOVA, etc, so as to find out whether these aspects affect students' capacity of information processing and searching. In terms of above research purposes and frame, hypotheses are formulated and relevant

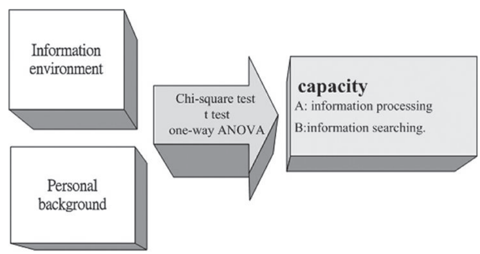


Figure 2. Research model.

tests are executed as follows: (1) Whether there are significant differences in senior high school students' capacity of information processing and searching. (2) Whether students with various backgrounds are significantly different in capacity of information processing and searching. (3) Whether students are significantly different in capacity of information processing and searching because of various information environments.

After pre-tests, item analysis, factor analysis, reliability analysis and validity analysis are carried out, so as to achieve the feasibility and correctness of test results after verification. The subjects of this study are students from various public and private senior high schools of Kaohsiung County, who constitute the 1150 subjects. Among the 1150 questionnaires sent out, 1117 ones are recycled with a response rate of 7.13%. After deleting incomplete and inappropriate answers to the questionnaires, there are 1099 valid ones and 18 invalid ones, with a availability ratio of 98.38%. Besides a reliability analysis of pretest results, this study also carries out a reliability analysis of formal questionnaire test results by "cronbach α" coefficient, so as to further verify the validity and reliability of formal *questionnaires*. The cronbach α coefficient of total scale is .923. The lowest cronbach α coefficient of sub-scale information processing capacity is .820, and that of information searching capacity is .786. Results show that formal questionnaires have good stability and internal consistency.

4 RESULT AND DISCUSSION

This part analyzes and discusses the data from questionnaire survey according to research goals and hypotheses, and then presents the results of actual questionnaire survey. This chapter is divided into 4 sections: section 4.1 analyzes current situations of students' information processing and searching capacity; section 4.2 analyzes students' cognitive differences in information processing and searching capacity when they have various backgrounds; section 4.3 analyzes students' cognitive differences in information processing and searching capacity when they are in various information environments; section 4.4 is limitations. The content is respectively presented as follows:

4.1 *Current situations of students' information processing and searching capacity*

Current situations of students' information ability include two aspects namely information processing and information searching capacity. This study classifies students' information ability into these two aspects. Among these aspects, the mean of "information searching capacity" is the highest, with average score of each question being 3.86. Then is the aspect of "information processing capacity", with average score of each question being 3.54.

After analyzing the investigation results from senior high school students' "A information processing capacity" and sequencing them in accordance with average score, the first two items are "A3 I can use word processing software, as well as file, store, utilize and share various types of information" (with an average of 3.72) and "A2 I can use presentation software and multimedia devices to integrate information communications, introduce information and share with others" (with an average of 3.57). The results show that these students have some ability of A3 and A2 aspects during their daily learning and dealing with personal affairs. Concerning "A4 I can develop the methods and procedures of problem solving in terms of features and limitations of various information media", the average 3.33 is lower than the mean score of overall questions, and this aspect can be a reference for courses, studies, contests or activities relevant to enhancing senior high school students' information processing and searching capacity.

After analyzing the investigation results from the aspect of "B information searching capacity" and sequencing them in accordance with average score, the top two items are "B3 I know what information resources, including Internet discussion groups, forums, search engines, etc, benefit my own problem solving" (with an average of 4.00) and "B2 I can participate in Internet discussion groups and forums and pick out the information needed for problem solving" (with an average of 3.85). Results show that most students can achieve the aspects of B3 and B2. Although the average of some question items such as "B1 know where to find out the search methods and information when in need of information" (with an average of 3.77) is above the middle degree, by contrast, students can still do a lot better in knowing where to find out the search methods when in need of information, and cultivating information ability.

4.2 *Analyzing students' cognitive differences in information processing and searching capacity when they have various backgrounds*

This section aims to find out whether senior high school students with various backgrounds are significantly different in information processing and searching capacity. It is simply divided into two aspects: A: information processing capacity, B: information searching capacity.After collecting data from questionnaire survey, data is processed by means of t test and one-way ANOVA. If the test results reach significant level, The Scheffe's test will be carried out to identify differences between groups. The verification of hypothesis 2 is described as follows:

Hypothesis 2: Senior high school students have significant differences in information processing and searching capacity due to their various backgrounds.

2-1 Students with different genders are significantly different in capacity of information processing and searching.

2-2 Students with diverse average monthly family revenue are significantly different in capacity of information processing and searching.

2-3 Students have significant differences in information processing and searching capacity if the highest educational degrees of their parents are different.

2-4 Students are significantly different in information processing and searching capacity if their parents' occupations are dissimilar.

2-5 Students from different public or private schools have significant differences in information processing and searching capacity.

2-6 Students from diverse schools at different districts are significantly different in capacity of information processing and searching.

The analysis of investigation results is presented as follows: Students with different genders are not significantly different in capacity of information processing and searching. Through analyzing the average of male and female students, it is found that senior high school students with different genders don't have gender differences in "information processing capacity" and "information searching capacity".

Students with different average monthly family revenue are significantly different in "capacity of information processing". The highest average score exists in students with average revenue of 50,000 to 99,999. About "information searching capacity", no difference results from diverse average monthly family revenue. Speaking from a single aspect, students with different average monthly family revenues are significant different in "information processing capacity".

Students' parents with highest educational degree of junior college are better in "information processing capacity" than those with highest degree of primary school and junior middle school, and better in "information searching capacity" than those with a highest degree of junior or senior vocational school as well uneducated parents. Obviously, students have significant differences in information processing and searching capacity if the highest educational degrees of their parents are different.

Students have significant differences in information processing and searching capacity if their parents' occupations are different. Students whose parents engage in business and service industry have higher average than other students. Obviously, students are significantly different in information processing and searching capacity if their parents' occupations are dissimilar.

Students from different public or private schools have significant differences in "information searching capacity", and are not significantly different in "information processing capacity". Speaking from the average, students from public schools are better in "information searching capacity" than those from private schools. Above results may be caused by different information device software and hardware, teachers, research subjects, students' degrees or regions of public and private schools.

Students from diverse schools at different districts are significantly different in "capacity of information searching" and not remarkably different in "information processing capacity". Students from Feng-Shan district have higher average in information searching capacity than those from Chi-Shan district; and students from Gang-Shan district are higher than those from Chi-Shan district. This shows that students from schools at different areas have significant differences in relevant aspects. Above results may be caused by different regions, urban-rural gaps or urban-rural digital divide.

Mostly, senior high school students are significantly different in information processing and searching capacity according to diversified genders, average family revenues per month, parents' highest educational degree, parents' occupations, types of public and private schools, districts of schools. Above results confirm the hypothesis 2 that senior high school students with various backgrounds are significantly different in information processing and searching capacity.

4.3 *Analyzing students' cognitive differences in information processing and searching capacity when they are in various information environments*

This section aims to find out whether senior high school students in various information environments are significantly different in information processing and searching capacity. It is simply divided into two aspects: A: information processing capacity, B: information searching capacity.

After collecting data from questionnaire survey, data is processed by means of t test and one-way ANOVA. If the test results reach significant level, The Scheffe's test will be carried out to identify differences between groups. The verification of hypothesis 3 is presented as follows:

Hypothesis 3: Senior high school students have significant differences in information processing and searching capacity because of their various information environments.

3-1 Students have significant differences in information processing and searching capacity dependent on whether their homes have personal computers (desktops or laptops) connected to the network.

3-2 Students with diverse average daily time of using computers are significantly different in information processing and searching capacity.

3-3 Students with diverse average daily time of using Internet have significant differences in information processing and searching capacity.

3-4 Students with diverse average daily time of using tool software (e.g. Word, Excel, PowerPoint, and Photoshop) are significantly different in information processing and searching capacity.

3-5 Students with diverse average daily time of using 3C multimedia devices (such as videos, projectors, MP3 Walkman) have significant differences in information processing and searching capacity.

3-6 Students have significant differences in information processing and searching capacity because of different average weekly time of going to library or utilizing library resources.

3-7 Students are significantly different in information processing and searching capacity due to different situations of using computers during studying at school.

3-8 Students have significant differences in information processing and searching capacity depending on whether their schools provide dedicated computers for schoolwork.

3-9 Students have significant differences in information processing and searching capacity depending on whether their schools offer computer-relevant training courses.

3-10 Students are significantly different in information processing and searching capacity depending on whether they take part in computer-relevant training courses after school.

Analysis results confirm the hypothesis 3-1 of "students have significant differences in information processing and searching capacity dependent on whether their homes have personal computers (desktops or laptops) connected to the network". Whether students' families have personal computers (desktops or laptops) connected to the network results in significant differences in aspects of "information processing capacity" and "information searching capacity". Speaking from the average, significant differences exist in both aspects, and students whose family owns computers connected to the network have higher average than those don't possess.

Analysis results partly accept hypothesis 3-2 of "students with diverse average daily time of using computers are significantly different in information processing and searching capacity." Students who spend different time in using computers are significantly different in "information searching capacity". Speaking from the average, significant difference exists in students' "information searching capacity", and students who averagely spend 2 hours and 2–4 hours in using computers have much better information processing and searching capacity than those who have no time of using computers. However, the ability doesn't become stronger as the time become longer.

Analysis results partly accept hypothesis 3-3 of "students with diverse average daily time of using Internet have significant differences in information processing and searching capacity". This factor causes significant differences in both "information processing capacity" and "information searching capacity". Analyzing from the viewpoint of average score, significant differences exist in both aspects, and students who averagely spend 2 hours and 2–4 hours in using Internet have much better information processing and searching capacity than those who have no time of using Internet. However, the ability doesn't become stronger as the time become longer.

Analysis results accept hypothesis 3-4 of "students with diverse average daily time of using tool software (e.g. Word, Excel, PowerPoint, and Photoshop) have significant differences in information processing and searching capacity". This factor causes significant differences in both "information processing capacity" and "information searching capacity". Speaking from the average, significant differences exist in both aspects, and students who averagely spend less than 2 hours in using Internet have much better information processing and searching capacity than those who have no time of using Internet. However, the ability doesn't become stronger as the time become longer.

Above results accept hypothesis 3-5 of "students with diverse average daily time of using 3C multimedia devices (such as videos, projectors, MP3 Walkman) have significant differences in information processing and searching capacity". This factor brings about significant differences in both "information processing capacity" and "information searching capacity". Speaking from the average, significant differences exist in "information processing capacity" and "information searching capacity". Students who averagely spend less than 2 hours, 2 to 4 hours or over 5 hours in using 3C multimedia devices have much better information processing and searching capacity than those who have no time of using 3C multimedia devices.

Analysis results partly accept hypothesis 3-6 of "students have significant differences in information processing and searching capacity because of different average weekly time of going to library or utilizing library resources". Students with different average weekly time of going to library or utilizing library resources present significant differences in "information processing capacity" and no significant differences in "information searching capacity". Speaking from the average, students with different average weekly time of going to library or utilizing library resources are significantly different in "information processing capacity". Students who averagely spend less than 2 days or 2 to 4 days in going to library to utilizing library resources have much better information processing and searching capacity than those who don't go to library or utilize library resources at all.

Analysis results accept hypothesis 3-7 of "students are significantly different in information processing and searching capacity due to different situations of using computers during studying at school." Students, who are different in using computers during at school, present significant differences in information processing and searching capacity. Analyzing from the angle of the average, significant differences exist in both aspects of "information processing capacity" and "information searching capacity". Students who frequently use computers at school have much better information processing and searching capacity than those use computers occasionally or those don't use at all.

Analysis results reject hypothesis 3-8 of "students have significant differences in information processing and searching capacity depending on whether their schools provide dedicated computers for schoolwork". Significant differences don't exist in "information processing capacity" and "information searching capacity". This may result from the following reasons: most students have computers at home, dedicated computers for schoolwork are seldom used inside school, and time limitations exist.

Analysis results partly accept hypothesis 3-9 of "students have significant differences in information processing and searching capacity depending on whether their schools offer computer-relevant training courses". Whether the schools offer computer-relevant training courses results in significant differences in "information searching capacity" but not so in "information processing capacity". Speaking from the average, students from schools that frequently open computer-related training courses have better information searching and processing capacity than those from schools that occasionally or never establish these courses.

Analysis results reject hypothesis 3-10 of "students are significantly different in information processing and searching capacity depending on whether they take part in computer-relevant training courses after school". This factor doesn't cause significant differences in both "information processing capacity" and "information searching capacity". It is shown that students who ever take computer-related training courses after school (17.2%) have similar information processing and searching capacity with those who never take after-class courses (82.6%). The reason of no difference may be that the proportion of students who take after-school courses is low (17.2%), or students have heavy schoolwork at school and spend time in dealing with personal affairs. This can be a reference for arranging courses, studies, contests or activities relevant to enhancing senior high school students' information processing and searching capacity.

Presently, senior high school students' capacity of information processing and searching varies significantly with whether they have computers connected to network at home, average time of using computer everyday, average time of using Internet everyday, average time of using tool software everyday, average daily time of using 3C multimedia devices, average time of going to library or utilizing library resources each week, the situation of using computers at school, whether the school establishes computer-related training schools, etc. However, significant differences are not caused by whether the school provides dedicated computer for schoolwork or whether the student takes computer-related training courses after school. Students who spend long time in using information devices or frequently use information devices have better capacity of information processing and searching. Above results confirm hypothesis 3 of "senior high school students have significant differences in information processing and searching capacity because of their various information environments".

4.4 *limitations*

Confined by factors including economy, manpower, material resources, time, etc, this study describes two kinds of constraints as follows, namely limitations of sampling scope and research subjects, and limitations of research tools:

(1) Limitations of sampling scope and research subjects

This study samples among various public and private senior high schools, and chooses students from grades 1 to 3 of these schools as the subjects. It is difficult to know other types of research subjects, and thus excessive inference is not proper.

(2) Limitations of research tools

This study mainly uses questionnaire survey to collect subjects' data. As the questionnaire scales answered by subjects are all self-report presentations and subjects fill out questionnaires independently, it is unable to determine subjects' context, psychology and reality. Therefore, this study can only make unilateral hypothesis: all the subjects should answer faithfully without being affected by social and self expectations. The research tools are constricted by whether the subjects' answers are correct, and thus, it should be prudent to explain and apply the research results.

5 CONCLUSIONS

Three conclusions are generalized and described as follows:

1. Students whose parents have higher educational degree are significantly different in information processing capacity. Research results show that students' backgrounds bring about significant differences in relevant aspects. Senior high school students whose parents have the highest educational degree of junior college have better "capacity of information processing" than those whose parents have educational background of primary school or junior high school, and are also better in other related aspects than students whose parents are uneducated people or senior high school graduates.
2. Types of public and private schools as well as districts of schools cause significant differences in "information searching capacity". Results show that students from different public or private schools have significant differences in information searching capacity. Students of public schools have better information searching capacity than those of private schools. Through analyzing the districts of various schools, it is found that students from Feng-Shan area have better information searching capacity than those from Chi-Shan area; students from Gang-Shan area have better information searching capacity than those from Chi-Shan area. Obviously, the districts of students' schools indeed cause significant differences in relevant aspects. Above analysis results may result from diversified information device software and hardware, teachers' qualifications, research subjects, degrees, regions or urban-rural gaps. This provides an information literacy reference for chief educational offices to shorten the gaps between public and private schools or urban and rural areas.
3. The information environment for the students is well designed and organized. Those students with better information environment are more competent in the capacity of Information Processing and information searching. The popularizing rate of subjects who have personal computers connected to Internet at home is quite high (85.9%), which means that computer and network play a role in students' daily life, and most students use computer, Internet and tool software everyday. Secondly, most students use computer to deal with their schoolwork and other relevant affaires everyday. Computer-relevant training courses are established inside the school occasionally or frequently, and most students participate in these courses and enjoy this. Among a few students who take relevant courses after school, only 17.2% are willing to strengthen computer-related conceptions or skills, which means a situation to be improved. Moreover, chief educational offices and schools should actively budget for various informationalization facilities year by year, buy more computer-related software and hardware devices, and improve the collocation of proper teachers. Overall, the information environment situation of school still complies with the standards.

REFERENCES

Brophy, J. & Bawden, D. 2005. *Is Google enough? Comparison of an internet search engine with academic library resources.* London: Aslib.

Callinan, J. 2005. Information-seeking behaviour of undergraduate biology students. *Library Review* 54: 86–99.

Driver, M. & Streufert, S. 1969. Integrative complexity: An approach to individuals and groups as information-processing systems. *Administrative Science Quarterly* 14: 272–285.

Griffiths, J. & Brophy, P. 2005. Student searching behavior and the web: use of academic resources and Google. *Library Trends* 53: 539–554.

Marchionini, G. 2006. Exploratory search: from finding to understanding. *Communications of the ACM* 49: 41–47.

McClure, C.R. 1994. Network literacy: A role for libraries? *Informationtechnology and Libraries* 13: 115–125.

McHarg, J., Goding, L., Caldarone, E., de Bere, S. & McLachlan, J. 2006. Availability of a virtual learning environment does not compensate for the lack of a physical facility. *Medical Teacher* 28: 258–263.

Miller, G.A. 1956. The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81–97.

Paul, S. & Nazareth, D. 2010. Input information complexity, perceived time pressure, and information processing in GSS-based work groups: An experimental investigation using a decision schema to alleviate information overload conditions. *Decision Support Systems* 49: 31–40.

Plotnick, L., Turoff, M. & Van Den Eede, G. 2009. Reexamining threat rigidity: Implications for design. In *42st Hawaii International Conference on Systems Science (HICSS-42)*. Waikoloa, Hawaii.

Rogerson, S. & Bynum, T.W. 1996. Information ethics: the second generation. *UK Academy for Information Systems Conference*, UK.

Smith, C. 1992. *Motivation and personality: Handbook of thematic content analysis.* Cambridge: Cambridge University Press.

Streufert, S. & Suedfeld P.D. 1965. Concept structure, information search and information utilization. *Journal of Personality and Social Psychology* 2(5): 736–740.

Tsai, M. & Tsai, C. 2003. Information searching strategies in web-based science learning: the role of Internet self-efficacy. *Innovations in Education and Teaching International* 40: 43–50.

Turoff, M., Chumer, M., Van de Walle, B. & Yao, X. 2004. The design of emergency response management information systems (ERMIS). *Journal of Information Technology Theory and Application* 5: 1–35.

Urquhart, C., Thomas, R., Spink, S., Fenton, R., Yeoman, A, et al. 2005. Student use of electronic information services in further education. *International Journal of Information Management* 25: 347–362.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Effects of short-term memory load and verbal transformation on working performance and EEG response

Tam Chan & Hsin-Yi Hsu
Institute of Industrial Engineering and Enterprise Information, Tunghai University, Taichung, Taiwan

ABSTRACT: In the modern time, people are required to deal with numerous messages in the working time very often. The message processing usually includes memorization and transformation in order to simplify the present of information. Thereby, design of message presentation has been a critical factor of improving working performance. Over the experiment, we acquired and recorded the data of working performance and subjects' brain wave during the task operating. This research is going to discuss the effects of short-term memory load and verbal transformation on working performance and EEG response. Working performance in this experiment consisted of finish time and error number. EEG response included θ rhythm and α signal. The results were analyzed with SPSS and showed that: With increased memory load, both finish time and error number significantly increased. Verbal transformation significantly affected finish time and error number. ETR caused the longest finish time, and TR caused the largest error number. Frontal-θ rhythm increased in strength with the difficulties of memorization and transformation. Occipital-α signal decreased in strength while the amount of characters remembered increased, and showed the strongest power for TR. The left hemisphere was more functional than the right side during verbal memorization and transformation.

Keywords: brain wave, EEG, working performance, memory, transformation

1 INTRODUCTION

As the science and technology come along, people increasingly rely on computers to acquire information and knowledge instead of papers. Thus, the memory load and the amount of information transformed have been a consequence of engineering in the present day. The requirement of precise and accurate memorization is getting more in our daily life since the information passes more and more speedily. Furthermore, most of the time, lots kinds or types of information have to be compressed into a simple sensitive unit from which is needed to transform. However, present studies had mainly concerned about the amount of sensitive units which has to be memorized, but lack of the synchronous transformation of data. Thereby, the volume of mental memorization and transformation has been increasingly demanded for better working performance. How to improve the working performance of computer by enhancing the recognition of text or picture on the screen has become more and more important today (Cushman, 1986).

In addition to human machine interface, working memory also plays a role in operation performance. Medically, brain is called the central nervous system, which is the most important organ in human body and responses to memory and cognition. The activities of neuronal cells could be detected and recorded as EEG signal. Since EEG signal is regarded as the presentation of activities of cortices, we diagnosed if the EEG response differs according to different task difficulty, and attempted to define the EEG signals.

2 LITERATURE REVIEW

2.1 *Memory*

The composition of memory system includes "memory storage" and "memory process". According to the sequence of stimulation, "memory storage" can be divided into sensory memory, short-term memory, and long-term memory. Memory process comprises four components, which are attention, rehearsal, encode, and retrieve. Obviously, memorization plays a critical role in the progress of comprehension. Short-term memory is also called working memory (Gevins, 1998), which is caused by cognitive stimulations and sustains for a while after the stimulation finished. With finite capacity, short-term memory is about the range of 7 ± 2 sense units, which are easily misplaced within 20~30 s without rehearsal and encoding (Miller, 1956).

2.2 *Brain wave*

Recent theoretical and experiment work has focused on the role of brain oscillations in working memory and target differentiation; there has been particular interest in distribution of the potential of brain cells. The diverse functions of different regions of cortex can be indexed:

1. Frontal lobe: forehead, in charge of sport coordination, thinking determination, and problem solving.
2. Parietal lobe: top of the head, responsible for tactile sensitivity.
3. Occipital lobe: hindbrain, reflecting vision.
4. Temporal lobe: bilateral sites of the head, responsible for auditory.

Brain wave can be approximately distinguished into four waves in frequency: beta (β) wave (12~32 Hz), alpha (α) wave (8~12 Hz), theta (θ) wave (4~8 Hz), delta (δ) (0.4~4 Hz).

2.2.1 θ *band*

Referring to EEG signals, lots of researches and studies had already constructed a few general theories. A frontal midline θ rhythm increased in strength with increased operating difficulty in a memorization and recognition task. In the experiment, the frontal θ rhythm was largely insensitive to the type of information. With increased practice on the task, the frontal θ signal apparently increased in amplitude (Gevins et al., 1997).

Klimesch et al. (1997) indicated that θ oscillations reflect processes of the WMS, whereas faster-α oscillations reflect retrieval of LTMS. Moreover, an event-related increase in θ oscillation reflects increasing short-term memory. A decrease in faster-α oscillation reflects semantic memory processes. A significant increase in θ probably results in the coherence between prefrontal and posterior electrode sites during the retention interval of a working memory experiment (Sarnthein et al., 1998).

There's a study which indicated that semantic processing does not draw selectively on the capacity of working memory, which means, different linguistic processes have no direct influence on the θ oscillation (Röhm et al., 2001). Similarly, Woodman et al. (2001) showed that Visual memorization doesn't attenuate the performance of simultaneously visual search, which means, no detrimental effect on visual search when subjects had to concurrently remember a visual object.

2.2.2 α *band*

Unlike the θ rhythm's side, there're several explanations and definitions to the activities of α signal. In the research from Jensen et al. (2002), the α peak, over the posterior and bilateral central regions, systematically increased with the number of items held in working memory.

Furthermore, α signal is generally supposed to be separated into slow-α (8~10 Hz) and faster-α (10~12 Hz). The slow-α signal decreased in strength with increased task difficulty, whereas it enhanced during a period of task operating or working time. Besides, just like the same way of θ rhythm in this part, the slow-α signal were largely insensitive to the type of information. That is, slow-α signal doesn't significantly change in strength when the type of stimulus changes. On the other hand, with large memory load, faster-α signal was suppressed more when attending to spatial than verbal information. Nevertheless, both the slow-α and faster-α signals increased in amplitude with increased practice on the task (Gevins et al., 1997). Klimesch et al (1997) indicated that θ oscillations reflect processes of the WMS, whereas faster-α oscillations reflect retrieval of LTMS.

In another standpoint, Pfurtscheller et al. (1996) had indicated that α oscillation represents the level of "idling" occurs in the cortex, which means, α oscillation is inversely related to the degree of functioning in brain region. The similar opinion came out in some other studies. There is a research showed that a decrease in faster-α oscillation reflects semantic memory processes. Besides, α activity signifies that some certain areas are not being used in processing and, in fact, are being inhibited by other active brain regions (Klimesch et al., 1999).

The increase in α activity with memory load appears to be a consequence of active inhibition of α-producing brain areas. Similarly, in a driving test experiment for domestic drivers, the power differed at the α band between over- and under-driving drivers in a VR-based driving experiment (Tung & Lin, 2005). Tremblay et al. (2008) also showed that activity in the α band may reflect attentional processes while activity in the β band may be more closely related to the execution and selection process.

2.2.3 *Verbal transformation*

The early studies indicated that when verbal transformation is implemented by brain, the information expressed by the object could totally differ from what it directly looks like. Furthermore, verbal transformation requires high level of concentration and attention, and curiously, it's not affected by how much you comprehend the object. In the experiment, the activities of EEG enhanced over frontal and central lobes when the subjects were operating verbal transformation (Galbraith et al., 1997).

In the experiments of verbal transformation and spatial transformation, a study revealed that there is a larger change in volt of ERP (Event-Relation Potential) at left-frontal when the verbal transformation is operated, whereas at occipital with the spatial transformation. In addition, the processing of short-term memory seems to be presented partially at the lobe which is supposed to be in charge of LTMS (Rolke et al., 2000).

A late research showed that the activities of EEG enhanced at left-IFC (inferior frontal cortex), ACC (anterior cingulated cortex), and left-prefrontal cortex during verbal transformation. Left-IFC and ACC reflect the load, and left-prefrontal cortex represents the occurrence of verbal transformation (Kondo & Kashino, 2007).

In sum, all above-mentioned opinions are somehow divided and even conflict with each other. Even though lots of factors have been discussed, the functional modulation of brain wave during verbal transformation which is seldom considered domestically could be still a special course to be investigated.

3 METHODS AND MATERIALS

3.1 *Objective*

The objective concerned in this study is to probe into the effects of the short-term memory load and verbal transformation level on the frontal θ and occipital α oscillation. In addition, according to the above-mentioned variables, the objective could be simply categorized as follows:

1. If different short-term memory load influences the finish time and error number of the tasks.
2. If different verbal transformation level influences the finish time and error number of the tasks.
3. If different short-term memory load influences θ and α signals in strength.
4. If different verbal transformation level influences θ and α signals in strength.
5. If different short-term memory load courses different θ and α signals in strength between hemispheres.
6. If different verbal transformation level courses different θ and α signals in strength between hemispheres.

3.2 *Subjects*

54 subjects who were aged 19–23 participated in this study. These subjects were all fully informed, healthy and no problem with hands or vision.

3.3 *EEG data acquirement and recording*

Subjects wore a movement-proof electrode cap with 36 sintered Ag/AgCl electrodes to measure the electrical activities of brain. The EEG electrodes were placed according to the international 10–20 system with a unipolar reference at the right earlobe.

The impedance between EEG electrodes and skin was kept to less than 5kΩ. Data were amplified and recorded by the Scan NuAmps Express system (Compumedics Ltd., VIC, Australia).

3.4 *Experiment design*

2 independent and 3 dependent variables were included in this experiment:

1. Independent variable

- Amount of characters remembered—3, 5, and 7 (English characters, including capital and lowercase)
- Level of verbally cognitive transformation—ETR, NT, and TR:
    - ETR (extra transformation required): recognizing the same letters, regardless of neither capital nor lowercase.
    - NT (no transformation): recognizing the same capital and lowercase letters.
    - TR (transformation required): recognizing the opposite capital and lowercase letters

2. Dependent variable

- Finish time—the finish time of the task operating.
- Error number—the number of the errors occurred during the task.
- Brain wave—recording the modulation of brain wave acquired from 17 electrodes FP1, FP2, FZ, CZ, CPZ, PZ, OZ, F3, F4, F7, F8, C3, C4, O1, O2, T5, and T6.

3.5 *Surrounding & apparatus*

The distance between screen and eyes was 30 cm. The conditions are as follows:

a. Temperature: 24~26 °C
b. Humidity: 43 ~54%
c. Luminance: 270~320 lux
d. Apparatus:

1. Acer Travel Mate C100 (Pentium® III 900)
2. Electrode cap: 32 electrodes according to the International 10–20 System.
3. Neuro-Scan 4.3.3: from NeuroScan Synamps

3.6 *Experiment process*

1. After the subject was fully instructed, the sequence of 9 combinations in the experiment was randomized.
2. Task began; the characters remembered concurrently appeared on the screen for 250 ms per character, the subject was asked to memorize what characters and what types (capital or lowercase) they were.
3. After a blank interval for 500 ms, 16 characters were showed on the screen. The subject had to recognize if all the characters remembered appeared according to the appointed type.
4. The trial varied 10 times per task.

4 RESULTS

The results showed the VDT performance, which is evaluated with finish time and error number. Furthermore, the modulation of brain wave would also be shown in this chapter.

4.1 *Working performance*

We analyzed if the finish time and error number were affected by the independent variables. The results of performance are showed as Table 1:

Table 1. Working performance.

	Finish time (s)	Error number
Amount of characters remembered		
3	38.38	0.72
5	50.76	1.61
7	56.10	3.45
Verbal transformation		
ETR	56.27	1.31
NT	40.54	1.85
TR	48.29	2.59

4.1.1 *Finish time*

The result presents that the finish time increased with the amount of characters remembered (7 = 56.10 > 5 = 50.76 > 3 = 38.38), and reached a level of significance. The verbal transformation also significantly influenced the finish time. Subjects spent the most time to finish the tasks for ETR, whereas the least for NT. (ETR = 56.27 > TR = 48.29 > NT = 40.54) The ANOVA table for finish time is shown as Table 2.

4.1.2 *Error number*

According to the result, error number increased with the amount of characters remembered (7 = 3.45 > 5 = 1.61 > 3 = 0.72), and reached a level of significance. The verbal transformation also significantly influenced the error number. Subjects made the most mistakes in the tasks for TR, whereas the least for ETR. (TR = 2.59 > NT = 1.85 > ETR = 1.31) The ANOVA table for error number is shown as Table 3.

4.2 *Brain wave*

In this investigation, we adopted the EEG response, which includes θ rhythm and α signal, to evaluate brain wave. The average power of EEG signals for different hemisphere would also be shown in this section.

4.2.1 θ *band*

We analyzed the average power and amplitude of frontal-θ rhythm for different amount of characters remembered. The results are shown as Figure 1~2. θ rhythm increased in strength with the amount of characters remembered at all frontal electrodes, and showed significances at FP2. Moreover, the average powers of θ rhythm were much stronger at FP1 and FP2, particularly FP1, than at other electrodes.

On the other hand, the average power and amplitude of frontal-θ rhythm for different verbal transformation level has also been analyzed shown as Figure 3~4. The result reveals that TR elicited the strongest power of θ rhythm at all frontal

Table 2. ANOVA for finish time.

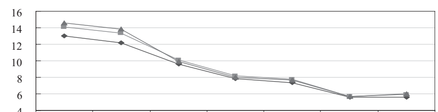
Source	ss	df	ms	F	p value
Amount of characters remembered	27572.8	2	13786.4	50.39***	<0.001
Verbal transformation	20805.5	2	10402.8	38.02***	<0.001
Interaction	4667.8	4	1167.0	4.27**	0.002
Error	135427.7	495	273.6		
Totals	188323.5	503			

*p < 0.05 **p < 0.01 ***p < 0.001.

Table 3. ANOVA for error number.

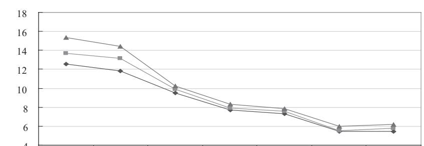
Source	ss	df	ms	F	p value
Amount of characters remembered	643.7	2	321.9	143.69***	<0.001
Verbal transformation	138.3	2	69.2	30.88***	<0.001
Interaction	73.6	4	18.4	8.21***	<0.001
Error	1108.2	495	2.2		
Totals	1962.5	503			

*p < 0.05 **p < 0.01 ***p < 0.001.



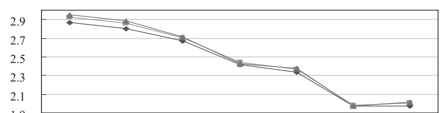
	FP1	FP2	FZ	F3	F4	F7	F8
3	12.959	12.193	9.573	7.833	7.335	5.577	5.558
5	14.083	13.364	10.085	8.137	7.719	5.694	5.877
7	14.55	13.82	9.939	7.969	7.703	5.707	5.975

Figure 1. Average power of frontal-θ rhythm for different memory load.



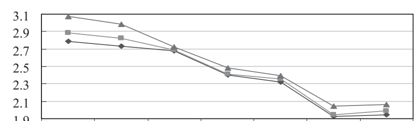
	FP1	FP2	FZ	F3	F4	F7	F8
ETR	12.534	11.814	9.507	7.729	7.287	5.434	5.456
NT	13.718	13.145	9.876	7.888	7.592	5.553	5.77
TR	15.339	14.418	10.214	8.322	7.878	5.99	6.183

Figure 3. Average power of frontal-θ rhythm for different verbal transformation level.



	FP1	FP2	Fz	F3	F4	F7	F8
3	2.866	2.799	2.671	2.419	2.337	1.97	1.976
5	2.928	2.857	2.706	2.445	2.365	1.983	2.007
7	2.949	2.886	2.709	2.427	2.373	1.97	2.014

Figure 2. Average amplitude of frontal-θ rhythm for different memory load.



	FP1	FP2	FZ	F3	F4	F7	F8
ETR	2.791	2.731	2.676	2.403	2.325	1.629	1.947
NT	2.883	2.826	2.685	2.409	2.357	1.948	1.99
TR	3.069	2.986	2.725	2.479	2.393	2.406	2.059

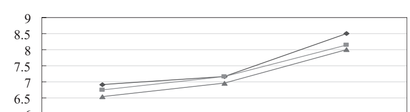
Figure 4. Average amplitude of frontal-θ rhythm for different verbal transformation level.

electrodes, whereas ETR elicited the weakest. The difference in power between verbal transformation levels reached a level of significance at FP1, FP2, F8, and marginally F7. In addition, all verbal transformation levels elicited much stronger power of θ rhythm at FP1 and FP2 than at other electrodes.

4.2.2 α *band*

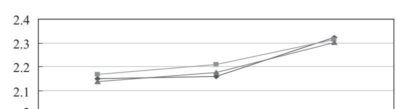
The average power and amplitude of occipital-α signal for different amount of characters remembered were analyzed shown as Figure 5~6. According the result, the average power of α signal increased with the amount of characters remembered at occipital electrodes. In addition, all amounts of characters remembered elicited much stronger power of α signal at O2 than at other occipital electrodes.

The average power and amplitude of occipital-α signal for different verbal transformation level (Figure 7~8) reveals that TR elicited the strongest



	OZ	O1	O2
3	6.92	7.173	8.496
5	6.745	7.172	8.142
7	6.549	6.949	8.005

Figure 5. Average power of occipital-α signal for different memory load.



	OZ	O1	O2
3	2.15	2.157	2.321
5	2.166	2.21	2.314
7	2.137	2.177	2.301

Figure 6. Average amplitude of occipital-α signal for different memory load.

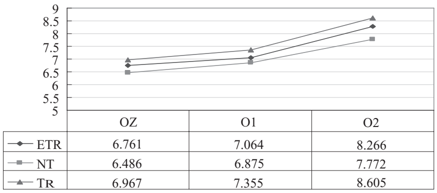


Figure 7. Average power of occipital-α signal for different verbal transformation level.

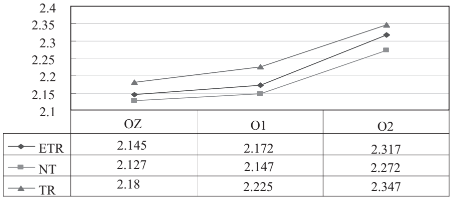


Figure 8. Average amplitude of occipital-α signal for different verbal transformation level.

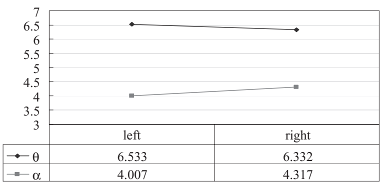


Figure 9. Average powers of θ and α signal for different hemisphere.

power of α signal at occipital electrodes, whereas NT elicited the weakest. The difference in power between verbal transformation levels reached a level of significance at FP2, and marginally FP1. Moreover, all verbal transformation levels elicited stronger power of α signal at O2 than at OZ and O1.

4.2.3 *Hemisphere*

According to the layout of 10–20 system, we selected FP1, F3, F7, C3, O1, and T5 to represent the left hemisphere, whereas FP2, F4, F8, C4, O2, and T6 to represent the right side. The average power and amplitude of EEG response, including θ and α band, acquired from these electrodes were averaged and are shown as Figure 9~10. The statistics shows that during the task operating, the average power of θ rhythm was stronger in left hemisphere than in right. On the contrary, the activities of α band were more inhibited in left hemisphere.

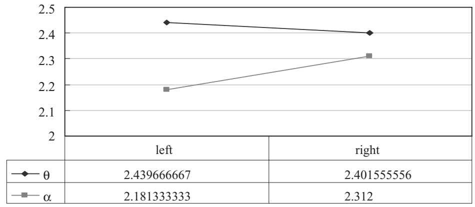


Figure 10. Average powers of θ and α signal for different hemisphere.

5 DISCUSSIONS & SUGGESTIONS

This investigation researches the effects of the amount of characters remembered and the verbal transformation on working performance and brain wave. The discussions and suggestions would be divided in finish time, error number, and EEG response.

5.1 *Finish time*

The amount of characters remembered "7" resulted in the longest finish time, whereas "3" caused the shortest. Based on the same size of matrix in the task, the subjects had to spend more time to search the appropriate characters in their short-term memory database when the amount of characters remembered increased.

On the other hand, ETR led to the longest finish time. ETR asked subjects to recognize the characters regardless of neither capital nor lowercase, which means, subjects not only had to transform the letters, but also keep the originals in mind at the same time during the tasks. TR led to the medium length of finish time because subjects only had to transform the letters and left the originals behind. Subjects spent the least time on NT because they just had to memorize the letters directly and no transformation needed.

In the further working design, designers should seriously consider about the volume of memory load and verbal transformation if they hope to shorten the working-hour.

5.2 *Error number*

The amount of characters remembered "7" resulted in the largest error number, whereas "3" caused the smallest. The more characters needed to be remembered, the more easily subjects failed or made mistakes in memorization and retention, and then recognized the wrong letters.

Referring to the level of verbal transformation, TR caused the largest error number because subjects felt difficult in transformation and easily got confused if they answered the correct answer. ETR and NT resulted in smaller error number because subjects could directly see the characters remembered on the screen, whereas they couldn't in TR tasks.

In the further working design, designers should consider about the degree of presentation and appearance of the sense units in order to avoid confusions, even if those sense units are so familiar.

5.3 *EEG response*

5.3.1 θ *band*

There were a few studies mentioned that the frontal-θ reflects WMS. (Klimesch et al., 1997; Sarnthein et al., 1998; Klimesch et al., 1999). Furthermore, Gevins et al. (1997) indicated that frontal-θ reflects the difficulty of memorization and level of attention. The similar result is shown in this research, the θ rhythm was stronger for the amount of characters remembered "7" than the others, which means, subjects felt difficult and paid more attention when the characters remembered increased.

The θ rhythm was strongest for TR and weakest for ETR. Consequently, subjects felt more difficult and paid more attention when they had to transform to something they couldn't see directly and easily got confused. This condition just fits in with the result of error number of which was mentioned above.

5.3.2 α *band*

Klimesch et al. (1999) and Jensen et al. (2002) had both shown that α signal reflects the inhibition level of brain region, which means, occipital-α signal increased in strength with the degree of inhibition of visual cortex. In this investigation, the amount of characters remembered "7" elicited the weakest occipital-α signal. Subjects had to watch more closely and focus on the characters remembered as the amount increased. Therefore, as the amount of characters remembered increased, the visual cortex in subject's brain was more inspired to work carefully in order to attenuate the confusion caused by memory overload.

The same situation happened for the level of verbal transformation, that is, TR elicited the strongest occipital-α signal. Since the strength of α signal is inversely related to the proportion of cortical resource required by task, TR required more mental effort transforming to the letter that subject didn't see on the screen directly. (Pfurtscheller et al., 1996) NT resulted in the weakest occipital-α signal because subjects just had to memorize and confirm those characters visually without transformation in mind.

5.3.3 *Hemisphere*

The average power of θ rhythm was stronger in left hemisphere than in right, which means, the left hemisphere was relatively active during the tasks. In addition, the weaker α signal in the left side indicates its less idling state than the right side. In sum, the above-mentioned statements about hemispheres suggest that the left hemisphere, which is in charge of language ability, was presumably more functional in the task of verbal memorization and transformation.

REFERENCES

Berger, H. 1929. Uber das elektroenzephalogramm des menschen. *Arch Psychiatr* 87: 527–570.

Cushman, W.H. 1986. Reading from microfiche, a VDT, and the printed page: Subjective fatigue and performance. *Human Factors* 28: 63–73.

Galbraith, G.C., Jhaveri, S.P. & Huo, J. 1997. Speech-evoked brain frequency-following responses during verbal transformations due to word repetition. *Electroencephalography and Clinical Neurophysiology* 102: 46–53.

Gevins, A., Smith, M.E., Leong, H., McEvoy, L., Whitfield, S., Du, R. & Rush, G. 1998. Monitoring working memory load during computer-based task with EEG pattern recognition methods. *Human Factors* 40: 79–97.

Gevins, A., Smith, M.E., McEvoy, L. & Yu, D. 1997. High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex* 7: 374–385.

Jensen, O., Gelfand, J., Kounios, J. & Lisman, J.E. 2002. Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex* 12: 877–882.

Klimesch, W. 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews* 29: 169–195.

Klimesch, W., Doppelmayr, M., Pachinger, T. & Ripper, B. 1997. Brain oscillations and human memory performance: EEG correlates in the upper alpha and theta bands. *Neuroscience Letters* 238: 9–12.

Kondo, H.M. & Kashino, M. 2007. Neural mechanisms of anditory awareness underlying verbal transformation. *Neuroimage* 31: 123–130.

Miller, E.K. & Desimone, R. 1994. Parallel neuronal mechanisms for short-term memory. *Science* 263: 520–522.

Miller, G.A. 1956. The magic number seven plus or minus two: some limits on our capacity for processing information. *Psychology Review* 63: 81–97.

Pfurtscheller G., Stancak A. & Neuper C. 1996. Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: a review. *Int J Psychophysiol* 24: 39–46.

Röhm, D., Klimesch, W., Haider, H. & Doppelmayr, M. 2001. The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters* 310: 137–140.

Rolke, B., Heil, M., Hennighausen, E., Haussier, C. & Rosler F. 2000. Topography of brain electrical activity dissociates the sequential order transformation of verbal versus spatial information in humans. *Neuroscience Letters* 282: 81–84.
Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G.L. & Stein A.V. 1998. Synchronization between prefrontal and posterior association cortex during human working memory. Proceedings of the National Academy of Science USA, 95, 7092–7096.
Tremblay, P., Shiller, D.M. & Gracco, V.L. 2008. On the time-course and frequency selectivity of the EEG for different modes of response selection: Evidence from speech production and keyboard pressing. *Clinical Neurophysiology* 119: 88–99.
Tung, H.W. & Lin, C.T. 2005. A study of EEG correlates of unexpected obstacle dodging task and driving style. National Chiao Tung University. (Master Thesis).
Wilson, F.A.W., Scalaidhe, S.P.O. & Goldman-Rakic P.S. 1993. Dissociation of object and spatial processing in primate prefrontal cortex. *Science* 260: 1955–1958.
Woodman, G.F., Vogel, E.K. & Luck, S.J. 2001. Visual search remains efficient when visual working memory is full. *Psychol Science* 12: 219–224.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The effects of whole-body motion direction on the efficiency of trackball interfaces

Yi-Jan Yau, Chin-Jung Chao, Wen-Yang Feng & Feng-Yi Tseng
System Development Center, Chung-Shan Institute of Science and Technology, Lung-Tan, Tao-Yuan, Taiwan

Sheue-Ling Hwang
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

ABSTRACT: The trackball is often used as a human interface input device in moving vehicles. However, few studies have explored the effects of Whole-Body Motion Direction (WBMD) on its manipulation. Fewer still have considered the possible detrimental effect of WBMD in designing the user interface and the arrangement of console location and orientation simultaneously. In this study, an experiment was conducted in a motion simulator to investigate the effects of WBMD and the Cursor Movement Direction (CMD) on the efficiency of manipulating a trackball. The results indicated that the WBMD and CMD both significantly affected the movement times but not the accuracy of target selection. The movement times were significantly longer under rolling and pitching motion, and the cursor movements along the horizontal and vertical axes from the start position were faster than those along the diagonals. The implications of these findings could be applied to optimize console and graphical menu design for use on moving vehicles.

Keywords: trackball, whole-body motion, roll, pitch, heave

1 INTRODUCTION

With the proliferation of Graphical User Interfaces (GUIs), Non-Keyboard Input Devices (NKIDs) such as mice, touch screens, trackballs and joysticks have become the major input devices. Numerous studies (Atkinsona et al., 2004; Burgess-Limerick and Green, 2000) have mentioned that the mouse was the most popular input device in GUIs; hence, many interface design principles and guidelines were generally developed for mouse users (MacKenzie and Buxton, 1992; Whisenand and Emurian, 1996). On the other hand, mouse not fixed on a console is easy to drop in a moving vehicle and needs room to manipulate; hence, mouse is neither suitable for use in motion environments nor appropriate for constrain work station. In such situations, trackball, joy stick or touch screen are more suitable in motion environments. However, in comparison with the researches on interface design for the mouse users, fewer researches have considered the trackball users.

Point-and-click task is one of the major input tasks on moving vehicles, and most of such tasks are executed on 2D visual display interface. Users need to manipulate NKIDs to move a cursor toward graphical objects on different angles of the interface and click a button to select a target or a command (Accot and Zhai, 2002). Some researches have indicated that the cursor movements along different angles would affect the input performance such as movement speed and accuracy. However, most of the researches focused on the manipulation of mouse rather than on the trackball. Whisenand and Emurian (1996) pointed out that the Cursor Movement Direction (CMD) affected the movement times of target selection with the mouse. Kurtenbach (1993) proved the hypotheses that the marked menu on-axis was faster and produced fewer errors than marks menu off-axis with pen-base computer. Very few studies have investigated the effects of CMD on the manipulation of a trackball.

Numerous studies have pointed out that the motion environments would affect the user's input performance. According to the reviews by Lewis and Griffin (1978) and McLedo and Griffin (1989), the whole-body vibration could interfere with the performance of continuous manual control tasks. McLedo and Poulton (1980) found that performance of using a joystick to execute the pursuing tracing tasks was affected by ship motion. Griffin (1990) indicated that the motion would affect the accuracy of small or precise manual movements. Hill and Tauson (2005) pointed out that the soldiers' input tasks were seriously affected by vehicle motion. Grandt et al. (2003) indicated that trackball was a conventional input device used on battleship and the users generally faced the targets located at different angles from a start position; hence, it is very important to understand the effects of different cursor movement direction of a trackball and different motion directions of the vehicle on manipulating a trackball. With respect the vehicle motion, ship motion includes not only heave (vertical) motion, but also roll, pitch, yaw, sway and surge motions. Pingree (1988) indicated that the heave, roll and pitch with varying frequencies of up to 1 Hz predominated over the ship motions. One can see that many studies have investigated the effects of motion on manual control tasks, nevertheless, most of them focused on the vertical vibration. Only few studies have investigated the effects of whole-body motion on manipulating a trackball (Yau et al., 2008).

This study aims to investigate the effects of WBMD (heave, roll and pitch) and CMD (0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°) on the movement times and error rate of manipulating a trackball on point-and-click tasks.

2 METHOD

2.1 *Subjects*

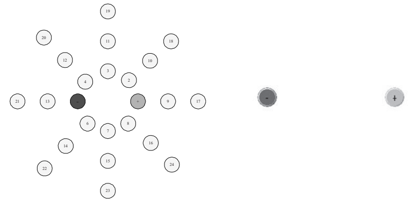
Twelve paid male participants were right-handed experienced mouse-users but no experience of using trackball. They all had corrected visual acuity of 0.8 or better as well as normal color vision. Their mean age was 26.5 years, raging from 22 to 30.

2.2 *Apparatus*

The experiment was conducted in a cabin loaded on a Stewart platform which has six degrees of freedom. Its capacity displacements of heave, sway and surge are ±200 mm, ±250 mm and ±250 mm, and the angles of roll, pitch, and yaw are ±15°, ±15° and ±17°, respectively. Point-and-click tasks were performed on a 17-in Color TFT LCD with 1024 × 768 resolutions. Input device was a thumb trackball and the experimental software was developed by Visual C++. The orientation of trackball was vertically aligned to the screen to keep the same movement angles of the cursor and trackball in this experiment.

2.3 *Procedure*

The whole-body motions were provided by the Stewart motion platform with heave (0.3 Hz, ± 100 mm), roll (0.3 Hz, ±8°) and pitch (0.3 Hz, ±8°). Participants manipulated the trackball to perform multi-directional point-and-click tasks in a static setting as well as these three motions. Participants were requested to evaluate their health condition and motion sickness by filling questionnaires at the beginning, middle, and end of experiment. Motion sickness was assessed on the Misery Scale (MISC), an 11-point discomfort rating scale from no problems (0) to vomiting (10), which can be filled out within seconds. After reading the experimental instruction, participants used their preferred hand to move the cursor to select the targets as quickly and accurately as possible. A 60-minute practice was offered prior to experiment for the subjects to be familiar with using trackball on the task under static and different motion directions. At the beginning of the experiment, participants positioned the cursor toward the center circle of operating interface and clicked it to commence a trial. A pair of circle targets (diameter = 7.5 mm) of which one was filled green with a black crosshair pointer (+) and another was filled red with a black minus sign (–) appeared. Participant moved and clicked the cursor from the green circle to red circle. The black crosshair or minus sign turned white while the subject successfully clicked inside the boundary of the green or red circle. An error was defined as clicking outside the red circle and a beep was sounded while click outside the green or red circles. Times were recorded between two successful clicks. Once successfully selected, another new pair of circles randomly appeared. If the movement of cursor over 7 seconds (the time limit), then the movement was counted as an error and the movement time was not recorded. Subsequently, the next pair of targets appeared. This process continued until all trials in a set were completed. A set consisted of twenty-four point-and-click trials (1 motion situation; 3 different distances; 8 different angles). The distances between the pairs of circles were about 75.7 mm, 151.4 mm and 227.1 mm and, the approach angles were 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° and 360° (see Fig. 1). Each participant completed ten repeat sets (a block) of point-and-click tasks under static, heave, roll and



(a) Interface design layout (b) A sample

Figure 1. Experimental interface layout and a sample.

pitch respectively. Totally there were 960 trials of point-and-click needed to be finished.

A fully within-subjects repeated measures design was applied. The independent variables, whole-body motion directions (4 levels, static, heave, roll, and pitch) and cursor movement angles (8 levels, 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° and 360°) were all within-subject factors. The dependent variables were the Movement Time (MT, millisecond) and Error Rate (ER, %). The movement time was the time between two successive clicks, and the error was defined as the target selected outside the red circle. The Mean Movement Time (MMT) is the arithmetic mean of movement times on each movement direction and the Error Rate (ER) was the percentage of error occurred on each direction of whole-body motion in each block. A two-way repeated measure ANOVA was used to examine the effects of the WBMD and CMD on the movement time and error rate.

3 RESULTS

3.1 *Movement time*

The results of ANOVA indicate that both the main effects of WBMD and CMD are significant on the MMT ($F(3, 33) = 4.774$, $p < 0.01$, $F(7, 77) = 6.439$, $p < 0.01$, respectively) whereas their interaction is not significant ($F(21, 231) = 0.061$, $p > 0.05$). The relationships between average MMT, and WBMD and CMD are shown in Fig. 2 and Fig. 3, respectively. The Fig. 4 presents the average MMT of cursor movement in various CMDs and WBMDs. The movement speed of manipulating a trackball on static is the fastest, followed by the heave, roll, and the pitch. Generally, the movements along the angles of 0°, 90°, 180° and 270° are faster than those along the angles of 45°, 135°, 225° and 315° under all motion directions. A Duncan multiple range test was conducted to compare the subsets among the means (Tables 1 and 2). The comparisons between different WBMD indicate that the effects on movement speed of the groups of pitch and roll are significantly larger than those of the groups of roll and heave, and the groups of heave and static. The comparisons between different CMDs reveal that the movements along the vertical and horizontal axes are faster than the movements along the diagonals.

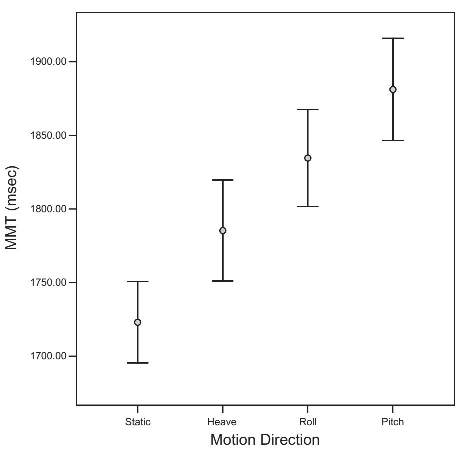


Figure 2. Average MMT for WBMD.

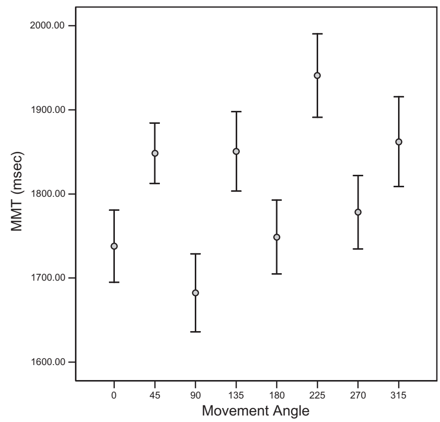


Figure 3. Average MMT for CMD.

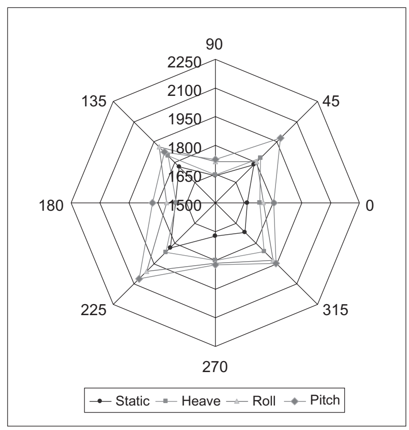


Figure 4. Average MMT in various CMD and WBMD.

3.2 *Accuracy*

Table 3 lists the error rate in various movement angles and motion directions. The results of ANOVA indicate that neither the motion direction nor the movement angle is significant on accuracy

Table 1. Duncan test of WBMD on average MMT (unit: msec).

Motion direction[b]	N	MMT	Duncan grouping[a]		
Static	96	1723	A		
Heave	96	1785	A	B	
Roll	96	1834		B	C
Pitch	96	1881			C

[a] Values with the same letter are not significantly different.
[b] Significantly different at $\alpha = 0.05$.

Table 2. Duncan test of CMD on average MMT (unit: msec).

Movement angle[b]	N	MMT	Duncan grouping[a]		
90°	48	1683	A		
0°	48	1738	A	B	
180°	48	1749	A	B	
270°	48	1778	A	B	
45°	48	1848		B	C
135°	48	1851		B	C
325°	48	1862		B	C
225°	48	1941			C

[a] Values with the same letter are not significantly different.
[b] Significantly different at $\alpha = 0.05$.

((F (3, 33) = 2.438, $p = 0.082$, F (7, 77) = 1.246, $p = 0.288$)). The interaction between motion direction and movement angle has no significant effect on the accuracy.

4 DISCUSSION

4.1 *Whole-body motion direction*

Although the degraded input performance caused by motion is similar to that reported in the previous studies (Bittner and Guignard, 1985; Griffin, 1990; Hill and Tauson, 2005; Lewis and Griffin, 1978; McLedo et al., 1980; McLeod and Griffin, 1989), the previous studies have not investigated the effect of WBMD on input performance and most of them focused on discussing the effect of the high frequency motion. The experimental results of current study reveal that the WBMD affected the speed of cursor movement by using trackball on point-and-click task, whereas the effects are not significant on the error rate. Generally, the effects of roll and pitch on the movement speed are larger than those of heave and static. The longer selection time for users implied that participants felt more difficult in selecting targets (Hancock and Booth, 2004). However, the results show that there is neither significant difference between pitch and roll nor between heave and static. With respect to the heave and static, this result is not consistent with the previous finding that the vertical vibration would affect the performance by using other input devices such as joystick or control lever to track targets (Lewis and Griffin, 1978; McLeod and Griffin, 1989). The reasons could be the heave motion of this experiment was a smoothly vertical movement with low frequencies (0.3 Hz), and the tasks as well as input device were also different from those of the previous studies; hence the effects of heave motion were few and not significant. In regard to the roll and pitch, the effects of pitch and row on movement time are not significant difference. It is possible because the effects of the WBMD are alleviated by different CMD.

Table 3. Error rate in various CMDs and WBMDs.

	0°	45°	90°	135°	180°	225°	270°	315°	Ave.
Static	3.7	5.6	3.7	2.8	6.5	7.4	1.9	4.6	4.5
Heave	4.6	4.6	3.7	5.6	3.7	4.6	8.3	3.7	4.9
Roll	8.3	6.5	4.6	9.3	4.6	10.2	6.5	3.7	6.7
Pitch	7.4	8.3	5.6	10.1	8.3	8.3	5.6	3.7	7.2
Ave.	6.0	6.3	4.4	7.0	5.8	7.6	5.6	3.9	5.8

Concerning the accuracy of target selection, the WBMD would not affect the accuracy of manipulating trackball to select targets. The possible explanations for this could be (1) the motion conditions such as frequency and amplitude are not severe to affect the accuracy of targeting, and (2) the participants could predict the motion direction in advance, hence they would know how to adjust their manipulation to respond to these motions. Although the effects of WBMD on error rate are not significant ($p = 0.08$), we observ a tendency of increasing error rate from static (4.51%), heave (4.86%), roll (6.71%) to pitch (7.17%) (see Table 3). In comparison with the error rates of the studies of MacKenzie and Buxton (1992) (4.6%) and Mackenzie et al. (1991) (3.9%), the error rates of the static and heave of the present study are very close to those of previous studies, but the error rates of roll and pitch are higher than those of their studies. If the motion amplitude and frequency of roll and pitch are increased, the effects of WBMD on accuracy could be more significant than those of the current study.

4.2 *Cursor movement direction*

The CMD significantly affects the movement time but not the accuracy. The experimental results indicate that the cursor movements along vertical and horizontal directions are faster than that of movements along diagonal directions. The results not only confirm the findings of MacKenzie and Buxton (1992) that the speed of cursor movement along 90° and 0° was faster than that of movement along 45° with mouse but also extend their findings to 270°, 180° and other diagonal directions with the trackball. This result is also consistent with the findings of Jagacinski and Monk (1985) that the movement along the horizontal and vertical axes were faster than that of along the diagonals with joystick and the findings of Whisenand and Emurian (1996) suggested that arranging target-type objects with the approach angles of 0°, 180° and 270° on interface would improve the target selection performance with a mouse. The more likely explanations of these results could be attributed to the nature of humans' physical structures that the movement is more rapid in certain directions than in other directions (McCormick and Sanders, 1992), and the differences in the muscle groups are required to manipulate cursor movements in different directions (Grossman and Balakrishnan, 2005). Although the movement controlled pivots of biomechanical system and muscle groups with a mouse and a trackball are different (Karlqvist et al., 1999), the results of this study (with trackball) and previous studies (with mouse) which are relevant to the speed of different movement angles are very similar. One possible reason could be explained is that the use of smaller muscle groups on axis movements would result in easier and faster movement than that of the large muscle groups required on the diagonal movements.

5 CONCLUSIONS

A more thorough understanding on the difficulty of cursor movement with trackball in motion environment was discussed in this study. The experimental results illustrated that the whole-body motion direction and cursor movement direction affected the speed of target selection. The effects of pitch and roll on movement times were bigger than those of static as well as heave, and the speeds of cursor movements along the diagonals were slower than those of movements along the vertical and horizontal axes. Our finding could help the designers to develop more efficient and comfortable interfaces to fit user needs; moreover, it could be applied to consol location/orientation arrangements.

REFERENCES

Accot, J. & Zhai, S. 2002. More than dotting the i's-foundations for crossing-based interfaces. In: *Proceedings of the CHI 2002 conference on human factors in computing systems* 73–80.

Atkinsona, S., Woodsb, V., Haslamc, R.A. & Buckleb, P. 2004. Using non-keyboard input devices: interviews with users in the workplace. *International Journal of Industrial Ergonomics* 33: 571–579.

Bittner, A.C. & Guignard, J.C. 1985. Human factor engineering principles for minimizing adverse ship motion effects: theory and practice. *Naval Engineering Journal* May: 205–213.

Burgess-Limerick, R. & Green, B. 2000. Using multiple case studies in ergonomics: an example of pointing device use. *International Journal of Industrial Ergonomics* 26: 381–388.

Grandt, M., Pfendler, C. & Mooshage, O. 2003. Empirical comparison of five input devices for anti-air warfare operators. In: *Proceedings of the 8th International Command and Control Technology Symposium* (ICCRTS).

Griffin, M.J. 1990. *Handbook of human vibration*. Academic press London, UK.

Grossman, T. & Balakrishnan, R., 2005. A probabilistic approach to modeling two-dimensional pointing. *ACM Transactions on Computer-Human Interaction* 12(3): 435–459.

Hancock, S.M. & Booth, K.S. 2004. Improving menu placement strategies for pen input. In: *Proceedings of the 2004 conference on Graphics interface*: 221–230. London, Ontario, Canada.

Hill, S.G. & Tauson, R.A. 2005. Soldier Performance Issues in C2 "On the Move". In: *Proceedings of the 10th International Command and Control Technology Symposium* (ICCRTS).

Jagacinski, R.J. & Monk, L.D. 1985. Fitts' law in two dimensions with hand and head movements. *Journal of Motor Behavior* 17(1): 77–95.

Karlqvist, L., Bernmark, E., Ekenvall, L., Hagberg, M., Isaksson, A. & Rosto, T. 1999. Computer mouse and track-ball operation: Similarities and differences in posture, muscular load and perceived exertion. *International Journal of Industrial Ergonomics* 23: 157–169.

Kurtenbach, G.P. 1993. *The design and evaluation of marking menus*. Ph.D. thesis, University of Toronto, Canada.

Lewis, C.H. & Griffin, M.J. 1978. A review of the effects of vibration on visual acuity and continuous manual control, part II-Continuous manual control. *Journal of Sound and Vibration* 56(3): 415–457.

MacKenzie, I.S. & Buxton, W. 1992. Extending Fitts' law to two dimensional tasks In: *Proceedings of the CHI 1992 Conference on Human Factors in Computing Systems* 219–226.

McCormick, E.J. & Sanders, M.S. 1992. *Human factors in engineering and design.* McGraw-Hill: New York.

McLedo, P., Poulton, C., Ross, H.D. & Lewis, W. 1980. The influence of ship motion on manual control skills. *Ergonomics* 23: 623–634.

McLeod, R. & Griffin, M. 1989. A review of the effects of whole-body translational vibration on continuous manual control performance. *Journal of Sound and Vibration* 133(1): 55–115.

Pingree, B.J.W. 1988. A review of human performance in a ship motion environment. *Warship Technology* May: 73–76.

Whisenand, T.G. & Emurian, H.H. 1996. Effects of angle of approach on cursor movement with a mouse: Consideration of Fitts' law. *Computer in Human Behavior* 12(3): 481–495.

Yau, Y.J, Chao, C.L. & Hwang, S.L. 2008. Effects of input device and motion type on a cursor positioning task. *Perceptual and Motor Skills* 106: 76–90.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The study of operation influence in electronic dictionary's button visual news

Tsu-Wu Hu, Li-An Yang & Ha-Na Shiung
Department of Industrial Design, Chaoyang University of Technology, College of Design, Taichung, Taiwan

ABSTRACT: In recent years emphasized the trend of international students and office workers will be learning a foreign language as a necessary condition for competition, and electronic dictionaries are learning English brought into one of the tools, along with the scientific and technical progress, will it was used by the dictionary, and converted to a mobile for the benefit of carrying electronic product, of consumers frequently used query "textexample textwords" function for major research projects, preceding questions and as far as test to improve a test. This study to research on usability engineering methods and criteria for the experiment to familiarise themselves with the expert that is, in so far as the overall improvement of the standards. Through measuring learned that before: 1. Repeat' for Y-400 Electronic dictionary in Chinese and English in function, "space bar" and the "enter key" function that cannot be easily resolved; 2. Secondly, in calculating the functional calculation symbol "Addition" and "Subtraction" button to find this difficult, depending on the typical work project, and by improving contrast analysis before and after and found that: 1. The same property entity keys placed in the same color, helps operators identify and reducing times; 2. Easy to be confusing the pressed key, the increase color and the writing are auxiliary, reduces the operation mistake probability effectively.

Keywords: usability, electronic dictionary, buttons visual news, user interface

1 RESEARCH PROJECT AND BACKGROUND

Government plans 90 fifth academic year, the official promotion of six grades in English, and in the third and fourth grade school year from 94 simultaneous implementation of teaching English, led early students have basic English communication skills, Send the results from the Ministry of the Interior Commission also pointed out that the ratio of children to learn Talent is the largest proportion of learning a foreign language, also in 2004 the purpose of college students participate in tutorial results, showed that 44.2% is to learn a foreign language, can know the importance of English proficiency. In 1988, the first electronic dictionary market, replace the original traditional English dictionary, enable students to easily carry it, In order to meet the demand for portable, many brands have introduced many of the models, but to many of the features indicated in the limited area, buttons differences will become important considerations affect the operation.

1.1 *Research object*

In the sample selection, based on the experimental purpose and convenience, the university student population (18 to 25 years old) dominated.

1.2 *Define the scope and limitations*

Because so many models of electronic dictionaries label, Therefore, the scope of the study set in the top three leading brands of electronic dictionaries are BESTA INSTANT TECH, HOT TECH. As many brands, for the research of consistency.

Will choose the price to 2,000 yuan to 5,000 yuan of the object model as the main research target, This range is in the class-based models, Due to the price relationship is acceptable to the public and the first reason for considering buying.

1.3 *Steps and implementation schedule*

Research outline	Research projects	Research explanation
Chapter 1 Foreword	1. Research Background and Motivation 2. Research purposes 3. Set the scope of the study 4. Research methods and procedures	Study electronic dictionary of current operations in the Entity button interface part, The button configuration type and description of keywords, For users in the operation and function of the distinction is influential.
Chapter 2 Literature	1. User Interface 2. Usability Engineering 3. HMI 4. Cognitive Design	Understand the existing interface specification, and electronic dictionary keypad interface problems. Explore the interface and user interaction.
Chapter 3 Research methods	Existing literature related to collection of electronic dictionary. Button interface classification of electronic dictionaries 1. Function buttons analysis. 2. Analysis of digital buttons. Characteristics of recombinant.	Electronic dictionary with physical buttons and other products (such as: computer keyboard, phone keypad, mp3 ... etc) Type of configuration and function buttons to do the analysis and classification, To identify current usage patterns useful or available.
Chapter 4 Analysis and discussion	1. Purpose of the experiment. 2. Experimental Methods. 3. Pre-test analysis and discussion. 4. Post-test analysis and discussion.	Use the usability method to evaluate. Test and evaluation of the original interface. Modify the interface after the experiments and evaluation
Chapter 5 Conclusions and recommendations	1. Research founding. 2. Practical application of research results. 3. Follow-up research and recommend.	The results of the study found, and give advice in the interface configuration. Description of the deficiencies, and provide direction for further research with the proposal.

2 RESEARCH PROPOSES

In recent years emphasized the need to have an international perspective trend, Students and office workers will be learning a foreign language as a necessary condition for competition, The electronic dictionary has become a carried tool for learning English, As technology advances, people used to use the traditional dictionary, Into a compact size to facilitate carrying of electronic products. Although the volume into a portable, However, many additional features also allow products to become not easy to use.

15-inch notebook button configuration, about 86 to 90 keys, for example Hot Tech Y400, 11 × 7 cm, a total of 64 key. Almost every key there are 2 or more functions, ex: Hot Tech Y400, enter key also represents the equivalent key, if you press the space bar there will be no action, even the space key also has other functions, however, in computing the keys it does not have equivalent function. But the unknown is, how can users know that, when operation a process, what kind of button is available and useful. Therefore, inefficient use will only increase the frustration of users.

The function of electronic dictionary are dictionary, learning, entertainment, IT, computing, Wikipedia, systems, Lin Hongxiang (2006) the use of electronic dictionaries for junior high students study that currently used by the basic type of electronic dictionary majority, Electronic dictionary to buy the main motivation to academic inquiry, Second is the entertainment uses; and check “English words” frequently used.

In this study, the visual message electronic dictionary operating mode button—English to Chinese function, Chinese to English function and Computing to conduct experiments, The main purpose of the user keys in the operation of electronic dictionary, Its operating performance and related factors affect the performance. In this study, reference samples of the products available in the market analysis and testing, as the models currently on the market have extended the brand habit, hope that this study can be summarized a usage model and interface design, for the further when extension making models have better recommendations and direction.

2.1 *User interface design*

If User interface design or user interface engineering is the design of computers, appliances, machines, mobil communication devices, software applications, and websites with the focus on the user's experience and interaction. The goal of user interface design is to make the user's interaction as simple and efficient as possible, in terms of accomplishing user goals—what is often called user-centered design. Good user interface design facilitates finishing the task at hand without drawing unnecessary attention to itself. Graphic design may be utilized to support its usability. The design process must balance technical functionality and visual elements (e.g., mental model) to create a system that is not only operational but also usable and adaptable to changing user needs.

There are several phases and processes in the user interface design, some of which are more demanded upon than others, depending on the project.

Functionality requirements gathering—assembling a list of the functionality required by the system to accomplish the goals of the project and the potential needs of the users.

User analysis—analysis of the potential users of the system either through discussion with people who work with the users and/or the potential users themselves.

Typical questions involve:

What would the user want the system to do?

How would the system fit in with the user's normal workflow or daily activities?

How technically savvy is the user and what similar systems does the user already use?

What interface look & feel styles appeal to the user?

Information architecture—development of the process and/or information flow of the system (i.e. for phone tree systems, this would be an option tree flowchart and for web sites this would be a site flow that shows the hierarchy of the pages).

Prototyping—development of wireframes, either in the form of paper prototypes or simple interactive screens. These prototypes are stripped of all look & feel elements and most content in order to concentrate on the interface.

Usability testing—testing of the prototypes on an actual user—often using a technique called think aloud protocol where you ask the user to talk about their thoughts during the experience.

Graphic Interface design—actual design of the final Graphical User Interface (GUI). It may be based on the findings developed during the usability testing if usability is unpredictable, or based on communication objectives and styles that would appeal to the user. In rare cases, the graphics may drive the prototyping, depending on the importance of visual form versus function. If the interface requires multiple skins, there may be multiple interface designs for one control panel, functional feature or widget. This phase is often a collaborative effort between a graphic designer and a user interface designer, or handled by one who is proficient in both disciplines.

2.2 *Mental model*

A mental model is an explanation of someone's thought process about how something works in the real world. It is a representation of the surrounding world, the relationships between its various parts and a person's intuitive perception about their own acts and their consequences. Our mental models help shape our behaviour and define our approach to solving problems (akin to a personal algorithm) and carrying out tasks.

2.3 *Principles of mental models*

Mental models are based on a small set of fundamental assumptions, which distinguish them from other proposed representations in the psychology of reasoning (Byrne & Johnson-Laird, 2009). Each mental model represents a possibility. A mental model represents one possibility, capturing what is common to all the different ways in which the possibility may occur (Johnson-Laird & Byrne, 2002). Mental models are iconic, i.e., each part of a model corresponds to each part of what it represents (Johnson-Laird, 2006). Mental models are based on a principle of truth: they represent only those situations that are possible, and each model of a possibility represents only what is true in that possibility according to the proposition. Mental models can represent what is false, temporarily assumed to be true, e.g., in the case of counterfactual conditionals and counterfactual thinking (Byrne, 2005).

3 RESEARCH METHODS

In this study adopt the observation and Performance Measurement as the use of assessment methods. In this study, the assumptions mentioned are: 1. Explore the product key on the graphic design and appropriateness of placement, and affect the user's knowledge and learning even the ease of operation. 2. Built-in does not change the original mode of operation, use color to group the related button to each other, it's contribute to product operation indicative. Borrow this suppose, and text by Hot Tech Y400 for pretest. Experimental place is in the product usability and ergonomics design laboratory, the subjects were based on fixed operating procedures. This experiment was set up around the corner of the 3 cameras, through on-site video record of the subjects the way of operation, recorded by the researchers. Scoring criteria according to experts of the time as a benchmark.

The scoring way is: Completed within standard time (0 points). More than the standard time to complete but for the permissible range (–3 points). Over time and after asking to be finalized (–5 points). Over time and think you do right on and proceed to the next step (–8 points). And on this basis to make improvements to posttest. Using usability way to do this research axis, First of all, is setting up the typical working item for research projects. Typical working item represents the product operation procedures, For example, electronic dictionary function of English to Chinese function: (a) power on (b) Catalog (English to Chinese function) (c) Chinese input: Happiness→ finding the happiness English (d) Esc.

To participate in this experiment testers were divided into pretest and posttest of all six, a total of 12, and 6 male, 6 female, Age range from 19 to 26 years of age; Another expert (acquaintances), pretest and posttest the two, a total of four. All Testers are Chao-yang University of Technology, department of Industrial Design students, Based on the Usability of the literature said that when the testers have five, Can be found in 75%–80% of usability problems (Nielson, 1993).

To do with the usability way of the main axis of this study, first of all, the subject matter of the typical experimental set up work projects, set the procedures for the Chinese to English Function, English Chinese to Function and computing three stages.

Chinese to English Function (1-c) when input Chinese, testing often confuse "enter key" and "space bar", and the misconception that the electronic dictionary is broken there is no response.

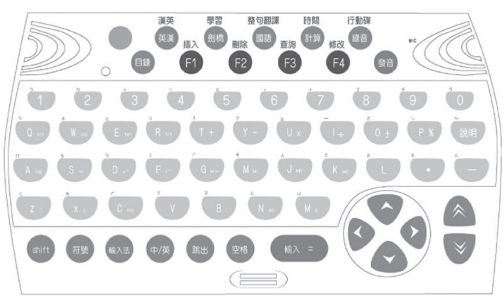


Figure 1. The original interface.

Computing (3-c) when calculated at the typed symbols plus or minus, Test will often be confused minus sign with the English link No, and the Plus confused with Cang Jie Input Method 「+」. From Table 2, Table 3, Table 4 can be know that In the steps 1-b, 1-c, 2-c, 3-c, the operate is not easy to use, This part of the improvement plan to strengthen the icon (see Figure 2).

Because this electronic dictionary BoPoMoFo phonetic symbols is light purple and printing in the machine body, Therefore, no change to the original operating principles, But in the semi-circle with the space bar next to the light purple color block and there BoPoMoFo words to guide users. Function

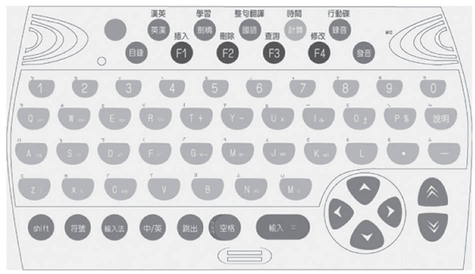


Figure 2. The improved user interface.

Enlarge map to improved interface:

in the Chinese to English press the space bar will take any action not the enter key.

Computing 3-c is also used the color to improve that will be groups the button of calculation, users can quickly find the function keys. (As Figure 4, Figure 5, Figure 6) Will group the Computing button by light green color, users can quickly find the

Improved interface buttons map (Computational domain).

Table 1. Number of people participating in the experiment.

	Before improving	Improved
Experts	2	2
Testers	6	6

unit: people

Table 2. The original time table for the operation interface (Chinese to English Function).

Working item testers	1-a power on	1-b Catalog→ Chinese to English function	1-c Chinese input	1-d esc
1	0:02	0:03	2:53	0:02
2	0:02	0:03	1:15	0:02
3	0:02	0:03	1:09	0:03
4	0:03	0:04	1:57	0:03
5	0:02	0:04	2:00	0:02
6	0:02	0:03	2:32	0:03
Weighted total	0	–6	–30	0
Expert	0:02	0:02	0:24	0:02

Table 3. The original time table for the operation interface (English to Chinese Function).

Working item testers	2-a power on	2-b Catalog→ English to Chinese function	2-c english input	2-d esc
1	0:02	0:02	0:14	0:02
2	0:02	0:02	0:27	0:02
3	0:02	0:03	0:21	0:02
4	0:02	0:03	0:25	0:02
5	0:02	0:03	0:27	0:02
6	0:02	0:03	0:31	0:02
Weighted total	0	0	–12	0
Expert	0:02	0:02	0:20	0:02

relevant keys, the equal sign on enter key is also group by light green.

3.1 *Conclusion*

According to pre-test found that 1-b, 1-c, 2-c, 3-c of the operation took more time, therefore, according to their original icons do to improve, but without changing the original positions of button, Reservation the original button's features and the

Table 4. The original time table for the operation interface (Computing).

Working item testers	3-a power on	3-b Catalog→ computing	3-c input formula	3-d esc
1	0:02	0:02	3:09	0:02
2	0:02	0:02	1:04	0:02
3	0:02	0:03	2:53	0:02
4	0:02	0:03	0:11	0:02
5	0:02	0:03	2:04	0:02
6	0:02	0:03	2:28	0:02
Weighted total	0	0	–30	0
Expert	0:02	0:02	0:23	0:02

Table 5. Schedule after the test operation (Chinese to English Function).

Working item testers	1-a power on	1-b Catalog→ Chinese to English function	1-c Chinese input	1-d esc
1	0:02	0:03	0:25	0:02
2	0:02	0:02	0:30	0:02
3	0:02	0:03	0:25	0:03
4	0:03	0:02	0:22	0:03
5	0:02	0:02	0:40	0:02
6	0:02	0:03	0:25	0:03
Improve the total weighted	0	0	–6	0
Original total weight	0	–6	–30	0
Expert	0:02	0:02	0:24	0:02

vocabulary, only to improve it next to join the auxiliary. A good interface design must have uniformity (consistency) and the difference (visual recognition of degrees), The target of this experiment do not have the above conditions, In the middle button area is composition by light gray button, the light gray keys are also arranged in different white text and symbols, degree in visual recognition is not easy to distinguish, therefore, the steps in the implementation of computing time, you can not immediately find the "+− × ÷" symbol, and easy to confuse the other symbols are plus or minus sign into the function.

Will be groups to the computing of related buttons, Use color to segment them, In the original all the computing buttons are light gray to group them in light green, To Identified as the degree is

Table 6. Schedule after the test operation (English to Chinese Function).

Working item testers	2-a power on	2-b Catalog→ English to Chinese function	2-c English input	2-d esc
1	0:02	0:02	0:10	0:02
2	0:02	0:02	0:20	0:02
3	0:02	0:03	0:21	0:02
4	0:02	0:03	0:15	0:02
5	0:02	0:03	0:17	0:02
6	0:02	0:03	0:22	0:02
Improve the total weighted	0	0	–6	0
Original total weight	0	0	–12	0
Expert	0:02	0:02	0:20	0:02

Table 7. Schedule after the test operation (Computing).

Working item testers	3-a power on	3-b Catalog→ computing	3-c input formula	3-d esc
1	0:02	0:02	0:12	0:02
2	0:02	0:02	0:24	0:02
3	0:02	0:03	0:20	0:02
4	0:02	0:03	0:17	0:02
5	0:02	0:03	0:23	0:02
6	0:02	0:03	0:15	0:02
Improve the total weighted	0	0	–6	0
Original total weight	0	0	–30	0
Expert	0:02	0:02	0:20	0:02

better than the old has been increased for many. The other functions on the Chinese to English input, The target of this research is set "space" to enter the phonetic to find the identification button of Chinese characters. But the study also found that all of the testers used the "Enter" key as the identification button, so that in operation, prone to make mistakes. It can be concluded that most people's mental model has been established that the "Enter" key to determine as the identification button, therefore, the use of this electronic dictionary of words is less appropriate, Therefore, the improve way is in the semi-circle with the space bar next to the light purple color block and there BoPoMoFo words to guide users. Function in the Chinese to English press the space bar will take any action not the enter key.

Therefore, group the button and differentiate color can helps users efficiency to operation, because it can increase visual recognition of degrees and the experimental results reveal, Using color segmentation to different attributes of button functions, a group of results is also easy for users to operate on a clear and easy to understand.

REFERENCES

Borgman, C.L. 1986. The user's mental model of an information retrieval system: An experiment on a prototype online catalog. *Journal of Man-Machine Studies* 24: 47–64.

Craik, K. 1943. *The nature of explanation*. Cambridge: Cambridge university press.

Fisher, B. 2003. CPSC 444: User interface design—lecture notes. Retrieved from *http://www.cs.ubc.ca/~cs444/lectures/444-4_mentalModels.pdf*

Fowler, S.L. & Stanwick, V.R. 1996. *The GUI style guide*. Cambrige: Academic press.

Gray, S.H. 1990. Using protocol analyses and drawings to study mental model construction during hypertext navigation. *International Journal of Human-Computer Interaction* 2: 359–377.

Johnson-Laird, P.N. 1983. *Mental models*. Cambridge: Cambridge university press.

Johnson-Laird, P.N., Girotto, V. & Legrenzi, P. 1998. *Mental models: A gentle guide for outsiders*. Retrieved from http://www.si.umich.edu/ICOS/gentleintro.html.

McDaniel, S. 2003. *What's your idea of a mental model?* Boxes and arrows.

Nielsen, J. 1993. *Usability engineering*. Morgan Kaufmann: San Francisco.

Norman, D.A. 1983. Some observations on mental models. In D. Gentner & A. Stevens (Eds.), *Mental models*. Hillsdale, NJ: L. Erlbaum Associates.

Norman, D.A. 1986. Cognitive engineering. In N.D.A. & D.S.W. (Eds.) *User-centered system design: New perspectives in human-computer interaction*. Hillsdale, NJ: L. Erlbaum Associates.

Norman, D.A. 1990. *The psychology of everyday actions. In the design of everyday things*. New York: Doubleday/Currency.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. 1994. *Human-computer interaction*. Harlow, England: Addison-Wesley.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Exploring the interaction problems of using ticket vending machines

Yu-Hsing Huang
National Pingtung University of Science and Technology, Pingtung, Taiwan

ABSTRACT: Ticket vending machines has been broadly used today for providing shorter waiting time for users and less man power for service providers. However, there are only 35% of ticket are sold via ticket vending machines in Taiwan High Speed Rail Company. This paper demonstrates a user centered study conducting at Taiwan High Speed Rail in Taiwan. The purpose of the study is to identify interaction problems that ticket buyers come across when they buy tickets via TVMs. Naturalistic observation and focus group interview are two methods used on collecting the interaction problems. The analysis of the identified problems generates several design suggestions which can be useful to the design of self service kiosks, especially TVMs. The design of future TVMs should support the various need of ticket buyers.

Keywords: ticket vending machine, observation, focus groups, kiosk

1 INTRODUCTION

Providing shorter waiting time for users and less man power for service providers, self service kiosks has been broadly used today, Automatic Teller Machine (ATM) is a good example. Except the deployment of ATM in banking, public transportation is another pioneer who introducing self service kiosk for selling tickets. Although Ticket Vending Machines (TVMs) have been used for many years, it is common to see a long queue at ticket counters in train/bus stations. According to the official statistics of Taiwan High Speed Rail Company, only 35% of tickets are sold via TVM. It is very interesting to know what bother these people in using TVMs, so that they would like to spend their precious time on queuing rather to challenge buying ticket from TVM.

This paper demonstrates a user centered study conducting at Taiwan High Speed Rail (THSR) in Taiwan. The purpose of the study is to identify interaction problems that ticket buyers come across when they buy tickets via TVMs. Naturalistic observation and focus group interview are two methods used on collecting the interaction problems. The analysis of the identified problems generates several design suggestions which can be useful to the design of self service kiosks, especially TVMs.

The Taiwan High Speed Rail Company provides passenger transportation service between two main cities, Taipei and Kaohsiung. The service links the two cities at a total length of 345 km with 90 minutes traveling time. THSR put into her first operation in March 2007. There are eight stations of the high speed rail is operated, namely, Taipei, Banciao, Taoyuan, Hsinchu, Taichung, Chiayi, Tainan, and Zuoying. This observation conducted in this study was at the THSR Zuoying station.

TVMs have been used in public transportation in Taiwan for more than 13 years. The first TVM was introduced by Taiwan Railway Administration (TRA) in 1996. TVMs have been also used by Taipei Rapid Transit (TRT) in 1996, by THSR in 2007 and by Kaohsiung Rapid Transit System (KRTS) in 2008. Some studies have been made on TVMs in Taiwan but mainly on the interface of TVM. Lin (2001) conducted a usability study on the interface of the TVMs of TRA. Wang (2005) and Lee (2005) did usability study on the interface of the TVMs of TRT. The study of Wu (2008) focus on the touch screen of TVMs.

This study examines not only the interface but also the interaction between ticket buyer and TVM. The interface of TVM is only a part of the ticket buyer-TVM interaction. Verhoef (1988) and Zwaga (1988) pointed out the main problem bothering people on using TVMs is to figure out what is the right procedure, e.g. destination first or coin first. Zwaga (1988) also showed an obvious difference on the time spending on getting a ticket between the experience buyer (17 seconds) and the novice buyer (200 seconds). Hisamune and Kishida (1997) conducted observation studies on three types of TVMs of Japan Rail East. Time spend, mistake and hesitating were recorded and analyzed. Connell (1998) did a study on errors made by ticket buyers on using TVMs. Hisamune and Kishida (1998) made a study on the experience, preference and frequency of ticket buyers to identify their references of ticket buying decision making. Connell et al (2004) took cognitive walkthrough method to identify interaction problems on using TVMs.

2 METHODS

Two methods, observation and focus group interview, were utilized in this study to collect the interaction problems that ticket buyers have during the use of TVMs. The two complementary methods are usually adopted at a same time.

2.1 *Observation*

The naturalistic observation method was adopted for collecting the first hand interaction problems during the use of ticket vending machine. The observation took place at the southern terminal (Zuoying) station of THSR in Taiwan. At the station, there are totally twenty machines installed. Five machines make a row. There two rows installed separately at the both side of the main entrance. Observations were done evenly between the four rows of the machines.

An observation team consists of two observers. They are equipped with a digital video recorder and a notebook. One observer is in charge of shooting the whole process of ticket buying. Another observer is responsible for taking note of problems which ticket buyers come across during ticket buying. All observations are conducted under the agreement of ticket buyers.

Each observation is end with a quick interview for collecting ticket buyers' subjective data. The contents of a quick interview depend on the time available and the will of each ticket buyer. The subjective data includes what problem they come across, why they run into the problem and how they resolve the problem.

There are totally twenty observations were made. The twenty participants consist of nine females and eleven males. They are age between 18 to 45 years old (average 28.8 years old). The observation took place in January, 2009. The day of the week of the observations are made are: three on Sunday, six on Tuesday, seven on Thursday and four on Thursday.

The videos shot at the observations were analyzed by using a qualitative analysis software called Noldus Observer XT.

2.2 *Focus group interview*

The focus group interview was also used for collecting the interaction problems during the use of ticket vending machine. Low cost, short time and rich interaction are characteristic of focus group interview.

In the focus group interview, a moderator is in charge of the flow of discussion and two assistants are responsible for taking note and video recording. The pictures and the films took at the observation study were used as reminders during the interview. A list of questions for guiding the discussion was prepared in advance. The discussion of the focus group interview was after turned into verbatim transcription.

There are seven participants, two females and five males, in the focus group interview. They are aging from 22 to 27 years old (average 24.6 years old). The participants were recruited at a local bulletin board system and were given an NT$500 worth reward for their participation. They all have the experience of using THSR TVMs.

3 RESULTS

The analysis of the data collecting from the observation and the focus group interview generates rich and fruitful findings. Below are some of the key findings.

3.1 *Not all ticket buyers are well prepared*

The current design of THSR TVMs, like many other TVMs, asks a ticket buyer to select "date, time, departure/arrival station, number of tickets, type of ticket" and enter the information according to the instructions on the screen. It also means that all ticket buyers must know what date, time, departure/arrival station, number of tickets and type of ticket they want to buy before they start to use TVMs. From the results of both the observation and the focus group interview studies, many ticket buyers do not have all the information they need.

3.2 *Some ticket buyers are active*

From both the observation and the focus group interview studies, active ticket buying behavior is identified. Some ticket buyers go back and forth several times at the different pages of the TVMs. They are not missing in following ticketing procedure. In fact, they know very well where they are in the procedure. They go back and forth at the pages for collecting information which they need to find a train which meet their specific goal, e.g. earlier arriving time, cheaper price.

3.3 *Give me backward not cancel*

Almost all participants make complaint against the missing of backward key. When ticket buyers mistakenly follow the ticketing procedure, there is no backward key but cancel key. They are willing backward to the previous step, so that they can correct their mistake. However, a cancel key will lead they go far back to the first beginning. They have to do all the selection once again.

3.4 *Card only, no cash*

All THSR TVMs accepts cash, credit card and ATM card for payment, but some TVMs are credit card and ATM card only. These credit card and ATM card only machines are differentiated from a sign at the top and the lack of billing device in the front (see Figure 1). Some participants mistakenly choice the credit card and ATM card only machine and notice the mistake when their operation reaches the last step, payment. They explain that they pay no attention to the sign and the lack of billing device but the touch screen at the center of the machine. The big screen attracts all their attention.

3.5 *Separated credit card paying display*

It is common that people pause for a while when they select to pay by credit card or ATM card. They still look at the main touch screen and wait for the next instruction. However, the paying information is not shown on the main touch screen but on another small display which locates on the top of the main touch screen (see Figure 1). It is also strange that the input of card payment is via a keypad located between the main touch screen and the small card paying display rather via the main touch screen.

Figure 1. The layout of a ticket vending machine which accepts credit card, ATM card and cash installed at THSR station.

3.6 *Ungrouping*

Some participants make a complaint about the forgetting of taking back their credit/ATM card. The slot for credit/ATM card locates beside the keypad which is between the main touch screen and small payment display. When the payment is accepted, people move their attention from the small payment display to the main touch screen for information. The information tells ticket buyers they can collect their ticket at the ticket slot which locates at the bottom of the main touch screen. Then they follow the instruction moving to pick up their ticket but forget taking back their cards.

4 SUMMARY

From the results of this study, it is no surprise that many travelers would like to spend their precise time queuing at ticket counter rather than challenge TVMs. When using THSR TVMs, ticket buyers are asked to answer a series of questions passively. If they fail to answer to the question, they fail to get their ticket. Nakamura (2004) pointed out the next generation of TVM is to support the various need of ticket buyers.

The application of supporting the various need of ticket buyer can be easily seen at the ticket counters. Some ticket buyers directly tell the ticket clerk a train number. Some ticket buyers express the information they have to the ticket clerk and ask for their suggestion on the rest information they do not have. Even some ticket buyers tell ticket clerk their criteria and ask them to filter out trains.

The use of large touch screen provides a direct manipulation interface. All information should be displayed and all data should be entered only at the large touch screen. The screen can also be used to point the location of card slot, coin slot and ticket slot. Figures and symbols instruction are much friendly than text instruction.

REFERENCES

Baber, C. & Parker, A. 1994. Buying a ticket on the under-ground, In *Proceedings of the Ergonomics Society 1994 Annual Conference*, University of Warwick.

Connell, I. 1998. Error analysis of ticket vending machines: comparing analytic and empirical data. *Ergonomics* 41(7): 927–961.

Connell, I., Blandford, A. & Green, T. 2004. CASSM and cognitive walkthrough: usability issues with ticket vending machines. *Behaviour and Information Technology* 23(5): 307–320.

Hisamune, S. & Kishida, K. 1997. A study on human behaviors observed in using ticket vending machines. *Journal of Japan Industrial Management Association* 48(4): 183–189. (in Japanese).

Hisamune, S. & Kishida, K. 1998. The effect for decision making of using vending machine by error experience: A case study on vending machines. *Journal of Japan Industrial Management Association* 49(1): 52–58. (in Japanese).

Li, H. 2005. *An Inquiry into the Universal Design Principles to the Automatic Ticket Issuing Machines of Taipei Mass Rapid Transit System*. Master thesis, National Taipei University of Technology, Taiwan. (in Chinese).

Nakamura, K. 2004. Transition in railway ticket vending machines: Greater convenience to passengers. *JR EAST Technical Review* 3: 15–18.

Verhoef, L. 1998. Decision making of vending machine users. *Applied Ergonomics*, 19(2): 103–109.

Wang, M. 2005. *An Interface Design Study of Automatic Ticket Machine for the Mass Transportation*. Master thesis, Tunghai University, Taiwan. (in Chinese).

Wu, L. 2008. U*ser Interface Design of Touch Screens for Seat Reservation of Mass Transportation System*. Master thesis, Tatung University, Taiwan. (in Chinese).

Zwaga, H. 1988. Buying a ticket for the Washington Metro, In E.D. Megaw (ed), *Contemporary Ergonomics*, 273–278.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A study on the preference of finger operation for computer graphics on the touchscreen

Hsinfu Huang & Chiulin Chiang
Department of Multimedia and Entertainment Science, Southern Taiwan University, Taiwan, ROC

ABSTRACT: In recent years, touchscreen is used in application to many electronic products design and it is popular by consumers. The usability of the interface in these products would influence the users' operation procedure and even effect the satisfaction. In this study, the independent variable of this experiment is three different icon control gestures. They includes move, rotation and zoom controls. The move gesture includes one-finger swipe, two-finger tap and two-finger swipe. The rotation gesture includes one-finger rotate, two-finger rotate and lock one and one finger rotate. The zoom gesture includes one-finger double tap, lock one and one finger drag and two-finger zoom. The dependent variable is the most suitable gesture match and the subjective rates. After the subjects experienced control types of each gesture, they had to pick up an ideal icon touch control match. The result of the study expressed that the users prefer the one-finger swipe on move gesture. The result of this gesture is one finger swipe > one finger tap > two finger swipe. Their preferring rotation control is the lock one and one finger rotate gesture. Also, the result of the gesture is one finger rotate > lock one and one finger rotate > two finger rotate. In addition, users tend to use the two-finger zoom gesture on zoom control. The result of this gesture is lock one and one finger drag > two finger zoom > one finger double tap. Moreover, there are significant difference between one finger and two fingers in every operate style ($p < 0.01$). The gender of subjects did not influence the preference result ($p > 0.1$). This result could offer interface designer for reference to their interface operation behavior planning.

Keywords: touchscreen, human-computer interaction, finger, operational preference

1 INTRODUCTION

In recent years, the development of touch screen technique has created a function-touch technique of display in many electronic products. There were touch machines, for example the ATM (Automatic Teller Machine) in the banks, kiosks, or the automatic ticket sellers in the shops and the MRT (Mass Rapid Transportation) stations. Moreover, the progress of these touch technique also started the application to the interface (Sears and Shneiferman, 1992; Benko, 2006; Levin, 2007). In home and personal consuming parts of the market to electronic products like cell phones, e-books, PDA (Personal Digital Assistant), and so on, were used the technique on the touchscreen. The technique helped users to use relative electronic systems and products in much convenient way. Therefore, the reason why the touch technology became a popular function was that it matched as better as users' intuitiveness. At the beginning, people did anything with their hands. But the traditional interface style of computer input has made users be accustomed to use mouse and keyboard gradually. When the touchscreen technique came out, people used their fingers to operate the computer naturally and it replaced the traditional interface with mouse. In addition, the input device such as mouse and keyboard was not that important since someone had the touchscreen, which saved units of the computer. The advantages mentioned above improving the portability and convenience to the touchscreen. So in the recent years, whether touchscreen were widely applied in each electronic device made users operate easier and much intuitive or not, was an issue which discussed widely in ergonomic field.

The screen supported the saturation with the improving display and the touch technique. From former black and white display to nowadays screen with 16.7 million colors, was a case of progress on LCD (Liquid Crystal Display). This could also offer users view and handle complicated information. Additionally, users had to be re-accustomed the new interface on the touchscreen to gradually replace the traditional HMI (Human-Computer Interface) with cursor and buttons. The application of HMI on touchscreen needed more researchers to perform such ergonomic study exploring and verifying that which HCI style matches human intuitive much more. Lifting users' operation efficiency and effectiveness depends on high-class HCI design (Shneiderman, 1992; Fung, 2003; Dix, et al. 2004). But there are few relative studies of recent HCI of the touchscreen. Levin (2007) pointed that when people faced more and more application on the touch screen, they needed to emphasize on usability and ergonomic issues of the interface in the future. Therefore, by improving the hardware and device level, the study on ergonomic of the touch screen should also be paid attention on and be developed. This kind of HCI design including giving appropriate feedbacks to users on the touch screen and other studies of design principle for facilitating users' efficiency on their jobs, and this had also become necessary issues in the touch ergonomics.

Lao et al. designed a series of gestures for interaction table commanding. They defined two styles of gestures at the beginning in the study. The two styles were continuous and discrete contacts. The continuous contact included finger, palm, half-palm, fists and vertical hand. The discrete contact included two, three, four and five fingers. They also defined three different moving methods like pressing, tapping and dragging in motion type. They combined these gesture styles, and they researched the interaction purpose and motivation, for example, move, zoom, rotation, duplication and delete to match each control purpose to their appropriate gesture styles. Finally, they used the analysis and result in the study for touch-applied on table and PDA. This study was done followed three main control items: change, zoom and rotate to perform the gesture preference research. Koskinen et al. (2008) focused on the control table to finish a gesture preference study for the intelligent gesture control in the future. They focused on pressing, sliding and rotating touch style to perform the research. They also focused on the variables such as surface, touch continuous time, intensity, accuracy, speed, repeatability, horizontal, vertical, direction, match and form to perform the experimentation. The result of the study represented that subjects preferred to control smaller graphics during the dragging. Subjects also preferred to use one finger to control the graphics than two fingers. The relative analysis and result also presented that the subjects preferred to use easy and uncomplicated gestures. With single finger intuitiveness result to the subjects, Koskinen et al. also suggested that the designers should be concerned about the single and double finger use in the touch design. But we had much interested in the analysis result of the single and double finger research and we hoped to verify this result in our study again. Therefore, the experiment of this study was designed couples of gestures of single and double fingers in "switch", "rotation" and "zoom" control types. We hoped we could verify the result of single and double finger use preference from the subjects.

Shneiderman (1992) considered that the color quality of the touch graphics was not the main factor to affect users' usability as our former study result. Moreover, in the option to touch graphic colors, we also agreed with Shneiderman that the hue did not influence the usability to users. Therefore, for not influencing the option result, we used the same series of pictures as the touch graphics in each gesture experiment. In addition, our former study also presented that the size of graphics were likely to influence the usability on LCD. This experiment used LCD to and finished the touch experiment. Therefore, we chose the graphic size as 500×500 pixels in the center of the 1600×900 pixels screen. The reason why we chose this proportion of size was that it was controlled the graphics freely and conveniently by the subjects because of enough room and size, and it could make the subjects not feel frustrated.

2 RESEARCH METHOD

2.1 *Subjects*

This study invited 60 university students to join this experiment. The age was from 18 to 31 years old (Mean = 22.05; SD = 3.15). They have normal vision and they are right-handed users. They can move their arms, wrists and fingers freely. Moreover, the subjects of this study were divided into three groups by touch experience, those three groups were types of "inexperienced", "not often used" and "often used" and there were 20 subjects in each group.

2.2 *Experiment design*

This experiment was a subject width design. The subjects had to operate three different touch control items, those items are "graphic switch", "graphic rotation" and "graphic zoom". And there were different gestures in each control items. On the other hand, there were three variables for the subjects in each control item. When the subjects performed the experiment, they had to experience nine different gestures with controlling the graphics on the center of the screen. The size of those graphics was 500×500 pixels. The set of the size because it could be appropriate on the center of the screen and could be controlled conveniently. The touch graphic in this experiment was square because we did not want the size and the shape influenced the experience of the subjects. Also, it could not influence the preference result.

In the "Graphic Switch" experiment, there were 17 different pictures presented on the graphic which on the screen. After the subjects done the "switch" gestures, the graphic would change to the next (or the last) pictures. There were three variables of the gestures in the experiment, they were "one-finger swipe", "one-finger tap" and "two-fingers swipe". The "one-finger swipe" was putting one finger (usually the index) on the touchscreen and moved it toward right or left then put the fingers left the screen, and the gesture finished; "One-finger tap" was to use one finger to tap on the left or right side of the screen smoothly; "Two-finger swipe" was similar to the "one-finger swipe". It put two fingers (usually be index and middle fingers) on the touchscreen, then swiped them to left or right and put the fingers left from the screen.

In the "Graphic Rotation" experiment, the subjects had to follow different gestures to drag the graphic to one curve direction to rotate the graphic. The variables as well as three gestures in the "Graphic Rotation" were "one-finger rotate", "two-finger rotate" and "lock one and one finger rotate". The "one-finger rotate" was to put one finger on the graphic of the center on screen, and dragged it to clockwise or counterclockwise direction, and the graphic will rotate to the same direction with the same angle as finger did; the "two-finger rotate" was that putting left and right index on the graphic at the same time, and dragged it to the same direction; The "lock one and one finger rotate" was that putting one finger on the graphic to fix it, and used another finger to perform the rotation. It also means that to drag the finger to clockwise or counterclockwise direction in curve move, and the graphic will rotate.

The third control item was the "Graphic Zoom". There were also three different gestures in this control item, they were "one finger double tap", "Lock one and one finger drag" and "two-finger zoom". The "one finger double tap" was to use the index finger to do the double tap on the graphic of the center on the screen. In this gesture, the graphic was set two kinds of size. When the subjects entered the experiment interface, the graphic was set in small size. When the subjects used one finger to tap on the graphic in twice, then the graphic will become the large size. If the subjects did the same action at that time, then the graphic will change to the smaller one as the beginning; the "lock one and one finger drag" was to use one finger fixing on the graphic, and use another finger to drag the graphic to zoom it; The "two finger zoom" was to press two fingers on the graphics and drag them to the same direction (to make the graphic smaller) and the different directions (to make the graphic larger) to perform the zoom gesture. The description of the nine gestures mentioned above showed in Table 1.

2.3 *Materials and stimulation*

All of the graphics during this experiment were set on the center of the screen. The graphic size was 500 × 500 pixels. The graphic was presented with 17 campus pictures orderly. The graphic was represented on the 1600 × 900 pixels screen.

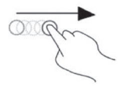
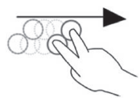
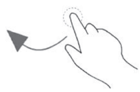
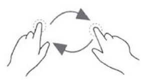
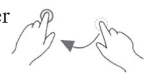
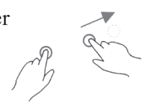
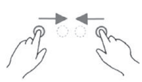
This experiment was programmed with the Flash actionscript 3.0, and the experiment was performed in the Windows 7 operation system. The factors of this experiment were not only the experience on the touchscreen and gender of the subjects. The main independent variables were the three control items: "switch", "rotate" and "zoom" which controlled with the graphic by subjects on the touchscreen. Moreover, the dependent variables were the subject rates and appropriate gesture types. This experiment was executed in an ergonomic laboratory of university, was used Touch Smart 300 touch tablet (HP IQ300-1038 W) as the operation device. As showed in fig. 1, the touch tablet was a 20 inches LCD, the height was 43.3 cm and the width was 27.1 cm, the resolution was 1600 × 900 pixels. The computer was set with the Windows 7 operation system and it was on a table which about 70 cm tall. The lab also offered an adjustable seat, could be adjusted to the most comfort angle when the subjects performed the experiment. The illumination of the lab was 500 Lux.

2.4 *Experiment tasks*

This experiment was divided into two parts and the first one was the experience task. After the subjects filled in their information about experience of touchscreen use, then they entered the experience task. The destination of the experience task was to let the subjects experience the control with using different gestures. The subjects had to sit on the computer and followed the instruction to use one or two hands to execute the gesture experience. There were three different gestures in each control item. Before the subjects finished the experiencing the three different types of gesture and generated the preference result, they could go back to experience until they find the answer of the gesture experience.

This study used the Lickerts' Scale with five-point rating to record the subject rating. The subjects wrote down the rate of 1 to 5 points after they experienced each gesture. The point "1" was the subjects feel the least appropriate and satisfied one to the control item rates; the point "5" was the most appropriate and satisfied one to the control item rates. The notice was that the subjects entered the part of questionnaire to rate the points

Table 1. The description of the nine gestures.

Control manner	Gesture name	Gesture model	Description
Type 01: Switch	1-1 One-finger swipe		• Using a finger to swipe from side to side on the screen for changing pictures.
	1-2 One-finger tap		• Using a finger to tap on the screen for changing pictures.
	1-3 Two-finger swipe		• Using two fingers to swipe from side to side on the screen for changing pictures.
Type 02: Rotate	2-1 One-finger rotate		• Using a finger to drag the graphic which at the center of screen and to move it in curve angle for rotating the picture.
	2-2 Two-finger rotate		• Using two fingers to drag the graphic which at the center of screen and to move them in curve angle at the same time for rotating the picture.
	2-3 Lock one and one finger rotate		• Using a finger to press and stay fix on the graphic which at the center of screen, while using the other one to drag the graphic and to move it in curve angle for rotating the picture.
Type 03: Zoom in/out	3-1 One finger double tap		• Using a finger to tap on the graphic which at the center of screen in twice for zooming the picture.
	3-2 Lock one and one finger drag		• Using a finger to press and stay fix on the graphic which at the center of screen, while using the other one to drag the graphic and to move it for zooming the picture
	3-3 Two finger zoom		• Using two fingers to drag the graphic which at the center of screen and to move them in different direction at the same time for zooming the picture.

after they had experienced three gestures of each control item. In this part, after the subjects finished the rate to the three gestures on the questionnaire, they could turn back to experience again and again until they got the difference between those gestures. The part we emphasized to let the subjects experience the gestures in easy condition to yield subject feeling and answer.

2.5 *Experiment process*

Before the experiment stated, we told the content and process instruction of this experiment to the subjects first. Then as the fig. 2 presented, the subjects were invited to sit in front to the computer to point at "graphic switch", "graphic rotate" and "graphic zoom" icons to enter the experience

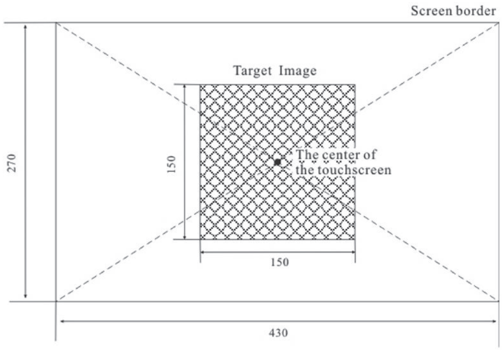


Figure 1. The experiment interface in this study (unit: mm).

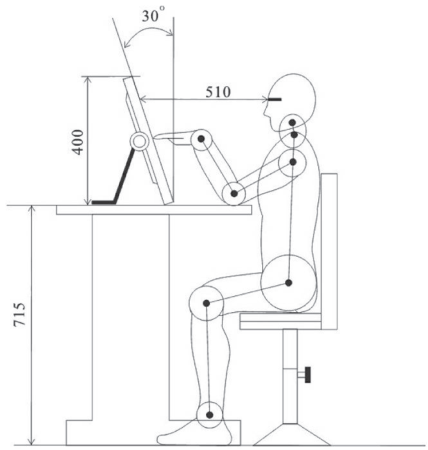


Figure 2. The experiment condition (unit: mm).

interface. When they entered the experience interface, they controlled the 500 × 500 pixels graphic to experience the gestures. There were three different types of gesture in each control. When the subjects finished experiencing an item, they had to give the rating point after experiencing. They gave those gestures the subject evaluation to the satisfaction and filled the result on the questionnaire with 1 to 5 points. 1 is the least and the 5 is the most appropriate and satisfied gestures.

The subjects could turn back to experience during filling the questionnaire to experience the difference between three gestures entirely. So the process of this experiment was: 1) Filling information. 2) Experience of graphic switch gesture. 3) Rating points for graphic switch gesture. 4) Experience of graphic rotate gesture. 5) Rating points for graphic rotate gesture. 6) Experience of graphic zoom gesture. 7) Rating points for graphic zooming gesture. 8) End the experiment. After experiencing and rating "graphic switch", "graphic rotate" and "graphic zoom" three gesture items, this experiment ended. The experiment took about 10 minutes.

3 RESULT ANALYSIS

3.1 *The preference of gesture type by subjects*

In the descriptive analysis we could discover that the subjects preferred the "one finger swipe" gesture (1-1) in the Type 1 (switch gesture); the Type 2 (rotate gesture) was preferred the "one finger rotate" gesture (2-1); and the Type 3 (zoom gesture) was preferred the "lock one and one finger drag" gesture (3-2) as shown in the Table 2. Except for this, there are not significant difference between each type of gesture ($p < 0.01$).

3.2 *The difference between the different experiences of subjects*

The result of the ANOVA Analysis represented that the three different type of subjects did not influence the preference to the gesture ($p > 0.5$) as the Table 3 showed. The Figure 3 showed the mean points of the preference to the gestures of

Table 2. Descriptive analysis to the nine gestures.

Type	Action	N	Mean	Std.
Switch	1-1 One-finger swipe	60	4.37	.780
	1-2 One-finger tap	60	3.83	1.060
	1-3 Two-finger swipe	60	1.97	.920
	Total	180	3.39	1.384
Rotate	2-1 One-finger rotate	60	4.22	.976
	2-2 Two-finger rotate	60	2.38	1.151
	2-3 Lock one and one finger rotate	60	3.42	1.306
	Total	180	3.34	1.371
Zoom in/out	3-1 One finger double tap	60	2.67	1.003
	3-2 Lock one and one finger drag	60	4.07	.972
	3-3 Two finger zoom	60	3.88	1.223
	Total	180	3.54	1.234
Total	1-1 One-finger swipe	60	4.37	.780
	1-2 One-finger tap	60	3.83	1.060
	1-3 Two-finger swipe	60	1.97	.920
	2-1 One-finger rotate	60	4.22	.976
	2-2 Two-finger rotate	60	2.38	1.151
	2-3 Lock one and one finger rotate	60	3.42	1.306
	3-1 One finger double tap	60	2.67	1.003
	3-2 Lock one and one finger drag	60	4.07	.972
	3-3 Two finger zoom	60	3.88	1.223
	Total	540	3.42	1.332

Table 3. Descriptive analysis to experience.

Subject experience	F	Sig.
Often use	.401	.671
Not often use	.391	.677
Never use	.177	.838

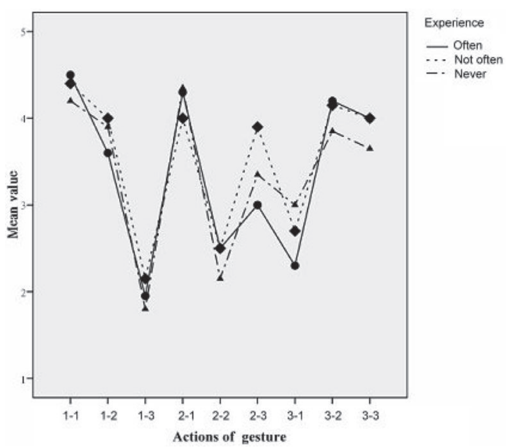


Figure 3. Mean value to the nine gestures.

"often use", "not often use" and "never use" three types, significantly, the result of the preference trend between these subjects are similar. Moreover, if we divide those subjects into "ever use" and "never use" two groups, the difference by experience also showed not significant result ($F = 0.568$, $p > 0.1$).

3.3 *The compare between one finger and two finger operation*

The independent variables of this experiment (9 type of gesture) could be divided into single finger and double fingers types, and there were not significant difference between these two types of preference ($p < 0.01$). In this part, the most of the subjects preferred to use one finger to operate just like the table 4 represented, and the subjects also expressed that the manner of using two fingers to operate graphics is much difficult. Figure 4 presented the result of mean point between one finger and two fingers, it also presented the subjects of the different three types preferred to use one finger to operate the touchscreen.

3.4 *The influence to gesture preference by gender*

In this study, the result of each subjects' gender presented that there was no significant difference on every gesture type ($p > 0.1$), it also expressed that the gender did not influence the preference to the gesture.

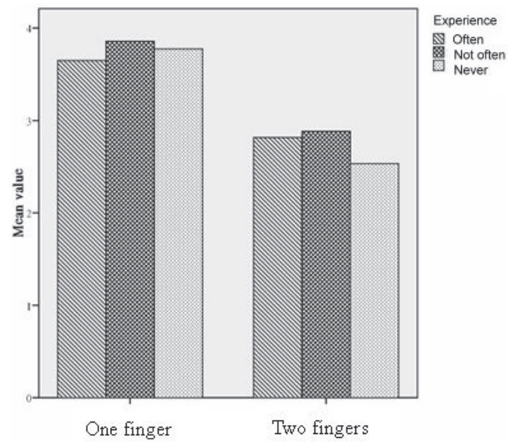


Figure 4. Mean value to the finger number.

Table 4. Descriptive analysis to finger number.

Type	N	Mean	Std.
One finger	360	3.76	1.172
Two fingers	180	2.74	1.375
Total	54	3.42	1.332

4 DISCUSSION

4.1 *Gesture preference of graphic switch*

In the experiment of the preference of graphic switch gesture, the subjects gave the first gesture item-"one finger swipe" the highest rates. This gesture was also the graphic switch gesture which used often on cell phones, PDA or other electronic products. Moreover, the result indicated that the "one finger tap" gesture was gave the second place.

The "one finger swipe" just used one finger to control the graphic which just like the "one finger swipe" gesture, the subjects just tapped on the sides of the screen then it could change the pictures when using the " one finger tap" gesture, but it was not that intuitive than using the "one finger swipe" gesture. In this part we discussed with the action "reading". When we read general books and papers, readers had to turn the pages with hands toward left or right. Therefore, if we view the pictures on the screen as textbooks or comic books, when we need to turn the page, then we are accustomed to wave hands from left to right or from right to left to turning action. So, the "one finger swipe" gesture was like the action and it was appropriate to the intuitiveness when reading. Relatively, the "one finger tap" gesture was like the "click" action which used the mouse to click on the left or right side on the screen. For the computer users, the "one finger tap" gesture was suitable for the intuitiveness when using mouse. But using mouse is still not that intuitive to human than using hands, therefore, the rates of the "one finger tap" gesture is lesser than the "one finger swipe" is.

Moreover, in the compare between the "one finger swipe" and the "two finger swipe" gestures, although these two kinds of gestures were execute in the similar way, the result of Koskinen et al. (2008) presented that the subjects preferred to use one finger to perform on the touchscreen. Also, the result of this study presented that using one finger were much appropriate to the preference and intuitiveness of the subjects just like Koskinen et al.'s research.

4.2 *Gesture preference of the graphic rotate*

In the experiment of graphic rotate gesture, the "one finger rotate" were rates the highest by the subjects. This result verified the results which presented that subjects would prefer one fingers to use the touchscreen in many former studies. Otherwise, when users used two fingers to rotate the graphic, then the graphic cannot turn around in 360 degrees. The subjects had to turn their arms back when they rotated their forearms to 180 or 270 degrees, then they just could turn the graphic to 360 degrees. This part presented that the subjects would use their ulna of forearm to rotate when they used double fingers to rotate the graphic. The ulna and the radial around the wrist do not have the joint or muscle which can act freely to turn 360 degrees just like each finger, so it is easier to hurt the arms or hands. According to Burgess-Limerick et al. (1998) presented that if an user over-extended wrist angle or twist their ulna when used touch device, then it would make muscular uncomfortable. So we guessed that the reason why the "one finger rotate" gesture got the higher rates than the two fingers did is the subjects can turn the graphic to 360 degrees at once.

4.3 *Gesture preference of the graphic zoom*

In the experiment of the graphic zoom gesture, the subjects gave the second gesture-"Lock one and one finger drag" the most points. Although the "Lock one and one finger drag" gesture was the highest rates, there were not much distance between it and the second rates gesture-"Two finger zoom". It expressed that the compare with "one finger double tap" gesture, most users preferred to use two fingers to execute zooming. Moreover, users had to control the vector of two sides (or two angles), they just can do the zooming action. Therefore, using two fingers are much appropriate and suitable for the intuitiveness than using one finger. Moreover, this study guessed that the "Lock one and one finger drag" is much power saving and stable (fixed vectors) than the "two fingers zoom", and it the reason why the subjects gave higher rates to the "Lock one and one finger drag".

4.4 *The influence of gesture preference by subject experience*

The result presented that there were no significant difference between the subjects who has often and seldom experience and no experience. This presented that the intuitiveness of the touch gesture preference to different experienced subjects are similar. The control type like graphic switch, rotate and zoom all referred former popular touch control style, so they are used in much intuitive. It perhaps that those control type all matches users' intuitive action, so whoever the users will feel the like "intuitiveness" feeling.

5 CONCLUSION

Before investigated the gesture preference, this study had observed many kind of touch gestures and then we summarized and designed the gesture of the experiment. The experiment result of this study was suitable for our estimate at the beginning. The estimate includes the preference of using single finger to do the graphic switch and rotation, and using two fingers to perform the graphic zoom. These studies verified the former study result of the gesture preference on the touchscreen, the result was also similar to nowadays custom preference.

Moreover, this study also investigated the difference between the touch gesture preference divided by the experience. The result represented that the touch experience do not influence the preference of gesture. On the other hand, when users used the touchscreen, their former experience would do not affect the judge of "intuitiveness". This result could be applied in the developed project for different experience of groups. Finally, this study just investigated for the preference index of the subjects. In the future, the relative experiment could focus on the error and efficiency detect of the different gestures on tasks or jobs.

ACKNOWLEDGEMENT

This study was partially supported by the National Science Council of Taiwan, R.O.C. under Grant no. 98-2218-E-218-003 NSC.

REFERENCES

Benko, Hrvoje, Wilson, Andrew D., Baudisch & Patrick, 2006, Precise selection techniques for multi-touch screens. In *CHI 2006 Proceedings on Interacting with Large Surfaces*, 1263–1272.

Burgess-Limerick, R., Shemmell, J., Scadden, R. & Plooy, A.1998. *Wrist posture during computer pointing device use*. Australia.

Dix, A.J., Finley, J., Abowd, G.D. & Beale, R. 2004. *Human-Computer Interaction.* Prentice Hall.

Epps, J., Lichman, S. & Wu, M. 2008. *A study of hand shape use in tabletop gesture interaction.* ACM: New York, NY, USA.

Lao, S.Y. et al. 2009. *A gestrual interaction design model for multi-touch displays.* British Computer Society: Swinton, UK.

Levin, M. 2007. Integrate tactile feedback into touch-screen HMI. *EE Times-Asia*, 1–3.

Shneiderman, B. 1992. *Designing the user interface: Strategies for effective human-computer interaction.* 2nd edition. Addison-Wesley: MA.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Effect of static, dynamic and video media on users' visual behavior in the Internet

Eric Min-yang Wang & Wan-ju Hsu
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan, ROC

ABSTRACT: Although multimedia technique provides more design options to web page designers, it is still require to understand how users' behavior affected by using multimedia technique in the Internet. Because inappropriate use of multimedia may confuses users and makes it harder for them to understand information. The purposes of this research are investigating and discussing effect of different combination of multimedia and different layout of webpage for information presenting on users' visual behavior. The results of this study can provide suggestions to web page designers and hope these suggestions can inspire designers with new design idea to create more user-friendly environment in the internet.

Keywords: human-computer interaction, web browsing, eye movement, multimedia, web page design

1 INTRODUCTION

According to the data on the Internet World States: Usage and Population Statics (Internet World Stats, 2009), the world's online population had reached 1.66 billion up to June 2009, and the growth rate of the online population was 362.3% from 2000 and 2009. Therefore, it is an explosive increase in the number of the Internet users all around the world, while it is also true for Taiwan. The report of the survey of the Internet use in Taiwan in the first quarter of 2009 showed that the Internet users had reached 10.57 million, with an Internet access and application popularity of 46%. With the increase of the online population and constant upgrade of the network technology, more information could be easily got from the Internet, therefore the means people using to search and get information is no longer limited to traditional communication media (such as newspapers and magazines, TV, broadcast and so on).

With the development of the technology, the Internet environment could provide users with large amount of diverse information. Apart from providing large amount of rich information like print media, it was also of the ability like rich media to represent dynamic text, pictures and even more realistic movie and music as well as static text. However, Nielsen (2000) pointed it out that although multimedia technology could provide more design choices for designers, further analyzes on its application were needed, or inappropriate use of the multimedia would confuse the users and make it difficult for them to understand the information it represents. Therefore, the understanding of how the users would be influenced by the website design before the website design work would not only provide web surfers with good experience, but also contribute to the effective communication between users and the web content (Landauer, 1995). In today's dissemination environment based on visual information, appropriate visual information design is able to ensure the information get across rapidly and accurately, at the same time it can even enhance the effectiveness of the information communication.

This research would explore the influences of different information presentations on users' visual behavior based on the eye movement data provided by eye tracking device. Meanwhile, through observing users' behavior, the researches try to find out whether users would show some specific behavioral patterns when using different medium and surfing such information as presented in different layouts and how to design can reach the purpose of promoting the more effective communication between users and the web content providers, that is, making users gain the information more rapidly and accurately. In brief, the goal of this study is to discuss the effect of different media combination for presenting on users' visual behavior and the effect of different layouts on the users' visual behavior.

2 LITERATURE REVIEW

2.1 *Multimedia web page*

"Media" was referred to a kind of channel to convey information, while multimedia was a combination of multiple and media, meaning such a media as combining with multiple information communication mediums. The concept of the multimedia means it was a project in essence, which was formed through any combination of the text, graphic art, sound, animation and video made by computer. This study would focus on discussing information presenting of web page in visual. Therefore it was based on the literature review and classified all kinds of media into three categories. The three categories were static media, dynamic media and video media. Static media included text, graph, static photograph, chart and figure. Dynamic media was composed of several pictures and continuously display these pictures. Video media was a continuously video clip with audio.

2.2 *Visual behavior*

When observing visual behavior, eye tracking should be often used to record eye movement data. The recorded eye movement data could be classified into three types: Fixation, Saccade and Scanpath. The detailed description is shown as follows:

1. Fixation: it meant when the eyeball keeps on a subject, the attention focused on it (Rayner, 1995; quoting from Josephson and Holmes, 2002).
2. Saccade: it meant the view sight move from a fixation to another fixation, and during this period, message processing was held back. (Stark and Ellis, 1981; quoting from Josephson and Holmes, 2002).
3. Scanpath: it was made up of a series of fixation and saccade (Noton and Stark 1971; quoting from Josephson and Holmes, 2002). It was conscious, related with the process of switching attention, higher level memory and cognition.

2.3 *Eye tracking related researches*

Rayner, Rotello, Stewart, Keir and Duffy (2001) found that watching print advertisements needed more time spent on the text but not pictures; while fixation duration and saccade length on the pictures were longer than that on the text. Because there were many fixation point in the text area, browser would not move between the text and the picture, but tend to read title first, then the content and pictures (although some browser would first scan pictures roughly). However, Rayner, Miller and Rotello (2008) modified this conclusion in the follow-up research, and they thought that watch sequence would be influenced by different purposes. In the said research, due to the experimental instruction requiring browser to watch the advertisement on the assumption of buying products, browser behavior appeared as first read the text; while in the new research which required browser to evaluate the quality or effectiveness of the advertisement, compared with the text area, pictures drew more attention and took more time of the browser. Meanwhile, Hyona, Lorch and Kaakinen (2002) also considered that when people scanned pictures or text, the visual behavior was determined by his personal strategy, knowledge and experience.

However, due to different media characteristics, even the same presentation design of both text and pictures would also lead to users' different visual behaviors. Berge, Collins and Dougherty (2000) found that, in the Internet environment, homepage writing was different from the paper writing, for when design the former, it was necessary to segment data into shorter paragraph or section, which would shorten the download time. Otherwise might be impatience and give up reading. Furthermore, the female was word-and accuracy-oriental and tended to accept text information carefully; while the male was picture-and loosing-oriental and would prefer to browse pictures rather than reading the words (Schiessl, Duda, Tholke and Fischer, 2002).

In respect of reading behavior for the website news, users would put the first-entry points in the text (Pan, et al., 2004). Tang and Jhuang (2005) had ever study the influence of photo location in the news on the reading sight and found that as for the symmetric pictures and texts layout design, different picture locations would bring about asymmetric attention distribution pattern. If the pictures were placed on the left of the text, browsers' eye sight leap span would be shorter and more fixation point would be appeared on the pictures, but with somewhat shorter gaze duration. On the contrary, if the pictures were placed on the right of the text, the eye sight leap would become larger, and lesser fixation point would be appeared on the pictures on right.

3 METHOD

The study was used experimental method, because the experimental method contributes to decrease unnecessary interference factors, build a causal relationship and understand different influences on the behavior caused by these factors. The study was mainly used an eye tracking device to record the users' eye movement data and had an interview with subjects to get a profound understanding the reason of the influence on users' visual behavior after experiment. The following would describe the subject, equipment, variables, experimental web pages, and experimental process.

3.1 *Subjects*

The report of Taiwan Network Information Center (TWNIC, 2008) showed that among 2036 broadband network users, those aged 16 to 25 who use broadband connections occupy the highest percentage (90.34%). On weekdays (from Monday to Friday), average hours online per day of 1 to 2 hours accounts for the largest proportion, 20.74%. On the weekend (from Saturday to Sunday), the proportion of the average hours online per day of 1 to 2 hours accounts for 13.93%, the highest percentage. Therefore, the 36 subject would be recruited by this research's interviewees are aged from 16 to 25, and they use the Internet each day average for more than 1 hour. Because the goal of this research is to discuss the users' visual behavior, the subject should be of normal eyesight or owe normal eyesight after correcting vision. Meanwhile, they must undergo necessary correcting process for the eye movement data recorded by eye tracking.

3.2 *Experimental equipment*

Equipments involved in this experiment included a eye tracking device (ASL Mobile Eye), and two personal computers (one was used for the subject to read experimental web page; another was used for researcher to observe and analyze eye movement data); software consist of Windows XP, Gaze Tracker (an eye movement data recording software) and Eye Vision (an eye movement data analysis software).

3.3 *Experimental variable*

There were two independent variables discussed in this experiment, both of which were within group design by the subject. Differential of multimedia combination was one of the independent factors. The three levels were: static media combination (texts and graph) dynamic media combination (texts and animation) and video media combination (texts and video). Another independent variable was differential of webpage layout. The two levels were: horizontal layout (as shown in Figure 1) and vertical layout (as shown in Figure 2).

3.4 *Experimental webpage*

Based on the independent variable, the experimental web pages would be designed by using

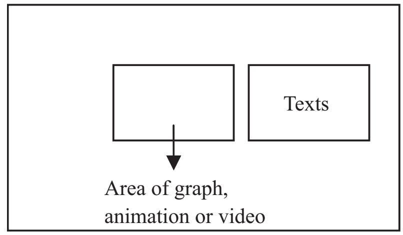


Figure 1. Horizontal layout.

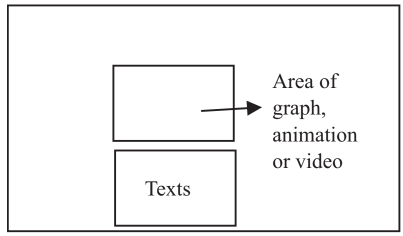


Figure 2. Vertical layout.

Dreamweaver. To reduce unnecessary disturbance, each web page would be designed in the standard format (as shown in Table 1).

3.5 *Experiment procedure*

The participants were welcomed and briefed about the experiment, which was followed by an explanation of the equipment to be used. It was explained that only the eye, voice and stimulus display would be recorded. Then researchers described the procedure and the aim of the experiment and guided the subject to fill in background questionnaire, which includes gender, age, and experience and usage of using the Internet. To avoid learning effects, the participants were randomly divided into several groups and different sequence for browsing webpage would also be assigned to each group. In this way, all the participants would be divided into six groups, with three male and three female in each group.

After calibration process of the eye tracking device finished, the participants started to brows the experimental web page. During the participants browsing the experimental webpage in the randomly arranged sequence, no disturbance should affect the subject's browsing and mood. When participants finished browsing the experimental web page, they would be interviewed. The interview included two parts: reading performance and explanation of their behavior. Detailed description is as follows:

1. Reading performance: the experimental in this part all consists of two single choices, one of which was designed based on the text content

Table 1. The standard format of the web page.

Content	Article: the content of each article was profile of Taiwanese, with 250 Chinese words. Graph: graphs were photos of Taiwanese, with the size of 360 pixels (wide) × 240 pixels (high). Animation: the animation was representing three photographs repeatedly, with a playback rate of 15 frames per second. And the size of the animation was 360 pixels (wide) × 240 pixels (high). Video: the content of the video was the profile, with the duration of 30 seconds and size of 360 pixels (wide) × 240 pixels (high).
Layout	The size of picture/animation/video would be the same with that of the text area, 360 pixels (wide) × 240 pixels (high). For the purpose of avoiding disturbance, no web page menu would be designed in it, but the reservation would simulate there is a menu in the left space. Users could browse all the webpage contents without using scroll bar.
Design	Font: all the content would be presented in Pmin gLiu, and the size of the title was 20 pt and 16 pt for the article. Color: background color would be white, and the font color was black.

of the page and another was designed based on the content of graph, animation, and video in the page. And the two choices were answered after the participants finished the webpage browse. The purpose of the experiment was to measure the reading performance of the participants after their browse. But in order to avoid that the participants would read the content and remember the information excessive carefully, the question was designed as single choice as for the content, letting the subject can find out answer by recalling from the options.

2. Behavior explanation: in this stage, there would be an interview between researchers and the participant. The participant would be asked about the explanation behavior of browsing the experimental webpage, for the purpose of gaining the subjective explanation for the subject's visual behavior to cover the shortage of eye movement data's objective exploitation. The content of the interview would be focused on the attention distribution to the picture and text, which was the fixation duration and scan-path to ask questions. And the questions can be divided into two dimensions: one is the user's description of their fixation duration and scan-path; the other is to explain the reason of their behaviors for users.

4 RESULTS AND DISCUSSION

4.1 *Participants*

There were 36 participants (18 males and 18 females) in this study and they are all college or graduate students of National Tsing Hua University. The average age of the participants is 21.78 year-old (std. = 2.87) and the average time of they using the Internet is 6.47 hours each day (std. = 3.33). The main three activities when they using the Internet is receiving e-mail (22.3%), browsing web pages (19.7%) and searching information (19.1%). Six participants were tested for the pilot study. Their right eye was calibrated for the eye tracking purposes.

4.2 *Difference analysis*

4.2.1 *Multimedia combination*

There are significant effects of different combination of media in the area of graph, animation, or video and texts on ratio of fixation duration and fixation counts. Table 2 is a summary of analysis of variance when the independent variable is multimedia combination.

Table 2 indicated that users would have different behavior when they browsing the web pages with different multimedia combination. There were three possible reasons. First, texts, graph, animation and video media have different levels of attracting user's attention and effect location and duration of first eye sight. It is also interfered users' attention when they look at another area. Second, different media could provide different levels of information volume. That would effect how users browse the web page and it could be found evidence from eye movement data such as duration of fixation and numbers of fixation. Third, texts, graph, animation, and video provide different levels of information quality. It was found that users tend to use the most efficient way to read the web page. For example, users would read one area carefully and then read another area when they browsing the web page designed with static media combination. But users would read both two areas when they browsing the web page designed with video media combination.

Table 2. A summary of analysis of variance (multimedia combination).

Dependent variable	Ratio of fixation duration: Area of graph, animation or video	Ratio of fixation duration: Area of texts	Fixation count: Area of graph, animation or video	Fixation count: Area of texts
Significant	yes	yes	yes	yes

Table 3. A summary of analysis of variance (layout).

Dependent variable	Ratio of fixation duration: Area of graph, animation or video	Ratio of fixation duration: Area of texts	Fixation count: Area of graph, animation or video	Fixation count: Area of texts
Significant	no	no	yes	no

4.2.2 *Layout*

There is a significant effect of different layout of web page on fixation counts in area of graph, animation or video. Table 3 is a summary of analysis of variance when the independent variable is layout of web page.

User would like to read content carefully at the beginning and continuously evaluate importance of the content. The importance of the content would effect users' decision whether they pay attention on this area. According to the results of the interview, users thought that they can concentrate more on reading one area when browsing the web page designed with vertical layout. In the web page with vertical layout, users would stop browsing the web page if they found that they were not interested in content of the web page. Meanwhile, users could read both two area at the same time when browsing the web page designed with horizontal layout.

4.3 *Analysis of scanpath*

The study was observed and summarized the eye movement record and description of participants' visual behavior from the interview. Table 4 shows that all scanpaths which participants had shown and lists the percentage of participants show this kind of scanpath when they browsing the experimental web page.

Compare to the web pages design with dynamic media and video media, users would concentrate on reading content of web page when they browsing the web page designed with static media. The behaviors of browsing the web page designed with static media combination could be classified two kinds of behaviors. The difference of these two behaviors is whether users looked at the graph again after reading the texts. The reason why some users looked at the graph again may be that they do not read the graph carefully at the first time and want to confirm and memorize information.

Users more concentrated on reading content when they browsing the web page with static media. In this kind of web pages, users can browse the web page more easily because static graph would not change the presenting information and interfere with users' attention. Users' visual behavior would be influenced by animation when browsing the web page with animation media design. There are 34.72% participants looked at the area of animation when reading the area of texts. The possible reason is that users want to confirm whether animation display new information or not because they did not be informed that the animation represents three photographs repeatedly. And also the animation repeatedly displaying different photographs may annoy and disturb users.

Users' visual behaviors would be affected by the content of video clip when browsing the web page designed with video media. About twenty-nine point seventeen percent of participants read both two areas at the same time and their scanpaths go back and forth between two areas. The main factor of deciding users' behavior is importance of the content on each area. Though users also showed similar behaviors in the web page design with dynamic media, users pay more attention on video. The possible reasons are that video media could provide more information than dynamic media and video media needs more time for presenting all the information. Therefore the scanpath of the users go back and forth between two areas if users have time pressure or want to browse web page more efficient.

Table 4. Description of participants' scanpath.

Variable	Description of participants' scanpath
Static media combination	The participants glanced over the graph and then read the texts. After reading the texts, they looked at the graph again. (45.83%)
	The participants looked at the graph and then read the texts (54.17%)
Dynamic media combination	The participants looked at the animation and then read the texts. (38.89%)
	The participants looked at the animation and then read the texts. After reading the texts, they looked at the animation again. (26.39%)
	The participants looked at the animation and then read the texts. When reading the texts, they looked at the animation several times. (6.94%)
	The participants looked at the animation and then read the texts. When reading the texts, they looked at the animation several times. After reading the texts, they looked at the animation again. (27.78%)
Video media combination	The participants watched the video and then read the texts. (54.17%)
	The participants watched the video first. Then they read the texts when watching the video. (29.17%)
	The participants read the texts and then watched the video. (16.67%)
Horizontal layout	The participants looked at area of the graph, animation or video in the left and then looked at area of the texts in the right. (49.07%)
	The participants looked at area of the graph, animation or video in the left and then looked at area of the texts in the right. They looked at area of the graph, animation or video in the left again when finished reading. (27.04%)
	The participants read area of texts in the right and then looked at area of the graph, animation or video in the left. (5.56%)
	The participants looked at area of the graph, animation or video in the left and then looked at area of the texts in the right. They looked at area of the graph, animation or video in the left several times while reading the texts. (12.04%)
	The participants looked at area of the graph, animation or video in the left and then looked at area of the texts in the right. They looked at area of the graph, animation or video in the left several times while reading the texts. After reading the texts in the right, they looked at area of the graph, animation or video in the left again. (9.26%)
Vertical layout	The participants looked at area of the graph, animation or video in the top and then looked at area of the texts in the bottom. (49.07%)
	The participants looked at area of the graph, animation or video in the top and then looked at area of the texts in the bottom. They looked at area of the graph, animation or video in the top again when finished reading. (27.04%)

(*Continued*)

Table 4. (*Continued*).

Variable	Description of participants' scanpath
	The participants read area of texts in the top and then looked at area of the graph, animation or video in the bottom. (5.56%)
	The participants looked at area of the graph, animation or video in the top and then looked at area of the texts in the bottom. They looked at area of the graph, animation or video in the top several times while reading the texts. (12.04%)
	The participants looked at area of the graph, animation or video in the top and then looked at area of the texts in the bottom. They looked at area of the graph, animation or video in the top several times while reading the texts. After reading the texts in the bottom, they looked at area of the graph, animation or video in the top again. (9.26%)

5 CONCLUSION

In this study, it was discussed the effect of using different web page design on users' visual behavior and investigate the reasons behind users' behavior. The results can inspire designers with new design idea to create more user-friendly environment on the internet. Moreover, this study can be also applied to not only webpage design but also the area which multimedia technology is used for transmitting information such as software interface design and educational material design.

REFERENCES

Berge, Z. Collins, M. & Dougherty, K. 2000. Instructional and cognitive impacts of web-based education. In B. Abbey (Ed.) *Design guidelines for web-based courses:* 32–40. Idea group.

Hyona, J. Lorch, R.F. & Kaakinen, J.K. 2002. Individual differences in rading t smmarize epository txt: eidence from ee fxation ptterns. *Journal of Educational Psychology* 94(1): 44–55.

Internet World Stats 2009. Internet usage statistics: Usage and population statistics. Retrieved 09/30/2009 from *http://www.internetworldstats.com*

Josephson, S. & Holmes, M.E. 2002. Visual attention to repeated internet images: Testing the scanpath theory on the world wide web. In *Proceedings of the 2002 symposium on Eye tracking research & applications*.

Landauer, T.K. 1995. *The trouble with computers: usefulness, usability, and productivity. MIT Press*.

Pan, B. Hembrooke, H.A. Gay, G.K. Granka, L.A. Feusner, M.K. & Newman, J.K. 2004. The determinants of web page viewing behavior: An eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications.*

Rayner, K. Miller, B. & Rotello, C.M. 2008. Eye movements when looking at print advertisements: the goal of the viewer matters. *Applied Cognitive Psychology* 22(5): 697–707.

Rayner, K. Rotello, C.M. Stewart, A.J. Keir, J. & Duffy, S.A. 2001. Integrating text and pictorial information: eye movements when looking at print advertisements. *Journal of Experimental Psychology-Applied* 7(3): 219–226.

Schiessl, M., Duda, S., Tholke, A. & Fischer, R. 2002. Eye tracking and its application in usability & media research. Retrieved 11/10/2009, retrieved from *http://eye-square.com/documents/Eye Tracking-ResearchApplications.pdf*

Shuman, J.E. 1998. *Multimedia in action.* Belmont CA: Wadsworth.

TNIC 2008. A survey of Taiwan broadband internet user in 2008. Taiwan Network Information Center.

Tang, D.L. & Jhuang, S.J. 2005. Exploring attentional effect of image position on news reading from eye-tracking method. *The Journal of Advertising & Public Relations* 24: 89–104.

Xu, Y.L. 2001. *A study on effect of using multimedia to present on learners' cognitive load and learning performance*. National Central University.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

An analysis of the effects of personal and design factors on comprehensibility of safety signs

Ka Lee Chan, Ke Chen & Alan H.S. Chan
City University of Hong Kong, Hong Kong

ABSTRACT: This study investigated the effects of personal and design factors on comprehensibility of safety signs. A survey was conducted to capture participants' particulars, ratings on sign features, and comprehension scores. Sixty-two Hong Kong Chinese participants guessed the meanings and rated the sign features of 25 Hong Kong safety signs. The results found that driving experience was a significant predictor of sign comprehensibility. Contrary to expectation, age group, gender, education background, period of time living in Hong Kong, occupational experience, safety professional background, and safety education background had no effect on comprehension performance. Sign familiarity, concreteness, and meaningfulness were correlated with comprehension score, whereas sign simplicity was not. The results of this study provide useful information and recommendations for the design of more user-friendly safety signs.

Keywords: safety signs, comprehensibility, sign feature, icon design, personal factor

1 INTRODUCTION

In Hong Kong, there were 14,932 industrial accidents in 2008 (Hong Kong Occupational Safety Health Council 2009). According to Whittingham (2004), the cause of about 80% of all accidents can be attributed to human error. Safety signs are thus needed to alert people of hazards and to provide clear instructions on what to do so as to avoid or minimize the occurrence of undesirable or fatal consequences. In a safety sign study done by Tam et al. (2003), thirty safety signs and symbols were distributed to construction personnel working in eight different construction sites in Hong Kong. The results of this study showed that the comprehension level of many safety signs do not meet the ISO 3864-3: 2006 and ANSI Z535.3-2002 comprehension criteria, and that the effectiveness of many safety signs is low in conveying safety messages. The results indicated that substantial problems exist in comprehending the signs and symbols that are posted in construction sites. Redström (2006) had stated that the idea of design is to communicate effectively and it is important to make a given design easy to understand and interpret. Industrial designers should make a sign as clear and unambiguous as possible so that the target group could understand the intended message easily (Mono 1997). Poor-designed symbols could lead to injury due to the lack of understanding or misinterpretation (Raskin 2000). The study of Tam et al. (2003) highlighted the importance in studying the sign comprehensibility of Hong Kong safety signs. However, not much study addressing the issue of sign comprehensibility of Hong Kong safety signs was found in the literature.

Some researchers have studied the relationships between the understanding of safety signs and personal factors such as age (Hancock et al. 2004), disability (Duarte & Rebelo 2005), cultural background (Leung & Hellier 1998), drinking habit, involvement in on-site safety promotion activities and job position (Tam et al. 2003). It was shown that personal factors did affect the sign comprehension performance. In this study, the eight factors investigated were: age, gender, education background, occupational experience, period of time living in Hong Kong, driving experience, safety professional background, and safety education background. In this study, participant who was a registered safety supervisor, safety officer for construction site, or safety officer for the shipyards or container handling workplaces was classified as one with safety professional background. Participant who had taken accredited courses in "construction safety" or "occupational safety and health", and awarded with any post-graduated diploma, degree, higher diploma, diploma and certification was regarded as one with safety education background. It was hypothesized that except for the gender factor, all the other factors will affect the sign comprehension performance.

According to McDougall et al. (1999), in addition to the consideration of sign features that are self-evident (e.g. colour and shape) or those that can be identified only in relation to other signs (e.g. distinctiveness), cognitive sign features like familiarity, concreteness, complexity, meaningfulness, and semantic distance are of the central concern in sign research. Familiarity is the frequency with which signs have been encountered. Signs are concrete if they depict real objects, materials, or people; those that do not depict real objects are abstract. Signs are regarded as complex if they contain a lot of details or are intricate, and simple if they only contain few elements or little details. Meaningfulness refers, rather obviously, to how meaningful people perceive signs to be. Semantic distance is the closeness of the relationship between what is depicted on a sign and what it is intended to represent.

Although there are studies to investigate the effects of sign design features on the comprehensibility of traffic sign (Ng & Chan 2007, 2008), systematic analysis and quantification of the five cognitive design features for safety signs was not found. The four cognitive sign features of familiarity, concreteness, simplicity, and meaningfulness were studied in this experiment. As questions about the comprehensibility of safety signs were set at the later part of the questionnaire, intended meaning of what was depicted on the sign had to be absent and thus semantic closeness was not studied. It was expected that participants will comprehend signs better if the signs are familiar to them. As concrete signs provide a direct visualization aid in helping participants to elicit the meaning, they were expected to have higher comprehension scores than abstract ones. Extraneous decorative parts on signs could confound the meaning of the signs (Bruyas et al. 1998), so it was hypothesized that simple signs will be easier to comprehend than complex ones. The meaningfulness of a stimulus depends on its associated imagery and refers to the ability with which a stimulus can elicit a meaning in one's mind (Preece et al. 1994). Hence, better comprehensibility was expected for meaningful signs. In summary, comprehension score was expected to be higher for familiar signs, concrete signs, simple signs, and meaningful signs. The results of this study should provide useful information and recommendations for the design of more user-friendly safety signs.

2 METHOD

A small scale survey based on convenience sampling was designed and conducted for capturing subjects' demographic information, colour vision status, personal factors, comprehension scores, and ratings on sign features.

2.1 *Participants*

Thirty-nine male and twenty-three female Hong Kong Chinese participants, aged 15 to 60, voluntarily participated in this experiment. None of the participants had colour vision deficiencies.

2.2 *Safety signs*

25 safety signs currently used in Hong Kong were selected for study in this experiment (Table 1). The choice of the 25 signs was based on three criteria: firstly, the signs are seen in Hong Kong public or workplaces areas; secondly, the message is conveyed with symbols only; thirdly, the signs are not used in conjunction with other signs for transmitting a message.

2.3 *Survey and procedure*

For effective communication, the Chinese version of the questionnaire was used. The first part of the questionnaire collected participants' personal and safety related characteristics. Participants were also asked to answer the question about their color vision status. The second part of the survey focused on safety sign features and comprehension. Participants were asked to give subjective ratings between 0 to 10 points for familiarity (0 = very unfamiliar, 10 = very familiar), concreteness (0 = definitely abstract, 10 = definitely concrete), simplicity (0 = very complex, 10 = very simple), and meaningfulness (0 = completely meaningless, 10 = completely meaningful) on the signs. They were also asked to complete a set of multiple-choice questions for evaluating their understanding of the safety signs. Six-option multiple-choice questions were used with one correct answer, three plausible distractors and two special answers in the six verbal labels in each question. The two special answers were designed based on ISO 9186-1: 2007. One special answer was "do not understand" and another special answer was the opposite of the intended meaning of the sign.

3 RESULTS

One-sample Kolmogorov-Smirnov test ($Z = 0.744$, $n = 25$, $p > 0.05$) showed that the comprehension score for all signs was approximately normally distributed. The mean and standard deviation of the comprehension score for all signs were 55.81% and 32.96%, respectively. Analysis of variance (ANOVA) and Kruskal-Wallis test were conducted to examine whether there were any statistically significant differences in comprehension performance among different levels of personal factors. The results of Kruskal-Wallis test showed

Table 1. 25 Hong Kong safety signs used in this experiment.

Reference number	Sign	Verbal label	Reference number	Sign	Verbal label
S1		Danger	S2		Wear ear protection
S3		Caution trip hazard	S4		No entry
S5		Electrical hazard	S6		Slippery floor
S7		Wear helmet	S8		No climbing
S9		Caution! High level work in progress	S10		Danger! No entry
S11		Emergency exit	S12		Slippery floor
S13		Mind the gap	S14		Danger! Overhead live wires
S15		Do not use lift when there is a fire	S16		Caution, slippery floor
S17		Do not use escalator when transporting any large package	S18		Deep water, Stay away
S19		Caution! Toxic	S20		Use safety belt when working at height
S21		No balloon	S22		Hold the hand rail
S23		Keep clear	S24		No smoking
S25		No standing			

that there were no significant main effects of gender group, occupational experience, and safety education background ($p > 0.05$). The results of ANOVA showed that there was significant main effect of driving experience [F (1, 61) = 13.41, $p < 0.005$]. Education level, age group, years of residence in Hong Kong, and safety professional background were non-significant ($p > 0.05$), showing that there was no significant correlation between comprehension score and these factors.

The mean ratings on familiarity, concreteness, simplicity, and meaningfulness for all signs were 4.52, 5.83, 6.14, and 6.20, respectively. These illustrated that the safety signs were perceived to be moderately unfamiliar, concrete, simple, and meaningful. Table 2 shows the signs with the lowest and highest ratings on familiarity, concreteness, simplicity, and meaningfulness. Pearson correlation analysis showed that there was a significant interrelationship among the safety sign features (Table 3). The highest correlation was found between meaningfulness and concreteness ($r = 0.977$, $p < 0.001$). The results of Pearson correlation analysis also showed that other than the simplicity rating, other three features were all significantly correlated with comprehension score. Concreteness was most highly correlated with comprehension score ($r = 0.710$, $p < 0.001$) while familiarity correlated with comprehension score at the lowest level of significance ($r = 0.458$, $p < 0.001$).

4 DISCUSSION

4.1 *Personal factors*

For the eight personal factors, interestingly driving experience was found to be a significant predictor of sign comprehensibility. It is known that participants with driving license must necessarily acquire traffic sign information because they are required to understand and obey the road users' codes and meanings of traffic signs prior to getting a driving license. Both the general safety signs and traffic signs are made of three components: image/text, border, and background. For example, triangular sign means warning, circular sign represents mandatory, and quadrilateral sign depicts guide information. This may explain why participants in this group performed better in comprehending the intended meanings of safety signs in the experiment.

No difference was found in comprehension performance amongst the three age groups of participants. This may be due to the lack of participants aged 50 years or above. It is suggested to increase the sample size to include older participants in future study. The findings of non-significant factors gender and education level are not surprising as Chen & Wang (2003) also found that gender and education level did not have significant effects on the comprehension of dangerous materials symbols.

It was found that 'occupational experience' and 'safety professional background' had no effect on comprehension performance. The expectation that participants with such experiences were familiar with the use of safety sign in their workplaces and thus could more easily interpret the safety signs was thus invalidated. Past studies had investigated the effect of context on sign comprehension (Wogalter et al. 1993) and it was found that the provision of context, for example photographs of potential locations where the signs might be viewed, can improve symbol comprehension (Wolff & Wogalter 1998). Bazire & Tijus (2009) pointed out that the environmental context in which the sign is perceived determines the sign meaning. However, the results here indicated that even though participants were familiar with the context of signs, they might not be able to perceive the importance of the signs and the associated hazards that the signs intended to represent.

Table 2. Signs with lowest and highest ratings on familiarity, concreteness, simplicity, and meaningfulness.

Sign features	Signs with lowest ratings		Signs with highest ratings	
Familiarity		S25–No standing (1.81)		S24–No smoking (9.34)
Concreteness		S16–Caution, slippery floor (3.10)		S24–No smoking (9.24)
Simplicity		S25–No standing (3.23)		S24–No smoking (9.16)
Meaningfulness		S25–No standing (3.55)		S24–No smoking (9.42)

Table 3. Pearson correlation coefficients amongst safety sign features and comprehension score.

	Familiarity	Concreteness	Simplicity	Meaningfulness
Concreteness	0.876**	1		
Simplicity	0.890**	0.815**	1	
Meaningfulness	0.878**	0.977**	0.809**	1
Comprehension score	0.458*	0.710**	0.344	0.700**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Safety education background was not a significant predictor of sign comprehensibility. This may be due to the lack of training course on safety signs given to the participants with safety education background. Wogalter et al. (1997) found that training would lead to a significant improvement in pictorial comprehension. This may explain why participants who had registered safety education qualification did not perform better than those had not. The factor 'period of time living in Hong Kong' had no effect on comprehension performance, suggesting that participants who had longer period of living in Hong Kong did not pay more attention to safety signs in their daily life. The above findings strongly caution the government about the importance of introducing effective education program for safety signs to the general public and even safety professionals.

4.2 *Sign design features*

Regarding sign features, comprehensibility was found high for familiar signs and low for unfamiliar signs. In a traffic sign comprehension study, Shinar et al. (2003) found that infrequent traffic signs are more likely to be miscomprehended and less likely to be correctly learned by drivers. Rosson (2002) specified that designers should use familiar symbols as much as possible. However, what is familiar to one person may not be familiar to another, making sign design decisions about familiarity is not always easy and simple. Concrete signs contributed to a higher comprehension scores than abstract ones. Where signs are abstract, interpretation of their meanings will be much more difficult. Users have to learn and understand an abstract sign without any assistance from representational elements within a graphic itself. Luftig (1983) noted that concrete concepts have been proved to be easier to learn than abstract ones. As symbols become more abstract and as phenomenal distance increases, difficulty in decoding concepts increases. Thus, the more abstract the concept, the greater the difficulty of the concept comprehension and the more intelligence and experience required for successful learning (Vygotsky 1962).

Regarding sign simplicity, Bruyas et al. (1998) suggested that extraneous decorative parts confounded understanding of the signs; and Mullet & Sano (1995) suggested that a simpler and more immediate interpretation of an icon can be achieved by removing all but the most distinctive and characteristic features of the icon. However, in this experiment, sign simplicity was found to have no significant correlation with sign comprehension score.

With regard to sign meaningfulness, the comprehension scores were high for meaningful signs and low for meaningless signs, probably because meaningful stimulus is related to its associated imagery and easily elicits a meaning in one's mind (Preece et al. 1994). Overall, the comprehensibility of safety signs is better when the signs are familiar, concrete or meaningful.

5 CONCLUSION

This experiment demonstrated that some personal and design factors are important for effective communication of safety sign messages. Driving experience was an important determinant affecting sign comprehension performance. Sign familiarity, concreteness, and meaningfulness were correlated with the comprehension score while simplicity was not. The findings provide useful information and recommendations for the design of more user-friendly safety signs. The findings may also serve as a useful guide for interface designers to design and evaluate icons across various types of consumer products.

ACKNOWLEDGEMENT

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [CityU 110508]. The authors thank for the data collection done by KH Ko.

REFERENCES

Bazire, M. & Tijus, C. 2009. Understanding road signs. *Safety Science* 47(9): 1232–1240.

Bruyas, M.P., Le Breton, B. & Pauzié, A. 1998. Ergonomic guidelines for the design of pictorial information. *International Journal of Industrial Ergonomics* 21: 407–413.

Chen, M.T. & Wang, A.H. 2003. Effects of dangerous materials symbols, educational background, and gender on conceptual compatibility. *Institute of Occupational Safety and Health Journal* 11(3): 188–196.

Duarte, M. Emília, C. & Rebelo, F. 2005. Proceedings of CybErg 2005. Comprehension of safety signs: internal and external variable influences and comprehension difficulties by disabled people. *The Fourth International Cyberspace Conference on Ergonomics*. Johannesburg: International Ergonomics Association Press. Available at: http://web.wits.ac.za/NR/rdonlyres/706425E4-FBAB-4CB2–9D3C-A233BA0AFCEA/0/access1.pdf

Handcock, Holly, E., Rogers, Wendy, A., Schroeder, Derek & Fisk, Arthur, D. 2004. Safety Symbol Comprehension: Effects of Symbol Type, Familiarity, and Age. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46(2): 183–19.

Hong Kong Occupational Safety Health Council 2009. *Occupational Safety and Health Statistics Bulletin Issue No. 9, August 2009.* Available at: http://www.oshc.org.hk/tchi/resource/health.asp?

ISO 9186-1:2007 (E) *Graphical symbols—Test methods—Part 1: Methods for testing comprehensibility.*

Leung, A.K.P. & Hellier, E. 1998. Perceived hazard and understandability of signal words and warning pictorials by Chinese community in Britain. *Contemporary Ergonomics*: 321–325.

Luftig, R.L. 1983. Variables Influencing the Learnability of Individual Signs and Sign Lexicons: A Review of the Literature. *Journal of Psycholinguistic Research* 12(4): 99–122.

McDougall, S.J.P., Curry, M.B. & de Bruijn, O. 1999. Measuring symbol and icon characteristics: norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, & Computers* 31(3): 487–519.

Mono, R. 1997. *Design for product understanding Liber*. Stockholm, Sweden.

Mullet, K. & Sano, D. 1995. *Designing Visual Interfaces: Communication oriented techniques.* Englewood Cliffs, NJ: Prentice Hall.

Ng, A.W.Y. & Chan, A.H.S. 2007. The guessability of traffic signs: effects of prospective-user factors and sign design features. *Accident Analysis and Prevention* 39: 1245–1257.

Ng, A.W.Y. & Chan, A.H.S. 2008. The effects of driver factors and sign design features on the comprehensibility of traffic signs. *Journal of Safety Research* 39: 321–328.

Preece, J., Rogers, Y., Sharp, H., Beryon, D., Holland, S. & Carey, T. 1994. *Human-Computer Interaction*. Addison-Wesley, England: 99–122.

Raskin, J. 2000. *Humane Interface: new directions for designing interactive systems*. ACM Press, Massachusetts, USA: 168–173.

Redström, J. 2006. Towards user design? On the shift from object to user as the subject of design. *Design Studies* 27(2): 123–139.

Rosson, M.B. 2002. *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*. Academic Press, San Francis: 119–125.

Shinar, D., Dewar, R.E., Summala, H. & Zakowska, L. 2003. Traffic sign symbol comprehension: a cross-cultural study. *Ergonomics* 46(15): 1549–1565.

Tam, C.M., Fung, I.W.H., Yeung, T.C.L. & Tung, K.C.F. 2003. Relationship between construction safety signs and symbols recognition and characteristics of construction personnel. *Construction Management and Economics* 21: 745–753.

Vygotsky, L.S. 1962. *Thoughts and language*. Edited and translated by Eugenia Harfmann & Gertrude Vakor. Cambridge: MIT Press.

Whittingham, R.B. 2004. *The blame machine: why human error causes accidents*. Elsevier Butterworth-Heinemann.

Wogalter, M.S., Kalsher, M.J. & Racicot, B.M. 1993. Behavioral compliance with warnings: effects of voice, context, and location. *Safety Science* 16(5–6): 637–654.

Wogalter, M.S., Sojourner, R.J. & Brelsford, J.W. 1997. Comprehension and retention of safety pictorials. *Ergonomics* 40(5): 531–542.

Wolff, J.S. & Wogalter, M.S. 1998. Comprehension of pictorial symbols: effects of context and test method. *Human Factors* 40(2): 173–186.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Cultural representation by card sorting

Javed A. Sheikh, Bob Fields & Elke Duncker-Gassen
Middlesex University, London, UK

ABSTRACT: The paper presents a concept of cultural representation for interaction design by sorting of cultural concept, integration of concepts into categorisation and interaction design for cultural representation. The representation used in interactive interfaces design causes problems for users from a different culture to that of the designer. A starting assumption for this research was that no scheme for organising information is likely to be equally effective for a range of cultural groups. The current research aims to make a contribution in this area, not by finding a universal way of classifying information, but by providing a method for investigating classification in a locale in order to generate localised interface designs. The expected solution will be based on local user access needs and capability of the local users.

Keywords: cross-cultural design; cultural information access; human-computer interaction, human factors

1 GENERAL INTRODUCTION

This paper presents the results of a pilot study. The principal aim of this pilot study was to clarify issues involved in research and to explore methods for main study. An effective research needs careful planning of all aspects of study and pilot study helps to achieve this goal. We designed small experiment to test our logistic, remove all discrepancies and get information before the main study to improve its quality and effectiveness. Therefore the lesson we learnt from pilot study will be incorporated in main study. The pilot study was undertaken to gather detailed evidence on how participants grouped the concepts and defined their categories. We also used pilot study to resolve following issues before main study:

- Review the instructions;
- Review the procedures;
- Check the correct cards;
- Check the reliability and validity of results.

We used card sorting technique to investigate differences and similarities in classification. Card Sorting experiments are used to understand users' perceptions of relationships among items. The participants sorted the cards into categories based on their perception and preference and then provided a label for each category.

In the forthcoming sections we will discuss the design of the studies, the data collection techniques employed, the intended data analysis, of pilot study.

2 BACKGROUND

Technically, the term culture is associated with anthropologists. They use it to describe a group of people, who have common characteristic. Therefore, every discipline has its own definition according to their specific area of research. Although there is no standard definition of culture, most researchers incorporate the Boasian definition.

"The system of shared beliefs, values, customs, behaviours, and artefacts that the members of society use to cope with their world and with one another and it is transmitted from generation to generation through learning."

Cultural data is to observe behaviours, performances and conceive ideas which are available only as tacit knowledge. Franz Boas gives the concept of cultural anthropology, cultural relativism and the participant-observation method of fieldwork. He integrates that specific cultural environment structure is based on individual behaviour.

The scope and reach of computing is truly global, spanning geographical and cultural boundaries, yet few scholars have investigated the influence of culture. The rapid growth of computing raises issue of cultural representation as there is no representation for all cultures. Cultural study is an important phenomenon in academic and research, but computer scientists neither consider it is essential element for them nor they think user is important (Fernandes, 1995; Sheridan, 2001).

The computing depends heavily on traditional classification system. The most important cultural clash is lack of culturally specific system. For in-stance, none of available systems are fully capable of classifying and organising material according to the culture. The existing classification schemes Dewey Decimal Classification (DDC), Library of Congress Classification (LC), and Ranganathan's Colon Classi-fication (RCC) are unsatisfactory, in one way or another. The two popular American origin systems, DDC and LC have great qualities of expansion, alte-ration and adjustment for new subjects. However, these schemes cannot classify eastern material in an appropriate way in general and specially in religion, language, literature, and history. Although the both systems are used internationally, they often falls short of meeting the needs of many cultures and countries. The important question is whether LC and DDC can continue to meet the needs of expanding knowledge and changing worldviews. The Middle East is gaining importance, and yet Syria, Lebanon, Palestine, Israel, and Jordan all share the classification of 956.9 same with religion where 90% of class 200 occupied by Christianity. In New Guinea "kobtiy" is an animal but culturally it is bird (Latour, 1987). Therefore, it is need to organize material based on cultural know-ledge. Duncker, (2002) emphasises on to create sys-tem for cross-cultural user.

The studies (del Galdo & Nielsen, 1996; Duncker, 2002) show that culture plays a vital role in research. The cross-cultural research is about people's perceptions and attitudes about what they do and how they behave. Cross-cultural research defines cultural similarities and differences in social behaviour. Therefore cross-cultural research demands to study real lives in real contexts. Cross-cultural research helps to create a better world by the study of people of different cultures. It also helps to improve knowledge and guide to survive in diversity. Investigating cultural variations is one of the goals in psychological functioning. But there is difficult to find variations in behaviour due to not direct personal contact between researcher and the people studied. Cross-cultural theories are often formulated in a manner that makes testing difficult. The cross-cultural research is gaining attention and new findings incorporating the culture specially language and thought. It is need to improve communication and share their visions among different cultures.

The researchers and designers unintentionally apply their own cultural values when design and develop information systems. Although Microsoft and other development organisations consider cultural issues but mostly involve language translation and visual aspects of the interface instead of the underlying structure of the application. (Russo & Boor, 1993; De Angeli et al., 2004).

Different peoples have different interpretations which sometime create confusion. This research is to find how non-western user structures their knowledge and to discover (if there are) differences among the knowledge of structure that are belong to western cultures. There are many methods for data collection on representation of knowledge (Rugg & McGeorge, 1997). One of the methods is card sort. Card sorting is used to explore knowledge reside in mind. Therefore we used card-sorting techniques to elicit information from participants by observing how they categorise. Card Sorting methodology used to conceive users' metal model and show the direction that how information can be structured in an interface (Faiks & Hyland, 2000). It is, systematic and easy to use (Rugg et al., 1992) technique which helps to find how much similarities and differences are in categories. Card sorting technique facilitate to develop and identify concepts, models, attitudes, trends, patterns and values for capturing information from participant's mental model. The mental model generates concepts and suggests possible taxonomies.

Card sorting technique is widely used in Human Computer Interaction, psychology, knowledge engineering for knowledge elicitation. It helps to evoke participant's domain knowledge (Rugg & McGeorge, 1997), distinguish the level of problem (Barrett & Edwards, 1995), reflects their idea about knowledge (Zimmerman & Akerelrea, 2002) and results help for further analysis (Upchurch et al., 2001) and can be used for further research. The card sorting technique is fast, cheap, and easy to handle for both researcher and participate (Zimmerman & Akerelrea, 2002, Rugg et al., 1992). The participants sort the card according to pre defined categories. (Upchurch et al., 2001) or categories according to own understanding. This is common phenomena that the people always disown and resist against the things which do not relate to their culture norm, tradition, sentiments and values. The existing Interactive Information Systems have several challenges, including effective organisation and efficient retrieval of information in general and particularly in cultural context. The limited extent of culture and unfamiliar terminology/classification create the demand of cultural oriented Interactive Information system. Therefore, this problem forces researchers and developers to devise a new mechanism to support interactive information system. This would be based on cultural controlled vocabularies. This research will explore this matter extensively find the way and techniques to overcome this problem as well as enhance the user access. For example during humanitarian crisis, volunteer teams need to obtain information about affected area's culture, languages, tradition etc.

There is no empirical study to design cultural based IIS as previous research shows that researcher used or extended other researcher's models. The use of cultural models in IIS is rare and few researchers used these business oriented cultural models. These models neither fulfil the user's need nor compatible with IIS. Whereas developers overlook cultural considerations to meet deadline and do not feel necessity to do research on cultures. There is a need for empirical study to provide a better understanding of cultural specific IIS to reduce the cultural gap. In the light of above-mentioned facts the problems emerge as locally specific content cannot be appropriately represented and classified in an Interactive design. In general, Interactive Information Systems (IIS) have followings issues;

While some Interactive Information systems are global, the users are always local;

- Interactive Information systems heavily depend on western classification;
- All users are not familiar with Interactive Information Systems' organisation;
- There is lack of culturally specific terminology for IIS;
- Due to global aspects of Interactive Information Systems, most users fail to get the desired result;

Interactive Information System designers do not pay attention on cultural aspect therefore there are difficulties to use Interactive Information Systems for people of a cultural background whose culture is different to the one of Interactive Information system.

A question of recurrent interest is how easily certain groups of users can retrieve information from web-based information sources. A wide range of online classification schemes can be found, of which some seem to have a wider applicability and acceptance than others. For example, UK online stores do not only sell different products compared to similar German online stores, they often classify their products differently.[1] Similar things can be said for classification systems in libraries. The Dewy decimal classification system is used worldwide and yet, it classifies books differently to the German library classification systems in general and in particular to specialty related classification systems (Heiner-Freiling, 1998; Marcus & Gould, 2007). This means that not only the content but also the way this content is organised and classified reflects the values and interpretive practices of the culture in which it was produced. Therefore, problems can arise, when content designed, organised and classified by members of one culture is used by members of another culture. Typically, web content, its organisation and its classification reflect values and interpretations of western cultures rendering it less appropriate to non-western cultural user groups. As part of a larger study, this research focuses on cross-cultural classification practices Interactive Information System. It also examines the way people classify differently and what are the causes of differences. As there is a strong need of Interact Information system based on cultural knowledge. In this research, I will study problems cause by the ambiguity of cultural terminology in Interactive information system. I will also propose and present a model to provide a solution to this problem.

The research is based on multi-disciplines. These include Information science, computer science, knowledge management, anthropology, psychology, sociology, and technology. The result of this research will provide an easy-to-use IIS. This research also will help to bringing the IIS benefits to the threshold of every culture. Therefore to overcome this situation there is a need to accommodate locally specific needs by using culturally specific classification. It is therefore important, to integrate the above mentioned feature. The Following tasks will need to be taken into consideration:

- Facilitate cultural terminology to use IIS efficiently;
- Better effective presentation of IIS;

To achieve the above mentioned aims we have to:

- build Cultural-specific controlled terminology;
- Integrate vocabulary with the IIS.

3 PILOT STUDY

The aim of the pilot study was to examine the differences in the way participants categorise food items. Categorisation is true reflection of specific concepts with some inconsistency. It was expected to find cultural and belief differences and result of the studies confirmed this hypothesis. The following method adopted to achieve the result; the participants of our study were being asked to classify items in several steps. First they group items and gave labels to each group. Then they were asked to further divide the group if they wish to do so. The steps of grouping and labelling are repeated until the subject no longer want to sub divide the groups of items.

Domain: The domain chosen for this research is food items as these are very common to everyone. Several other domains were discussed at the beginning, and after detailed review fooditems

[1] For a comparison see for instance the Galeria-Kaufh of website and the Debenhams website.

was the chosen one. Other domain like natural disaster, combination of food item and natural disaster also considered. They were rejected as participants face difficulty to perform them. We used 39 cards, so participants complete the task in adequate time. Every card had its own number for reference.

Participants: There is no census on how many participants to include in open card sorting study. Some card sorting guides suggest (Gaffney, 2000; Robertson, 2002), 4 to 6 (Maurer et al., 2004; Nielsen, 2004), 10 to 15, (McGovern, 2002; Tullis et al., 2004) 20 plus and Tullis and Wood (2004) suggested 20 to 30 participants. It is also to be noted that larger sample size increase the cost and analysis time. However, informal polls conducted at the 2007 Information Architecture Summit and local Usability Professionals Association meetings proposed 6 to 12 participants.

A total of sixteen Middlesex University students participated in this pilot study and were selected based on their ethnicity and duration of their stay in the UK. They were familiar with all the items on the cards. They were divided into two groups; Asian participants, who had stayed in the UK for less than six months and English participants with no apparent foreign background.

- The participants were born and raised in a target culture.
- Were literate and over 18 year of age.
- Were naive to the purpose of the experiment and not familiar with the testing.

Procedure: We provide brief introduction of research and instruction. The instructions for participants were as follows:

1. Group the cards in a way that makes sense to you. (Start by placing the first card on the table and then look at the second card to see whether it belongs in the same group or if it deserves its own category—and so on through the set of cards).
2. Once the groups are established, label each group.
3. You are allowed to make sub-groups if you feel that's appropriate.
4. Feel free to ask questions during the exercise if you feel the need. I can't guarantee that I can answer them during the experiment, but I'll do my best to answer them when you're finished.
5. Your categorisation will be kept confidential and will not be linked to you personally; it will be reported anonymously.
6. You can withdraw at any time without any further consequences. Your participation poses no risks to you but the benefits will contribute to the research that is the subject of the experiment.

4 RESULT: INFORMAL ANALYSIS

This informal analysis of the pilot study has revealed that Asian and British participants differed in their categorisation judgments. However, they shared a common representation structure in a few categories. The differences are also noticed within the culture. The pilot study gives a valuable result regarding food classification with relation to different cultures (Asian and British) such as

- *Differences: Asian and British Culture*
  - Halal, Kosher and Non-Vegetarians food items are mainly noticeable.
  - Chicken and Turkey are both put together in different category label for example Poultry (British), Meat (Asian).
  - Pork and bacon are both put together in different category label for example Meat (British) Non-Vegetarian (Asian).
- *Differences: Asian and British Beliefs*
  - Pork and bacon are both put together in different category label for example Non-Kosher (British), Non-Halal (Asian).
  - Egg put in different category label for example Dairy (British), Poultry (Asian).
- *Differences: Asian Culture*
  - Chicken and Turkey are both put together in different category label for example Non-Vegetarian (Indian Jain), Meat (Indian Hindu).
  - Fish and Prawn are both put together in different category label for example Non-Vegetarian (Indian Jain), Non-Halal (Indian Hindu).
  - Pork and bacon are both put together in different category label for example Non-Vegetarian (Indian Muslim), Non-Halal (Pakistani Muslim).

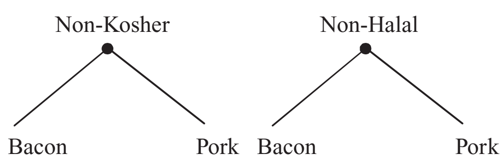


Figure 1. Result regarding food classification within a single culture (Asian) such as.

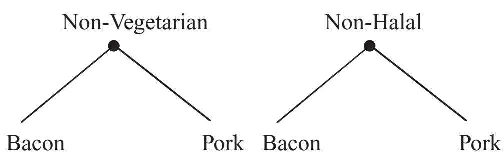


Figure 2. Result regarding food classification within a single culture (Asian) such as.

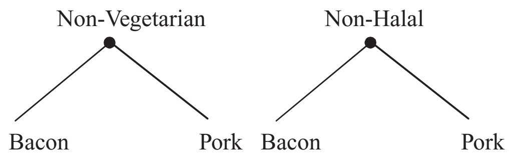


Figure 3. Result regarding food classification within a single culture (Asian) such as.

Table 1. Hypotheses.

Culture	Belief	Classification
Same	Same	Very similar
Same	Different	Moderately similar class
Different	Same	Some similarities some differences
Different	Different	Very different

- *Differences: British Culture*
  - Coffee and Tea are both put together in different category label for example Caffeine (Atheist), Drink (Chruch of England).
  - Chicken and Turkey are both put together in different category label for example Meat (Chruch of England), Poultry (Roman Catholic).
  - Egg put in different category label for example Dairy (Chruch of England), Poultry (Roman Catholic) Egg (Jewish).
- Poultry, Halal, Kosher and Non-Vegetarians food items are mainly noticeable.
- The food classification is based on individual belief; for example
  - Kosher, Halal and Non-Vegetarian items.
  - Pork and bacon are both put together in.

The pilot study results suggest that members of different cultures systematically organise and categorise items differently.

The pilot study helped to obtain an initial understanding of the problem. The result of the studies suggested that both the 'national culture' and the 'belief system' of a participant shape the way they categorise items. By 'belief system' here, we refer roughly to religious background as this is a highly significant factor in the way people understand food and the various domestic practices that surround it. If items other than food stuffs had been used for the study, it seems likely that other elements of culture, such as professional cultures or membership of communities of practice would gain greater significance.

In our hypothesis independent variables are Culture and Belief System. The dependent variable is the differences and similarities in classification. Our hypotheses are:

Table 2. Null hypotheses.

Culture	Belief	Classification
Same	Same	Moderately similarly or very different
Same	Different	Either very similar or very different
Different	Same	Either very similar or very different
Different	Different	Moderately or very similar

Therefore the null hypotheses will be as shown in Table 2.

5 CONCLUSION

We believe that from this pilot study there is a huge cultural and belief difference in categorisation among cultures and beliefs. The informal results of the pilot study encourage a further an in-depth study to verify the generality of these claims. In this context, a major field study has been planned; Pakistan and UK.

The contribution of this pilot study is to investigate how people of similar cultures and with similar faith perceive and classify same elements. The result of the studies reveals that the categorisations made by the participants differed not only between cultures and belief systems but were also found to be different even within a single culture. The pilot study observations helped to obtain an initial understanding of the issue.

We presented our research methods and explained why we choose them for this research. We hope this will enable reader to understand the problem and its solution will give them better overview of this research. The research aim is to propose a design for all cultures. The further field study and its analyses help to increase usability interaction in different cultures.

This research has threefold Contribution to knowledge

1. A theory of differences in culturally based classification between Culture A and Culture B.
2. A method for doing it how to capture and how to write it down in terms if tree structure and edit distance.
3. Partially/proposed solution for best representation to use and organise web page.

As this research proposes the Cultural based Interactive Information System, where users, without any discrimination have access to the processes of creation and exchange of knowledge. Therefore this research will help to bringing the

Interactive Information System benefits to the threshold of all cultures. In this way, people take charge of their own knowledge and technology transfer needs, narrowing the digital divide.

The solution will contribute:

- To raise awareness about the value of cultural knowledge,
- To increasing the efficiency and functionality for accessing non-familiar culture resources,
- To bridge the gap among different cultures,
- To enhance sharing of knowledge among cultures.

The abovementioned observations indicate that research relating to culture and technology needs more attention. Therefore, it is need to consider culture influences, cultural needs, and cultural difference when building interfaces. I would like to investigate, how can culturally specific variables increase the effectiveness, efficiency and satisfaction of computing user?

REFERENCES

Akerelrea, C. & Zimmerman, D., 2002. A group card sorting methodology for developing informational web sites. In *Proceedings of the 2002 IEEE Professional Communications Conference,* 437–445.

Barrett, A.R. & Edwards, J.S. 1995. Knowledge elicitation and knowledge representation in a large domain with multiple experts. *Expert Systems with Applications* 8(1): 169–176.

Del Galdo, E. & Nielsen, J. 1996. *International User Interfaces*. New York: Wiley publishing.

Duncker, E. 2002. Cross-cultural usability of the library metaphor: In *Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries*, 223–230.

Faiks, A. & Hyland, N. 2000. Gaining user insight: a case study illustrating the card sort technique. *College & Research Libraries* 61(4): 349–57.

Fernandes, T., 1995. 'Global interface design'. Human Factors and Computer Systems.

Firlej, M. & Helens, D. 1991. *Knowledge elicitation, a pratical handbook*. New York, Prentice Hall.

Gaffney, G. 2000. What is Card Sorting? *Information & Design.* Retrieved June 6, 2010, from http://www.infodesign.com.au/usabilityresources/cardsorting

Heiner-Freiling, M. 1998. DDB und DDC – Die Deutsche Bibliothek und die Dewey-Dezimalklassifikation Perspektiven, Befuerchtungen, Hoffnungen in Bibliotheksdienst, Heft 12. (in Dutch)

Katz, I., Mack, R., Marks, L. (Eds.) 1995. Conference companion on Human factors in computing systems. New York: ACM Press, 363.

Latour, B. 1987. *Science in action: How to follow scientists and engineers through society*. Harvard: Harvard University Press.

Marcus, A. & Gould, E.W. 2000. Cultural dimensions and global web user-interface design: What? So what? Now what? Retrieved June 6, 2007, from http://www.amanda.com/resources/hfweb2000/AMA_CultDim.pdf

Maurer, D. & Warfel, T. 2004. Card sorting: a definitive guide. *Boxes and Arrows*, retrieved June 6, 2010, from *http://www.boxesandarrows.com/view/card_sorting_a_definitive_guide*

McGovern, G. 2002. Information architecture: using card sorting for web classification design. Retrieved June 6, 2007, from *http://www.gerrymcgovern.com/nt/2002/nt_2002 _09_23_card_sorting.htm*

Nielsen, J. 2004. Card sorting: How many users to test. *Jakob Nielsen's Alertbox,* retrieved June 6, 2010, from *http://www.useit.com/alertbox/20040719.html*

Robertson, J. 2002. Information design using card sorting. *Intranet Journal,* retrieved June 6, 2010, from *http://www.intranetjournal.com/articles/200202/km_02_05_02a.html*

Rugg, G. & McGeorge, P. 2005. The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems* 22(3): 94–107.

Russo, P. & Boor, S. 1993. How fluent is your interface? Designing for international users. In *Proceedings of the conference on Human factors in computing systems*. Boston: Addison-Wesley Longman Publishing, 342–347.

Sheridan, E.F. 2001. Cross-cultural web site design: Considerations for developing and strategies for validating locale appropriate on-line content. MultiLingual Computing, Inc. Retrieved June 6, 2007, from *http://www.multilingual.com/articleDetail.php?id = 595*

Tullis, T. & Wood, L. 2004. How many users are enough for a card-sorting study? In *the proceedings of Usability Professionals Association Conference*. Minneapolis, MN.

Upchurch, G.R., Welch, G.N. & Fabian, A.J. et al. 1997 Homocyst(e)ine decreases bioavailable nitric oxide by a mechanism involving glutathione peroxidase. *J. Biol. Chem.* 272: 17012–17017.

Part IV: Musculoskeletal disorders & ergonomics intervention

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Managing work-related MSDs in Europe—Lighten the Load

Z. Podniece & T. Tregenza
European Agency for Safety and Health at Work, Bilbao, Bizkaia, Spain

ABSTRACT: MusculoSkeletal Disorders (MSDs) are the most common work-related health problem in the EU, affecting millions of European workers across all employment sectors. The European approach to MSDs, promoted during a European Campaign on MSDs (Lighten the Load), recommends that risk-assessment based preventive measures are put in place alongside with appropriate rehabilitation and reintegration of those workers suffering from MSDs. Effective occupational safety and health management, information for and consultation with workers, training and health surveillance are also part of such measures. This paper reflects on the Lighten the Load Campaign and includes information on evidence-based workplace interventions to prevent MSD risks and to retain employees with musculoskeletal problems at work. In addition, the paper provides good practice examples on how to tackle MSDs and information on what tools in this area are available for experts and those in the workplace.

Keywords: work-related musculoskeletal disorders, prevention, return-to-work, multidisciplinary approach, good practice, campaign

1 INTRODUCTION

MusculoSkeletal Disorders (MSDs) are the most common work-related problem in Europe. According to a European survey, up to 25% of those workers in the 27 Member States of the European Union (EU27) who perceived impact of work on health reported back pain and 23% muscular pain (European Foundation for the Improvement of Living and Working Conditions (EUROFOUND) 2007a). Any worker can be affected by these disorders. It is not a problem only for elderly blue collar male workers (EUROFOUND 2007b). MSDs and exposure to MSDs risk factors are increasing in younger working population. Women are also considerably exposed, but the effects are still under-recognised (European Agency for Safety and Health at Work (EU-OSHA) 2010).

Impairments to an individual's musculoskeletal system cause not only personal suffering and loss in family income, but also constitute high costs for businesses and state economies. Precise figures on overall MSD costs do not exist; however, estimates show that their impact is huge. For example, in some states, 40% of the costs of workers' compensation are attributed to MSDs, and up to 1.6% of the Gross Domestic Product (GDP) of the country itself (Takala 2000).

MSDs have commonly been associated with physically demanding working conditions (EUROFOUND 2007b). However, other factors such as vibration, cold or excessive heat, lack of control over the tasks performed and a low level of job satisfaction also play a significant role (Buckle & Woods 2002, EU-OSHA 2007a, EUROFOUND 2007b). In addition, new aspects of MSD risks emerge. European experts have identified lack of physical activity as a top emerging risk related to MSDs (EU-OSHA 2005a). It is not surprising taking into account that work with computers and automated systems is becoming more and more prevalent in Europe (49% of women and 43% of men work on computers for at least a quarter of their working time) (EUROFOUND 2007a) exposing the workers to both physical inactivity and repetitive work (Wahlström 2005). Multi-factorial MSD risks, especially those that include human, social and organisational factors, are perceived by European experts as important issues to be tackled in the future. Furthermore, they indicated that factors such as job insecurity and fear of the future resulting from the unstable labour market accentuate the effect of physical risk factors such as poor ergonomic design, thus contributing to an increase in the incidence of MSDs (EU-OSHA 2005a).

Lighten the Load Campaign, co-ordinated by the EU-OSHA, promoted an integrated management approach to prevent and manage work-related MSDs across Europe. This approach should consider both the risk-assessment based prevention of new disorders and the rehabilitation and reintegration of workers who already suffer from MSDs (EU-OSHA 2007b).

2 LIGHTEN THE LOAD-AN INTEGRATED MANAGEMENT APPROACH TO TACKLE MSD

2.1 *European MSD campaign*

Lighten the Load campaign 2007 (http://osha.europa.eu/en/campaigns/ew2007/) built on the first European campaign for safety and health at work in 2000 which focused on the theme 'Turn your back on MSDs'. The campaign, coordinated by the EU-OSHA, took place on a decentralised basis in all EU Member States, candidates and EFTA (European Free Trade Agreement) countries, beginning with simultaneous launch events and concluding with a closing summit. Furthermore, the Agency made greater efforts to engage partners at a European level, and these efforts paid off in ensuring the involvement of European multipliers such as professional and sectoral organisations. A number of organisations were involved in the European network meetings and as official campaign partners. As an indication of the scale of the Lighten the Load campaign, there were some 194,000 participants in campaign conferences and events.

The Good Practice Award was an important part of the campaigning activities with a strong dimension of European added value. A total of 36 entries from the Member States were forwarded to EU-OSHA, which later on were disseminated at European and national level.

EU-OSHA provided information materials for the campaign and certain activities were arranged at a European level. With its campaign EU-OSHA sought to encourage an integrated management approach to tackling MSDs where risk-assessment based prevention activities would be implemented along with return-to-work measures (Centre for Strategy and Evaluation Services 2008).

2.2 *Promoting an integrated management approach*

2.2.1 *Comprehensive risk assessment*

Proper risk assessment is a key for the preventive actions being successful. It should be based on a holistic approach and consider the total load on the body. Not only should bio-mechanical risk factors (e.g. having to lift heavy loads, exposure to vibration) be considered, but also the work environment (e.g. excessive cold making it harder to grip an object), organisational and psychosocial factors (e.g. high pace of work) and individual factors (e.g. physical capacity, age, experience, organizational factors, the need to wear personal protective equipment during the work activity). Where there are combinations of risk factors, then all potential health outcomes should be considered and addressed together. For example, physical risks for MSDs and work-related stress are thought to be interrelated and could give better results when tackled together (Bongers et al. 2006, EU-OSHA 2005b, EU-OSHA 2007a).

In carrying out the risk assessment, care should be taken to ensure that gender differences in exposure to risk factors or in health outcomes are identified and addressed. Similarly, special attention should be paid to workers who may be at increased risk or have particular requirements such as young workers, migrant workers, home carers, short-term contractors and workers with health problems. Risks causing some overlooked MSDs, such as lower limb disorders also need to be assessed (EU-OSHA, in press). Consultation with workers and their involvement throughout the risk assessment and management process is both essential for effective interventions and a requirement of EU directives (The Council of the European Communities 1989, 1990a, b).

The risk assessment must be completed by a set of appropriate actions targeted to elimination, were possible, or reduction of the risks to musculoskeletal health.

2.2.2 *Prevention of work-related MSDs*

Preventive measures should be based on the holistic approach and address the whole load on the body that may harm the worker's musculoskeletal system. Normally there is no single factor that causes MSDs—for example, manual handling alone is rarely the cause of back pain; there are many other factors that may contribute to its development such as stress, vibration and work organisation (De Beeck & Hermans 2000). Therefore, it is very important to assess the full range of MSD risks and to address them in a comprehensive way.

Effective workplace interventions exist. Results of a literature review carried out by the EU-OSHA on the effectiveness of such interventions revealed the following (EU-OSHA 2008):

- There is strong evidence that technical ergonomic measures can reduce the workload on the back and upper limbs without the loss of productivity and moderate evidence that these measures can also reduce the occurrence of MSDs.
- There is moderate evidence that a combination of several kinds of interventions (multidisciplinary approach) including organisational, technical and personal/individual measures is better than single measures; however, it is not known how such interventions should be combined for optimal results.
- There is some evidence that a participative approach which includes the workers in the

process of change has a positive effect on the success of an intervention.

- Physical training can also reduce the recurrence of back pain and neck-shoulder pain; in order to be effective, however, the training should include vigorous exercise and be repeated at least three times a week.
- There is limited scientific evidence that a reduction in daily working hours can reduce MSDs and that extra pauses for recovery can often be added in an industrial setting without loss of productivity.
- There is strong evidence that training on working methods in manual handling is not effective if it is used as the only measure to prevent low back pain.
- There is no conclusive evidence to support back belt use to prevent work-related low back pain.

Financial commitment from management in the form of appropriate material and human resources that are made available for improvements to the working environment is essential for success of the interventions.

2.2.3 *Improving return-to-work*

Maintaining workers with MSDs at work should be an integral part of workplace MSD policy. Particularly important is the role of social and organisational support in enabling workers with MSDs both to return to work and to stay in work. The active support and involvement of workers at risk and other stakeholders in the organisation is important. Results of a literature review carried out by the EU-OSHA on the effectiveness of measures to keep workers with MSDs at work revealed the following main findings with respect to particular body parts (EU-OSHA 2007b):

Back pain

- There is clear evidence that it is important for patients to stay active and return to ordinary activities as early as possible.
- A combination of optimal clinical management, a rehabilitation programme and workplace interventions is more effective than single elements alone.
- Taking a multidisciplinary approach offers the most promising results, but the cost-effectiveness of these treatments needs to be examined.
- Temporarily modified work is an effective return-to-work intervention, if it is embedded in good occupational management.
- Some evidence supports the effectiveness of exercise therapy, back schools, and behavioural treatment.
- Lumbar supports such as back belts and corsets appear to be ineffective in secondary prevention.

Upper limb pain

- A multidisciplinary approach involving a cognitive-behavioural component might be the most effective type of intervention.
- There is limited evidence on the effectiveness of some technical or mechanical interventions and exercise therapy.
- In the scientific literature, sufficient evidence is not available for the effectiveness of psychosocial interventions.

Lower limb pain

- No information on work-related intervention strategies has been found.
- The results of studies concerning lower limb treatment in general indicate that exercise programmes might be effective for hip and knee problems.

Although many studies have been carried out (Boocock et al. 2007; Hayden et al. 2005; Van Tulder et al. 2006), the evidence for the effectiveness of interventions, in particular regarding interventions aimed at upper limb symptoms, is somewhat limited (Breen et al. 2005; Karjalainen et al. 2003; Meijer et al. 2005). A possible explanation for this lack of demonstrated efficacy of interventions could be that the quality criteria used in scientific reviews may not be applicable to often-complex workplace interventions. For example, randomization often is not feasible. Therefore, studies of successful interventions may not be included in a review or are considered of low quality.

In spite of the lack of strong scientific evidence, anecdotally many of the above work-related interventions are reported as being effective (EU-OSHA 2007b). Therefore, policy-makers and employers should not be discouraged from carrying out preventive action simply because there is no 100% scientific proof. Moreover, secondary and tertiary prevention should go hand in hand with primary prevention to prevent the recurrence of MSD episodes.

2.3 *Results of the campaign*

According to evaluation report, the theme of this campaign was overwhelmingly seen as having been appropriate and in line with national priorities. Besides, following on from 2000, when the theme was last addressed through a European campaign, it was seen as a very good time to re-sensitise the European public to MSD issues and take stock of developments.

In many ways, Lighten the Load campaign was the most successful campaign to date with regard to reaching target groups. This was in part due to its broad appeal and the high importance of the issue of musculoskeletal disorders across Europe, but also because the organisation of the campaigns has become increasingly well institutionalised.

2007 campaign activities were generally considered to have been useful in raising awareness of health and safety issues. The campaign materials produced by the EU-OSHA were also generally seen as useful. According to evaluations, the Good Practice Award (GPA) scheme was perceived as one of the most effective European campaigning tools. More than 80% of GPA participants indicated that they would enter the competition again.

The feedback from evaluation was also quite positive with regard to the sustainability of the campaign results (Centre for Strategy and Evaluation Services 2008).

3 GOOD PRACTICE EXAMPLES AND FURTHER INFORMATION

3.1 *An example of MSD prevention-an innovative solution to eliminate manual handling in the production of construction materials*

The Dutch company Dycore manufactures ribbed floor components. It uses a wooden pallet system to transport these products, in which workers have to handle awkwardly shaped pallet blocks weighing around 25 kg when dry (see Figure 1).

The risk of MSDs arising out of the manual handling of these heavy pallet blocks was recognised during an assessment based on the "NIOSH method" (http://www.cdc.gov/niosh/topics/ergonomics/). A traffic light system was used to indicate risk levels: red for a major risk requiring action, amber for a risk where action is desirable and green for a situation where no action is required.

Potential solutions included changing the design or materials of the pallets and using mechanical lifting aids. However, the company concluded that the best solution was to get rid of the pallets altogether.

A brainstorming session with the employees identified the solution: incorporating a recessed element into the ribbed floor components themselves, rendering the wooden pallets redundant (see Figure 2). The feasibility of this solution was tested, in particular, whether the redesigned components maintained the required quality and construction standards.

The result is that neither Dycore employees nor other workers, such as sub-contracted drivers, have to expend physical effort handling pallets and risk injury.

Other results include:

– At an annual cost of €137,000, the redesign saved €91,000 a year in other component use and €22,000 in the purchase of pallets.

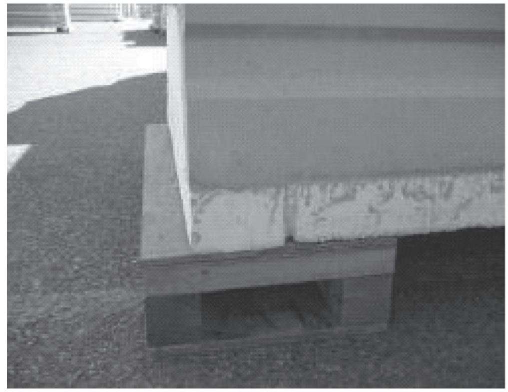


Figure 1. The components before modification sitting on a pallet block.

Figure 2. The components after modification with the recess.

– Improved handling capabilities also meant that three fewer workers were required in this process, saving €120,000.
– In total, there is a net benefit of €96,000 a year.
– A cut in noise as nail guns are no longer needed to repair pallets.
– Quicker deliveries on site because there is no need to load pallets onto trucks (EU-OSHA 2008).

3.2 *An example of back-to-work policy*

With the aim of addressing the growing problem of MSDs, the German government has begun to transfer obligations regarding the participation of people with disabilities in work, from the State and/or social insurance to employers. The focus now is on early recognition and avoidance of long-term incapacity at work. If an employee is unfit for work for more than six weeks within a year, a meeting between the employer and the member of staff must initially be convened in consultation with the works council (workers representatives), in order for constructive and integrative solutions to be reached with the insurers at a subsequent stage. Disability managers support employers in their new role as 'early warning systems' (EU-OSHA 2007b).

3.3 *Further good practice information*

Good practice resources are available on the EU-OSHA MSD Single Entry Point http://osha.europa.eu/en/topics/msds. Furthermore, useful information on MSDs prevention can be found in the databases of case studies http://osha.europa.eu/en/good_practice/index_casestudy and risk assessment tools http://osha.europa.eu/en/good_practice/index_ralink

4 CONCLUSIONS

MSDs are highly prevalent and complex, yet preventable, work-related health conditions. Their prevention is vital to stop the suffering of millions of European workers and to save the costs for businesses and State economies. To reach this objective, it is important that the government, employers and employees, assisted by professionals from a range of disciplines, work together and join their efforts in response to the challenge. At the workplace level, an integrated management approach is needed. It covers both prevention of MSD risks, thus protecting the workers from becoming ill and retaining at work the workers who are already ill. A holistic approach, which considers the total load on the musculoskeletal system, is an efficient way of assessing and addressing MSDs risks. Finally, an eye should be kept on the new and emerging risks, originating from the changing work environments, in order to eliminate such risks from the workplace or to keep them under control.

REFERENCES

Bongers, P.M., Ijmker, S., Van den Heuvel, S. & Blatter, B.M. 2006. Epidemiology of work related neck and upper limb problems: psychosocial and personal risk factors (part I) and effective interventions from a bio behavioural perspective (part II). *J Occup Rehabil.* 16(3): 279–302.

Boocock, M.G., McNair, P.J., Larmer, P.J., Armstrong, B., Collier, J., Simmonds, M. & Garrett, N. 2007. Interventions for the prevention and management of neck/upper extremity musculoskeletal conditions: a systematic review. *Occup Environ Med* 64: 291–303.

Breen, A., Langworthy, J. & Bagust, J. 2005. *Improved early pain management for musculoskeletal disorders.* Sudbury Suffolk: HSE Books.

Buckle, P. & Woods, V. 2002. *Work, inequality and musculoskeletal health*. Sudbury Suffolk: HSE Books.

Centre for Strategy and Evaluation Services 2008. *Evaluation of European Campaigns and Healthy Workplace Initiative 2006 & 2007.* Retrieved from http://osha.europa.eu//en/publications/evaluation_reports/evaluation_of_european_campaigns_and_healthy_workplace_initiative_2006–2007.pdf

Council of the European Communities 1989. Council Directive 89/391/EEC of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work. *Official Journal* L 183: 1–8.

Council of the European Communities 1990a. Council Directive 90/269/EEC of 29 May 1990 on the minimum health and safety requirements for the manual handling of loads where there is a risk particularly of back injury to workers. *Official Journal* L 156: 9–13.

Council of the European Communities 1990b. Council Directive 90/270/EEC of 29 May 1990 on the minimum safety and health requirements for work with display screen equipment. *Official Journal* L 156: 14–18.

De Beeck, R.O. & Hermans, V. 2000. *Research on work-related low back disorders*. Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2005a. *Expert forecast on emerging physical risks related to occupational safety and health (OSH)*. Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2005b. *Priorities for occupational safety and health research in the EU-25*. Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2007a. *Factsheet 71 - Introduction to work-related musculoskeletal disorders*. Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2007b. *Work—related musculoskeletal disorders: Back to work report* Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2008. *Work-related musculoskeletal disorders: prevention report*. Luxembourg: Office for Official Publications of the European Communities.

European Agency for Safety and Health at Work 2010. *OSH in figures: Work-related musculoskeletal disorders in the EU—Facts and figures*.

European Foundation for the Improvement of Living and Working Conditions 2007a. *Fourth European Working Conditions Survey*. Luxembourg: Office for Official Publications of the European Communities.

European Foundation for the Improvement of Living and Working Conditions 2007b. *Managing musculoskeletal disorders.* Retrieved from http://www.eurofound.europa.eu/docs/ewco/tn0611018s/tn0611018s.pdf

Hayden, J.A., Van Tulder, M.W., Malmivaara, A.V. & Koes, B.W. 2005. Meta-analysis: Exercise therapy for non-specific low back pain. *Ann Intern Med* 142: 765–775.

Karjalainen, K., Malmivaara, A.O., Van Tulder, M., Roine, R., Jauhiainen, M., Hurri, H. & Koes, B. 2003. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. *Cochrane Database of Systematic Reviews* Issue 2. Art. No.: CD002194. DOI: 10.1002/14651858.CD002194.

Meijer, E.M., Sluiter, J.K. & Frings-Dresen, M.H.W. 2005. Evaluation of effective return-to-work treatment programmes for sick-listed patients with non-specific musculoskeletal complaints: a systematic review. *Int Arch Occup Environ Health* 78: 523–532.

Takala, J. 2000. Indicators of death, disability and disease at work. *Asian-Pacific Newsletter on Occup Health and Safety* 7(3): 4–8.

Van Tulder, M.W., Koes, B. & Malmivaara, A. 2006. Outcome of non-invasive treatment modalities on back pain: an evidence-based review. *Eur Spine J* 15 Suppl 1: 64–81.

Wahlström, J. 2005. Ergonomics, musculoskeletal disorders and computer work. *Occupational Medicine* 55: 168–176.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Musculoskeletal disorders risk assessment

Kamiel G.A. Vanwonterghem
Cergo International-Mensana, Sijsele-Damme, Belgium

ABSTRACT: In contemporary professional industrialization human health is challenged by high physical workloads due to materials handling, awkward and sustaining postures, sensory overload and repetitive (sometimes high speed) movements. Muscle overloading due to static and dynamic contractions result in a cumulative poisoning of the vasco-muscular system with serious individual, social, and economic consequences. Prevention of MSD requires a precise assessment on risks occurring in man-at-work systems.

Keywords: muscle load, risk evaluation, checklists KIM, MAC, cumulative aspects, prevention

1 BACKGROUND

MusculoSkeletal Disorders (MSDs) from occupational activities have been an increasing problem for some decades now. The incidence and severity MSD rates are serious concerns for:

– the employees: because the disorders bring direct financial losses and, perhaps more importantly, put continuing and future employment prospects at risk,
– the enterprises: because absenteeism adversely affects, directly and indirectly, production efficiency which causes costs to rise—also a negative outcome,
– the society in general: because of the socio-economic consequences on employment and the financial compensation payable for both temporary and, in the longer term, permanent injury.

The nature of musculoskeletal disorders causes specific immediate effects and side-effects:

– the long convalescence or resting periods for full recovery from MSDs can take weeks or even months the efficiency of other rehabilitation methods is very low. Resting is as efficient as kinesthetic treatments although the latter may reduce the pain/discomfort,
– the risk of the injury recurring is very high if the employee returns to the same job,
– for some injuries surgery may bring relief but it incurs a similar risk of a relapse as above,
– finally, there is the obvious potential problem arising for future employment opportunities.

The seriousness and large number of MSD problems—widely known, although in most cases largely underestimated—have led to many strategic actions and prevention policies. The scientific analysis of MSD cases delivers a long list of primary and secondary causes, mostly referring to imbalances in the "man-at-work" systems.

The registered MSD cases are proof of failures in system-design and are concrete FACTS. It is logical that in a first approach to prevention, the analysis of the recognized injuries and diseases should be a prime technique. The injury data base can provide statistically processed incidence and prevalence rates as well as evidence about the seriousness of injuries. Both may be used to support arguments to intensify preventive actions. In the author's view, such case evaluation grossly underestimates the MSD-problem because analyzing only the reality, i.e. proven injuries, is retrospective and is unable to capture an important concept in the definition of risk, namely the prospect or the "probability of an injury or loss". Basing analysis on accidents and injuries alone does not take into account the information available from "almost accidents", in airline jargon, "near misses". Omitting the "luck or bad luck"-factor is a missed opportunity to collect relevant information.

2 METHOD

The analysis of imbalances in "man-at-work" systems disclosed two types of risk assessment which depend on how the system is approached. Either one starts from the impact of "work on people", or conversely the impact of "people on work". Each of these ways uses a different method but has the same objective of identifying which factors caused the injury.

a. Logically, the first approach is to start from the known, registered MSD cases. The evaluation is then based on the analysis and interpretation of the history of preceding facts or events and prevention is focused on the working conditions to which the victim was exposed. The objective of this method is trying to find the combinations in the causal risk-structure, in particular where the problems arose, and subsequently, to eliminate the most adverse element(s).

Statistical processes to assess the validity of causes and facts must be scientifically reliable and predictive in order to propose well-founded risk reduction measures (see 2.1).

"Fitting the job to the worker" by design as re-design is technical-organizational in nature and conclusions, recommendations and solutions are common for a group of employees performing a similar job.

b. The alternative approach is "fitting people to the job" which is based on the individual human operator being THE element which integrates all work-related components. The individual perceives and evaluates the cognitively and subconsciously perceived events and conditions—including risks—and this finally, results in the operational behaviour. The balance with the workload depends on the personal capacities and skills (see 2.2).

Both technical and integrating approaches to risk management have their own specific methods. The ultimate and rather obvious goal is that the most rational prevention strategy should be based on the development of a combined methodology. This should result in an easy, simple, valid, reliable, and acceptable technique to be applied in wide range of industrial and professional settings.

2.1 *Technical risk assessment: Fitting the job to people*

The technical risk assessment should include all the work-related factors (machinery, equipment and materials, environment, organization and tasks), and should further include the interactions with the exposed workers. In principle, the work-related factors must guarantee a secure, safe and healthy job performance.

The causal factors, known from the case analysis, should, if possible, be eliminated by adjusting the work-system (for example any inappropriate lighting can be intensified, vibrations absorbed, noise absorbed, heavy loads split into smaller weights, confined spaces widened, etc…). Such a design can be realized only when taking into account the specific human characteristics and capacities (body dimensions, reaching distances), the physical capacities (force, power, energetic potentials), the sensorial perception (vision, audition, etc.), the cognitive and intellectual skills.

Many of the "external" factors and effects on people are known from field studies and laboratory tests and have been included in lists serving as tools for controlling the risks imposed on people in task-processing procedures, job organization and environmental conditions.

An almost uncountable number of checklists—in some cases also called "ergonomic" checklists—have been developed during the last decennia. MSDs, as a hot issue, were initially detected through the high prevalence of Low-Back Pain (LBP) in materials handling, pushing-pulling and lifting. Their understanding evolved together with the mechanization and automation of production processes to reduce upper limb disorders (Thoracic Outlet Syndrome), RSI (Repetitive Strain Injuries), tennis/golfers elbows, etc…

A systematic assessment of 19 of the most used checklists was recently conducted and published with an extended bibliography (Takala et al. 2010). It included well known methods such as the AET (1979) general workstation analysis, the NIOSH equation (1981–1991) relating to materials handling, OWAS (1974, 1996) to assess the body posture, and OCRA (2002) (this provided the basis for ISO 11228-3 on repetitive work).

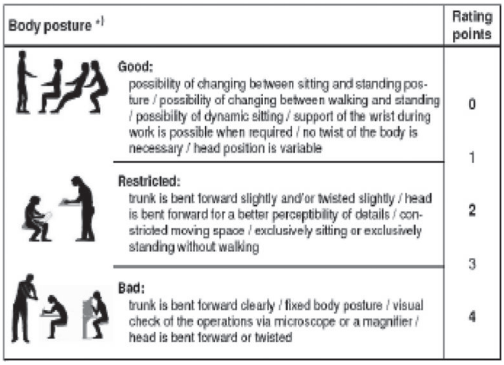
Some of the checklists served as bases for guidelines or standards. More recent methods—developed for specific activities or risks—have been included by the European Agency for Safety and Health in their Risk Assessment Campaign, namely—the Key Indicator Method—KIM. (see 2.1.1) and the Manual Handling Assessment Charts—MAC (2.1.2.), launched in 2002. The main objectives of both are reducing MSDs in "material handling".

2.1.1 *KIM—Key Indicator Method manual handling operations*

The Key Indicator Method for Manual Handling Operations is a German method compiled from an intensive scientific literature survey (1994) in the context of the Committee for Safety Work Hygiene and Health protection at Work to advise the European Commission on upper limb disorders and RSI (Steinberg et al. 2008).

For manual handling tasks: lifting, holding, carrying and for pushing and pulling, KIM includes a series of work-related factors such as cycle time, movement frequency, level of action force, mode of force transfer, type of grips, work organization, working environment, both posture and hand-arm position. Working time is a multiplier factor.

The load-specific hazards for physical overload factors are quantified in individual scales rated in a range from minimum/optimal to maximal/unacceptable load. An example of the table showing postural load-scales is given in Figure 1.



Body posture *)	Rating points
Good: possibility of changing between sitting and standing posture / possibility of changing between walking and standing / possibility of dynamic sitting / support of the wrist during work is possible when required / no twist of the body is necessary / head position is variable	0
	1
Restricted: trunk is bent forward slightly and/or twisted slightly / head is bent forward for a better perceptibility of details / constricted moving space / exclusively sitting or exclusively standing without walking	2
	3
Bad: trunk is bent forward clearly / fixed body posture / visual check of the operations via microscope or a magnifier / head is bent forward or twisted	4

Figure 1. KIM—Evaluation body posture.

On the basis of the rating points calculated and the table below it is possible to make a rough evaluation.

Risk range [2)]	Risk score	Description
1	< 10	Low load situation, physical overload unlikely to appear.
2	10 to < 25	Increased load situation, physical overload is possible for less resilient persons[3)]. For that group redesign of workplace is helpful.
3	25 to < 50	Highly increased load situation, physical overload also possible for normally resilient persons. Redesign of workplace is recommended.
4	x 50	High load situation, physical overload is likely to appear. Workplace redesign is necessary.

2) The boundaries between the risk ranges are fluid because of the individual working techniques and performance conditions. The classification may therefore only be regarded as an **orientation aid**. Basically it must be assumed that as the number of risk scores rises, so the risk of overloading the musculo-skeletal system increases.
3) Less resilient persons in this context are persons older than 40 or younger than 21 years, newcomers in the job or people suffering from illness.

Published by: Federal Institute for Occupational Safety and Health and Committee of the Laender for Occupational Safety and Health (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Postfach 17 02 02, D - 44061 Dortmund and Länderausschuss für Arbeitsschutz und Sicherheitstechnik (LASI), Franz-Josef-Roeder-Str. 23, D - 66119 Saarbrücken)

Figure 2. Total KIM score-evaluation.

The final outcome of the KIM assessment is a sum of the rating points relevant to the activity performed, multiplied by the time rating points and adjusted to gender (female = *1.3) . The total score is then evaluated by a table indicating the risks and recommendations for action (Figure 2).

Example: a) for moving 1 to <4 km (score duration = 4), b) use of rails, hand carts, fixed rollers (score 1.5) and c) speed = fast (score 2), d) posture slightly bending forward (score = 2), working conditions restricted (score = 2). Sum (a, b, c, d = 1.5 + 2 + 2 + 2) = 7.4 × 4 = 30.

For example when the total score is 30 re-design is recommended (See Figure 2).

2.1.2 *Manual handling Assessment Charts (MAC)*

The British HSE (Health and Safety Executive) developed an assessment tool in 2002 to have an easy and fast to use tool for detecting high risks in material handling (lifting, carrying and team handling) (Monnington et al. 2002) followed in 2007 for registering pulling-pushing risks (Ferrera & Smith 2007). A classic "traffic light" risk indication color (green for safe, red for danger) is used to classify the risks.

The tasks are split into operations specific for lifting, carrying, handling as part of a team and pushing-pulling. Minimum and maximum scores are defined from an extensive literature survey. Each zoning is defined by a specific load value (Figure 3).

For example: combinations of weight and lifting frequency 0 = safe (good), orange 4 (reasonable), red 6 (poor, danger), and purple = 10 overexertion.

These values are presented in an eye-catching summary sheet, color or load digits for the selected type of work and operations, indicating the "hot spots" which require priorities for action (Figure 4).

2.2 *Integrated risk assessment*

As the technical checklists are mainly evaluated sequentially the mathematical averaging is not a correct method for the "integration" of a total load. There is only one integration tool which combines

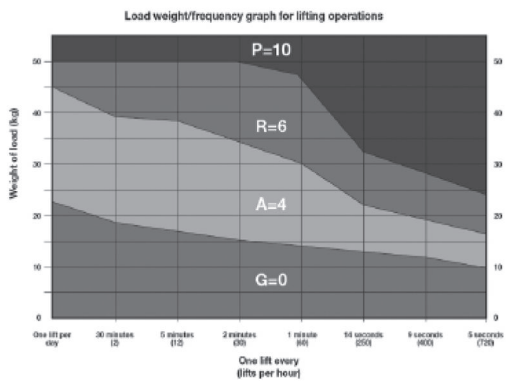


Figure 3. MAC—Load quoting weight and lifts/hour.

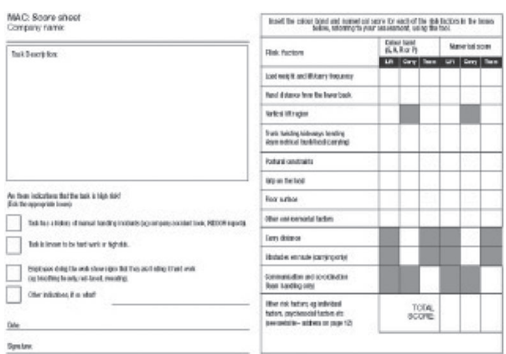


Figure 4. Summarizing table MAC-scores.

load and capacity, the endo—and exogenous factors: the human operator.

The operator evaluates the occupational risks via cognitive and subconscious processes—mainly steered by subjective experiences such as nuisance, annoyance, fatigue and pain—resulting in a final outcome of safe or risky behavior.

The strictly individual decision making and developed coping strategy depends on learnt complex signals which often precede the evidence of imbalances.

Because of the fast advancement in new working systems in which the human operator faces the challenge of coping with new less subtle risks—indeed the heavy physical work load may be replaced by high-speed repetitive work, adverse body postures, mental strain—a quasi-permanent adjustment of the behaviour becomes essential.

The operator has to learn to recognize subjective pre-signals indicative of a failing coping strategy (e.g. fatigue, annoyance, pain symptoms) as they occur during work in order to anticipate the critical moment in the evolution towards an injury or disease. As the operators are confronted with cumulative aspects over longer time such as general and local fatigue, they should adjust their behavior based on a suspicion that possibly had evidence.

At present MSDs caused by "over exertion" tend to evolve to a CTDs (Cumulative Trauma Disorders) in which muscle strain, due to postural static contraction and repetitive work of small muscle groups, evolves to a critical acidosis. Via this slow cumulative (poisoning) process muscles, tendons, joints and bones will be affected, exploding in an sudden occurring injury once a critical level is reached. A traditional evolution towards an MSD/CTD (Vanwonterghem 2004) is drawn in Figure 5 in which 3 cases are represented: 1 most of the cases evolve asymptotically to an injury over a cumulative periods varying from 9 months to 6–8 years. (line 1). Some individuals develop their own coping strategy and can perform the job over many months or years (dotted line 2) whereas others evolve in less than one year (3-dot/dash) due to overexertion. (Note that in this case many persons denied the subjective adverse effects because of an inappropriate motivation).

The 4 indicated phases cover i) an adaptation phase (few months, in which the body is adapting itself to the muscle requirements), ii) an adapted phase (body is able to do the job without problems), iii) a cumulative phase (which can vary from a few months up to 6–8 years) and iv) a critical phase (very short: days, weeks …) when the poisoning effect evolves asymptotically to an injury/disease. The loss in force/power shows an almost mirror image to these graph-fluctuations. The declining

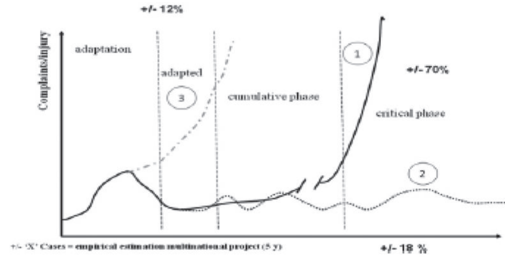


Figure 5. MSD evolution.

tendency after starting can be explained by the muscle training effect achieved during work.

This evidence-based model emphasizes the need for a prevention strategy which anticipates the state of the system before the progression has become critical. This can be obtained only by a method which includes the operators' expressions of complaints and experiences and the assessment information contains not only the objective criteria (technical risk assessment) but also the "subjective criteria" to accommodate and reflect the integrated stress-strain level.

The methods available to collect the subjective load are almost as plentiful as the technical checklists, but are based on questionnaire and interview techniques.

The sensorial perception and information processing are integrated with specific human characteristics such as motivation, social behavior etc. of which some are directly linked to workload, others more indirectly to the individual. Integrated risk-assessment methods should concentrate on the outcome at the workplace and should—to increase confidence levels—give confirmation of the objective measurements.

Questionnaires such as SWAT (Subjective Workload Assessment Technique) and NASA-TLX (Task Loads Index) are focused on "mental workload" (pilots) which have only an indirect effect on the CTDs where in the Borg CR10 (category ratio), "fatigue" is assessed on a 10 point scale (Borg 1970).

The SWI (Subjective workload Index) also uses the 0–10 scale for "no problems" to "unbearable" (Vanwonterghem 1985). It is used for fatigue, concentration, task complexity, work rhythm and problems with responsibility. In addition two factors which can be estimated as "compensating" factors are evaluated as well: interest in the job and the degree of autonomy. The SWI-index considers, in detail, all aspects of the technical risks (posture, movements, environment etc.) and is weighted to the tasks' exposure time and subjective rating.

Both Borg's CR10 and SWI are more appropriate to assess the MSD/CTD.

3 CONCLUSIONS

The technical risk assessment checklists have undoubtedly many merits for those concerned with the adjustment of workload to the human capacities (design and control) and refining the methods should continue as new evidence becomes available.

The merits of guidelines lie in *designing jobs for groups of employees* if build on scientific rules of statistical validity and probability as well as on good advices and experiences.

However, in MSD there is an aspect which cannot be predicted from external observations or from scientific studies of groups. This is because of the inter—and intra-individual differences which are inherent in the functioning of an individual and these may weaken, or even destroy, statistical relevance.

The coping process of individuals (cognitively and subconsciously) results in the operational behaviour being, to a large extent, determined by the individually perceived discomfort and personal evaluation of the working conditions.

Finally, an efficient risk assessment should call upon both principles: a technical evaluation integrated with subjective experiences. This will deliver a new strategy in which the objective screening of the critical jobs (tasks, organization and environment) and the subjective assessments of the exposed personnel will be able to re-evaluate each design (or in most cases re-design) on better evidence-based information.

REFERENCES

Borg, G. 1970. Perceived exertion as an indicator of somatic stress. *Scandinavian journal of Rehabilitation Medicine* 2(2), 92–98.

Ferrera, J. & Smith, M. 2007. *Evaluating the feasibility of developing assessment charts for high risk pushing and pulling operation*. HSL-Report RR562.

Monnington, S.C. et al. 2002. *Development of an inspection tool formanual handling risk assessment*. HSL ERG/02/20 Report. Health & Safety Laboratory, Sheffield.

Steinberg, U. et al. 2008. *Key Indicator Method Manual Handling Operations design and testing aid for assessing working conditions*. Research Report F-1994. Bundesanstalt für Arbeitsschutz une Arbeitsmedizin DE44149 Dortmund (Germany).

Takala, E.-P. et al. 2010. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scand. J Work Environ Health* 36(1): 3–24.

Vanwonterghem, K. 1985. Validation of a checklist composed of different evaluation methods by means of ergonomic principles. Cahiers Ecotra, Univ. Geneva. (in French).

Vanwonterghem, K. 2004. Considerations about results and Project team work. In Strambi F. et al. (eds), *Risk assessment of biomechanical damage risks in small and medium-sized enterprises*. Moretti & Vitali: 155–165.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Checklists for WMSD hazard evaluation

Yen-Hui Lin
School of Occupational Safety and Health, Chung Shan Medical University, Taichung, Taiwan

Chih-Yong Chen & Yi-Tsong Pan
Institute of Occupational Safety and Health, Council of Labor Affairs, Executive Yuan, Taiwan

ABSTRACT: This study explores the suitability of the OSHA MusculoSkeletal Disorders (MSDs) checklist for use in the semiconductor manufacturing industry. Six companies were enrolled in the current study, and 122 subjects were evaluated through a questionnaire survey and field observation using the OSHA MSDs checklist. Experimental results demonstrated that nearly 50% of subjects reported physical discomfort, with shoulder symptoms (38.5%) being the most common complaint. Additionally, the aggregated sensitivity, specificity, and positive predictive values for the MSDs checklist in the six companies were 47.3%, 62.4% and 30.0%, respectively. This study concluded that the OSHA MSDs checklist appeared to contain reliable indicators for capturing musculoskeletal discomfort, and found that the OSHA MSDs checklist provided an easily administered, proactive surveillance instrument to assist in early identification of musculoskeletal discomfort.

Keywords: ergonomic checklist, musculoskeletal disorders, questionnaire

1 INTRODUCTION

Semiconductor manufacturing is a leading high-tech industry in Taiwan. The workforce in this industry has increased significantly during the past two decades, and the technological and economic benefits associated with the industry have been accompanied by growing scientific and public concern for worker health and safety (McCurdy, et al., 1989). Numerous studies have already reported the prevalence of musculoskeletal disorders and Cumulative Trauma Disorders (CTD) in semiconductor manufacturing workers (Wu, et al., 2009; Chee & Rampal, 2004; Li, et al., 1997). Work-related risk factors in the semiconductor-related manufacturing industry may include awkward postures and long working hours, prolonged standing and frequent walking, high production demands, poor equipment design and repetitive wafer-handling activities (Chee & Rampal, 2004; Yu, et al., 2003).

In numerous industries or enterprises, ergonomic checklists have been designed to identify musculoskeletal injuries resulting from awkward posture or highly repetitive motions (Armstrong, et al., 2001; Kerserling, et al., 1993; McAtamney & Corlett, 1993; OSHA, 1995). Although these checklists are useful in identifying actual and potential risks of musculoskeletal discomfort in factories (Armstrong, et al., 2001; Lu, et al., 1999), the data on musculoskeletal discomfort among semiconductor manufacturing workers is limited. Therefore, semiconductor manufacturers still need to know the actual effects of a proposed ergonomic checklist to judge the confidence of usefulness. This study presents an empirical study examining the diagnostic effects of the MSD checklist in the semiconductor-related manufacturing industry.

2 METHODS

2.1 *Study subjects*

Study subjects comprised full-time workers in clean rooms of firms in semiconductor and related industries in Taiwan. A total of 122 subjects, from six different companies, participated voluntarily. Subjects in each selected company were responsible for wafer container carrying, transporting, loading and unloading tasks among various wafer processing tools and storage racks.

2.2 *Procedures*

The research procedures comprised four steps, including (1) ergonomic-based checklist training course, (2) MSD questionnaire survey, (3) MSDs checklist use, and (4) data analysis.

2.2.1 *The ergonomic-based checklist training course*

Six trainees (observers), from six different companies, were enrolled in this study. All trainees were females and their main tasks were handling questions related to occupational safety and health issues, including health promotion, health care and ergonomic interventions. The training course comprised six 6-h workshops, conducted over one week (36 hours).

The training courses were designed to: (1) identify the risk factors of MSD, (2) clarify ergonomic principles for reducing MSD, and (3) provide practice in using checklists to evaluate MSD of workspaces. The training course began by presenting work-related risk factors for the onset of identified MSDs and ergonomic principles for reducing MSD for semiconductor-related manufacturing workers. Second, the disciplines of occupational biomechanics, work physiology, engineering psychology, applied anthropometry, and work-space design were provided. Third, practice was provided in using the checklist for workspace MSD evaluation.

2.2.2 *MSD questionnaire survey*

A Nordic Musculoskeletal Questionnaire (NMQ) survey was employed to obtain information from the selected subjects. Furthermore, 122 subjects (by observers) from six different companies participated voluntarily. Before the survey, all subjects were informed of the study objectives and participated voluntarily.

2.2.3 *MusculoSkeletal disorders checklist*

Musculoskeletal discomforts were assessed using a Chinese version MSDs checklist, revised from the MusculoSkeletal Disorders (MSDs) checklist by the Occupational Safety and Health Administration (1995). The MSDs checklist comprises three parts: Part A checks risk factors associated with upper extremity problems; Part B checks risk factors associated with back pain and lower extremity disorders; and Part C evaluates risk factors for manual material handling tasks. Action is taken if five risk score are identified. If the evaluated risk score of Part A equals or exceeds five score, then ergonomic interventions should be performed to protect workers from upper extremity problems. Meanwhile, if the sum of the risk score of Parts B and C equals or exceeds five, then improvements must be implemented to protect workers against back pain and lower extremity disorders.

2.2.4 *Data analysis*

All analyses were performed using SPSS Release 11.5.0 (SPSS Institute, 2002). First, descriptive statistics were calculated for all the questionnaire responses. Next, based on the results of the NMQ questionnaire survey, this study ascertained the ability of this MSDs checklist to accurately categorize individuals, and determined the sensitivity, specificity, and positive predictive value indexes using the formulations of Marley & Kumar (1996).

3 RESULTS

3.1 *Characteristics of the study population*

Table 1 lists the study population characteristics. Male (48/122) and female (74/122) workers comprised 39.3% and 60.7% of subjects, respectively. Nearly two-thirds (61.5%) of subjects were in the 20–29 years old age group. Moreover, 57.4% of subjects had worked for less than 5 years in the semiconductor industry, and over 85% of subjects had worked in their present company for less than 5 years. As for weekly time spent on physical exercise, 20.5% of subjects reported engaging in physical exercise 'often'. Most subjects were non-smokers (77.0%), and right-handed (90.2%).

3.2 *Physical discomfort prevalence*

Table 1 also lists the percentage distributions of physical symptoms. Overall, nearly 50% of the subjects reported physical discomfort. The prevalence of upper extremity discomfort was 46.7%, with shoulder symptoms (38.5%) being most common, followed by hand or wrist (36.9%), neck (34.4%), and elbow (16.4%). Meanwhile, the prevalence of lower extremity or back discomfort (48.4%) was slightly higher than that of upper extremity discomfort, with the most common symptoms being lower back or hip pain (36.1%), followed by the lower leg (24.6%), knee (12.3%), and thigh (11.5%).

3.3 *MSDs checklist analysis and sensitivity, specificity calculation*

The sensitivity, specificity, and positive predictive value of the checklist were considered based on the results of the NMQ questionnaire survey. Table 2 lists the sensitivity, specificity, and positive predictive value for the MSDs checklist. Initially, the aggregated rates of sensitivity, specificity, and positive predictive value for the MSDs checklist for all companies were calculated. Overall, taking the results of the NMQ questionnaire as the standard, the average sensitivity, specificity, and positive predictive values for the MSDs checklist in the six companies were 47.3%, 62.4% and 30.0%, respectively. Furthermore, to reveal the significant

Table 1. Demographic characteristics and prevalence of physical discomforts.

	Number	Percentage (%)
Gender		
Male	48	39.3
Female	74	60.7
Age (years)		
20–29	75	61.5
30–39	38	31.1
40–49	6	4.9
>50	3	2.5
Work experience		
<1 year	16	13.1
1–5 years	54	44.3
6–10 years	32	26.2
11–20 years	20	16.4
Working experience of current job		
<1 year	35	28.7
1–5 years	69	56.6
6–10 years	15	12.3
11–20 years	3	2.4
Weekly physical exercise time		
Seldom	33	27.0
Sometimes	63	51.6
Often	25	20.5
Smoking		
Seldom	94	77.0
Sometimes	5	4.1
Often	23	18.9
Handedness		
Right-hand	110	90.2
Left-hand	12	9.8
Physical discomfort		
Neck	42	34.4
Shoulder	47	38.5
Elbow	20	16.4
Hand or wrist	45	36.9
Lower back or hip	44	36.1
Thigh	14	11.5
Knee	15	12.3
Lower leg	30	24.6
Summary		
Upper extremity	57	46.7
Lower extremity or back	59	48.4

differences in sensitivity, specificity, and positive predictive value among individual companies, all cases were disaggregated by company. The disaggregated rate of each variable was calculated exactly as the aggregated rate. Table 2 also lists the percentage sensitivity, specificity, and positive predictive value for the MSDs checklist among the six companies. From Table 2, the sensitivity, specificity, and positive predictive value for the MSDs checklist in the six companies ranged from 26.5% to 74.2%, 42.2% to 86.4%, and 10.9% to 46.9%, respectively.

Table 2. Sensitivity, specificity and predictive value by MSDs checklist and NMQ in six companies.

Company	Sensitivity	Specificity	Positive predictive value
A	26.5%	86.4%	37.5%
B	40.7%	65.8%	44.6%
C	60.5%	66.2%	31.1%
D	53.8%	42.2%	10.9%
E	74.2%	56.0%	46.9%
F	44.4%	63.0%	31.8%
Total	47.3%	62.4%	30.0%

Sensitivity = Number of true positives/Number of true positives + Number of false negatives.
Specificity = Number of true negatives/Number of true negatives + Number of false positives.
Positive predictive value = Number of true positives/Number of true positives + Number of false positives.

4 DISCUSSION

4.1 *Prevalence of physical discomforts*

The proportions of discomfort ratings for various body parts were compared with statistics from the Institute of Occupational Safety and Health (IOSH) of Taiwan (2002) of all industries. The prevalences of discomfort in the hand or wrist (36.9% in this study versus 26.6% for IOSH), lower back or hip (36.1% in this study versus 33.8% for IOSH), and lower leg (24.6% in this study versus 18.2% for IOSH) significantly exceeded the average prevalence for all industries as reported by IOSH. The overall prevalence of upper extremity discomfort, including that involving the neck, shoulder, elbow, and hand/wrist, was 46.7% in this study (Table 1). This prevalence supports the finding of Hsu & Wang (2003) regarding upper extremity discomfort (42%) for semiconductor workers. Furthermore, the finding that shoulder symptoms (38.5%) were the most frequently reported form of physical discomfort in the semiconductor industry was consistent with Li et al. (1997).

Although the questionnaire survey results identified a high prevalence of musculoskeletal discomforts in the semiconductor manufacturing industry, the findings are limited because they are based on personal subjective judgments. As Hsu & Wang (2003) reported in an earlier study, the prevalence rates based on a questionnaire alone tend to exceed those based on questionnaires with physical examinations. Accordingly, the questionnaire survey alone is insufficient for determining the prevalence rates. The evidence would be more convincing if objective measures, including identified risk factor ratio and prevalence of MSDs, could support the findings of the questionnaire survey.

4.2 *The sensitivity, specificity and positive predictive value of MSDs checklist*

The findings regarding the sensitivity and specificity were 47.3% and 62.4% for the MSDs checklist in the semiconductor manufacturing industry, respectively, slightly less than the finding of Lin et al. (1999) regarding the sensitivity (53.7%) and specificity (65.4%) for the MSDs checklist in the manufacturing industry. Therefore, the MSDs checklist appeared to provide a reliable indicator for capturing musculoskeletal discomfort. This phenomenon indicated that the MSDs checklist could be utilized as an easily administered, proactive surveillance instrument to assist in early identification of musculoskeletal discomfort.

5 CONCLUSION

This study found that work-related musculoskeletal discomforts are common in the semiconductor manufacturing industry. The prevalence of upper extremity, and lower extremity or back discomforts were 46.7%, and 48.4%, respectively. Although the analytical results also indicated that the MSDs checklist is a sensitive and useful tool for identifying musculoskeletal discomfort in semiconductor manufacturing workers, the findings are somewhat limited because the sensitivity, specificity, and positive predictive value measurements are based on personal subjective judgments (i.e. questionnaire surveys). Future studies thus measure musculoskeletal discomfort based on medical tests.

REFERENCES

Armstrong, T.J., Franzblau, A., Haig, A., Keyserling, W.M., Levine, S., Streilein, K., Ulin, S. & Werner, R. 2001. Developing ergonomic solutions for prevention of musculoskeletal disorder disability. Assist Technol 13: 78–87.

Chee, H.L. & Rampal, K.G. 2004. Work-related musculoskeletal problems among women workers in the semiconductor industry in peninsular Malaysia. International Journal of Occupational and Environmental Health 10: 63–71.

Hsu, W.H. & Wang, M.J. 2003. Physical discomfort among visual display terminal users in a semiconductor manufacturing company: A study of prevalence and relation to psychosocial and physical/ergonomic factors. AIHA 64: 276–282.

Institute of Occupational Safety and Health of Taiwan 2002. Survey of employees' perception of safety and health in the work environment in 2001 Taiwan (in Chinese).

Keyserling, W.M., Stetson, B.A., Silverstein, B.A. & Brower, M.L. 1993. A checklist for evaluating ergonomic risk factors associated with upper extremity cumulative trauma disorders. Ergonomics 367: 807–831.

Li, K.W., Yeh, W.Y., Chu, J.C. & Chen, C.Y. 1997. Investigation of current status or cumulative trauma disorders in semiconductor manufacturing industry. Journal of Occupational Safety and Health 5: 1–14 (in Chinese).

Lin, Y.H., Yeh, W.Y. & Chen, C.Y. 1999. A study on the suitability of work-site ergonomic checklists. Journal of Occupational Safety and Health 7: 45–60 (in Chinese).

Lu, C.W., Chang, K.F., Chang, C.H., Lin, Y.H. & Chen, H.G. 2005. The suitability for the BRIEF applied in the steel industry. Asian Journal of Ergonomics 6: 97–114.

Marley, R.J. & Kumar, N. 1996. An improved musculoskeletal discomfort assessment tool. International Journal of Industrial Ergonomics 17: 21–27.

McAtamney, L. & Corlett, E.N. 1993. RULA: a survey method for the investigation of work related upper limb disorders. Applied Ergonomics 24: 91–99.

McCurdy, S.A., Schenker, M.B. & Lassiter, D.V. 1989. Occupational injury and illness in the semiconductor manufacturing industry. American Journal of Industrial Medicine 15: 499–510.

Occupational Safety and Health Administration (OSHA) 1995. Draft: Instructions for completing the risk factor checklist. OSHA: 24–32.

SPSS Institute, Inc. 2002. SPSS user's guide, Release 11.5.0.

Wu, H.C., Chen, H.C. & Chen, T. 2009. Effects of ergonomics-based wafer-handling training on reduction in musculoskeletal disorders among wafer handlers. International Journal of Industrial Ergonomics 39: 127–132.

Yu, C.Y., Chen, Y.C., Yeh, W.Y. & Lin, Y.H. 2003. A preliminary investigation of the causes of musculoskeletal fatigue in semiconductor workers. Journal of Occupational Safety and Health 11: 1–18 (in Chinese).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Ergonomic practices for packing tasks in a printing ink manufacturing factory

Hsin-Chieh Wu & Peng-Cheng Sung
Chaoyang University of Technology, Wufong, Taichung County, Taiwan

Yi-Tsong Pan
Institute of Occupational Safety and Health, Xizhi, Taipei County, Taiwan

ABSTRACT: The purpose of this study was to assess the operational workloads for packing tasks in a printing ink manufacturing factory located in northern Taiwan, with an aim to lower the risk of MusculoSkeletal Disorders (MSDs) among the workers. The packing task operators pack 2000 1-kg ink cans per day on average, with sets of 12 ink cans packed into paperboard boxes. The Key Indicator Method (KIM) was used to evaluate and identify the MSD risk factors associated with the ink packing tasks. The data collected include material weights, work posture, work pace, vertical distances, vertical lifting displacement, and lifting frequency. The mean lifting frequency was 4166 lifts/day and the average lifting weight was 1.44 kg/lift. The KIM risk score of the packing task was evaluated as 40, indicating a risk level of 3. To improve working conditions and reduce health and safety risks for this packing task, a 70 cm height worktable was suggested to replace the original 39 cm height worktable. This simple intervention can allow the operators to move the ink cans more smoothly and efficiently without bending their backs as they pack the cans. In addition, the estimated operator risk score was reduced to 24, indicating a risk level of 2. This suggested change has been adopted by the ink company and implemented for their packing tasks.

Keywords: manual lifting tasks, workstation design, ergonomic risk identification

1 INTRODUCTION

According to occupational injury records, MusculoSkeletal Disorders (MSDs) in Taiwan account for 33% of total work injuries in recent years (Taiwan IOSH 2005). Musculoskeletal disorder is also the most widespread occupational disease in Taiwan (Fig. 1) and the related medical loss of NT$ 2 billion was approximately 0.67% of Taiwan's GDP for the year 2009 (Taiwan IOSH 2010).

The occurrence of MSDs is usually a result of factors such as excessive force, repetitive movements, awkward postures, prolonged exposure, and exposure to low temperature or vibration (Putz-Anderson 1988). The body parts most often affected by MSDs are the lower back, neck, and upper extremities (shoulder, arm, and wrist). MSDs not only endanger worker's health but also reduce their productivity. Therefore, addressing the lowering the risks for developing MSDs of the workers is crucial to improving labor health and safety and increasing business productivity.

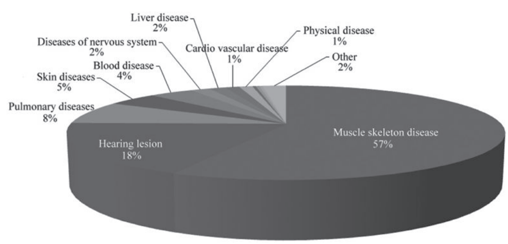


Figure 1. Occupational disease category distribution in 2009.

Several ergonomic interventions, such as workstation redesign, employee training, and working condition improvement, have been reported to mitigate risk factors causing MSDs. This paper presents a case study in which risk factors of MSDs in an ink manufacturing company were evaluated and identified. In addition, simple interventions were also recommended and implemented to enhance the health, safety, and productivity of company employees.

2 METHOD

2.1 *Ergonomic assessment tools*

This research used the following two checklists to assess the MSDs risks at the worksite of a printing ink manufacturing plant located in northern Taiwan.

2.1.1 *Baseline Risk Identification of Ergonomic Factors (BRIEF)*

The BRIEF checklist was designed to quickly identify occupational engineering risk factors in the workplace (Humantech Inc. 1993). This checklist can evaluate six body parts: hands and wrists, elbows, shoulders (the above three parts, all have right and left sides) along with neck, back, and legs. The main risk factors evaluated by the BRIEF tool are force, posture, repetitiveness, and work duration. Tasks associated with other risk factors, such as exposure to vibration, mechanical pressure, and low temperature, should be referred to professional personnel for further analyses. According to the guidelines of the BRIEF method, if two or more risk factors are identified for a given body part, this body part has a strong possibility of developing MSD.

2.1.2 *Key Indicators Method (KIM)*

KIM was developed by Germany's Federal Institute for Occupational Safety and Health (BAuA) and Committee of the Länder for Occupational Safety and Health (LASI). It has been adopted and used by several other EU countries for risk assessment of manual load handling (SLIC 2010). This method evaluates relevant activity data called key indicators. The indicators include duration, frequency and distance with a multiplier for weight, posture and working conditions. The method is quite suitable for application to industrial practice. KIM simplifies the assessment procedure into the three following steps:

1. Determine the time rating point: Identify the task as a lifting, holding, or carrying task and then determine the corresponding time rating point from the KIM rating table, according to the number of daily lifts, the holding duration, or the carrying distance.
2. Determine the rating points for load, posture, and working conditions: With the KIM rating table, look up the load rating points for male or female operators that correspond to the effective load (the real action force that is necessary for moving the load). Identify the typical working posture and load position of the task in order to determine the rating points for posture. The rating points for working conditions account for the presence of environmental hazards such as workspace limitation, physical obstacles, uneven or unsteady flooring, inadequate lighting, or poor gripping conditions, etc.
3. Compute risk score and determine risk level: Compute the total risk score of the task by multiplying the time rating point by the sum of the rating points for load, posture, and working conditions. Measure the task's risk level by consulting the KIM defined risk score range (level 1: risk score <10; level 2: 10≤ risk score <25; level 3: 25≤ risk score <50; level 4: risk score ≥50).

This classification of the risk scores gives an indication of any load bottlenecks.

2.2 *Ergonomics investigation process*

The ergonomics evaluation process included a formal meeting and a tour of the shop floor (onsite inspection). In the meeting, the chief inspector first introduced the ergonomics investigation process and then dialogued with the employers and major crewmembers to gather an overview of the work and the key problems experienced by the inspected company. During the tour, inspectors collected field data of specific tasks by talking with individual operators, taking photographic and video evidence of manual materials handling tasks, and measuring task demands and physical dimensions of inspected workstations. BRIEF and KIM tools that enabled the quantification of data were used for explaining problems to management. The inspectors drafted and sent the company an inspection report with adjustment suggestions, and a follow-up meeting was arranged afterward to discuss the efficacy of the suggested adjustments.

3 ERGONOMIC PRACTICES

3.1 *Ink packing tasks*

In the factory's ink packing station, operators filled ink into 1-kg cans and packed every 12 of those ink cans into a paperboard box. The working area was about 33 m^2, with a mixer barrel, a filtering tank (Fig. 2), a capping machine on a worktable (Fig. 3), and paperboard boxes on pallets. The filtering tank was fixed on a steel frame (120-cm from the floor) of 50-cm height and 50-cm diameter. The dimensions of the capping machine were 52-, 46-, 60-cm in length, width, and height, respectively, and it was situated on a worktable of 150-, 70-, and 54-cm in length, width, and height, respectively. The dimensions of the pallets and paperboard boxes used for packing ink cans were 1.2 m × 1.2 m × 15 cm and 53 cm × 40 cm × 16 cm in length, width, and height, respectively. A team of

Figure 2. Ink filtering tank.

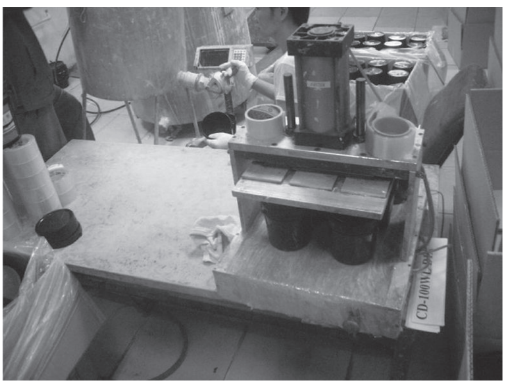


Figure 3. Capping machine.

2 to 3 operators, aged from 40 to 50 years, was employed to complete the ink packing tasks.

The shop floor tour inspected the ink filtering, can filling and weighting, and paperboard box labeling procedures. When the semi-finished material in the mixing barrel was pumped into the filtering tank, one operator stirred the ink in the filtering tank with a stick (Fig. 4) to reduce the viscosity of the ink. Then, the second operator sitting in front of the filtering tank filled the ink can to 1 kg weight. After filling, the second operator moved the can onto the table at her left hand side. Then, the third operator cleaned the can body, capped the can using the capping machine, packing the cans into the paperboard box, and piled the paperboard box on the pallet.

During the stirring operation, the worker was standing with one hand above his shoulder level.

Figure 4. Work environment and working postures of operators performing stirring and filling tasks.

For the filling and weighing operations, the worker was sitting with neck flexed to gaze the weight scale and to control the ink filling valve at the bottom of the filtering tank. The third operator packed the cans into the paperboard boxes with a flexed trunk.

3.2 *Problems statements*

The operator who stirred the ink with an over shoulder posture could encounter muscle fatigue in his upper limbs from prolonged operation. Nevertheless, this action of reducing ink viscosity was not frequent. Therefore, investigators did not suggest any amendment to this task.

The operator conducted the filling task with her neck flexed looking at the weight scale. This static and non-neutral posture could lead to muscle fatigue of the neck and shoulders. Moreover, in order to accurately filling 1-kg of ink into each can, the operator repetitively flexed and extended her dominant wrist to control the outlet valve of the ink (Fig. 4). According to BRIEF tool analysis, the operator had a risk score of 2 at her wrist due to a radial deviation and holding duration >10 sec; a risk score of 3 at her shoulder due to an elevation angle >45°, and maintaining an award posture >10 sec and 2 times/min; a risk score of 3 at her neck and back due to maintaining a forward bending and twisting posture >10 sec and 2 times/min. The operator had regional risk score at wrist, shoulder, neck, and back all greater than 2, indicating high probability of musculoskeletal disorders and need for working-posture adjustment. For the same operator, KIM assessment showed that lifting a 1-kg can for 2000 times/day with slightly bending and twisting trunk had rating points of 10 (time), 1 (load), and 2 (posture). Due to movement confinement caused by the workspace arrangement, the rating point of the work environment was estimated as 1. Her total KIM risk score was computed as 40 (risk level 4), indicating a highly increased load situation, with physical overload possible for the average persons. Hence redesign of the workstation was recommended according to KIM assessment.

The third operator made repetitive lifting movements: packing the paperboard boxes with cans and laying them on pallets. He handled 2000 1-kg cans per day, on average, and lifted and laid 12-kg boxes on pallets over 167 times per day. KIM assessment showed that his task risk score was 24 (risk level 2), indicating a possibility of physical overload only for a weaker person. Therefore, investigators did not suggest any adjustment to this task.

3.3 *Ergonomics interventions*

The inadequate table height (57 cm) caused the operators to adopt awkward postures. A new workstation (Fig. 5) was thus devised to reduce stress on the operators due to inappropriate working posture. The newly designed workstation was L-shaped with 70 cm table height; the base of the mixer tank was raised from 70 cm to 140 cm accordingly. Two fillisters, one on each end of the table, were designed to position the weight scale and capping machine, respectively. The sunken depth of each fillister was designed to keep the surface of the scale and the capping machine at the same level of the table surface so that the ink cans could be moved easily on the surface level (Fig. 6). The capping machine was rotated 90° from it original orientation to have its entry aligned with the moving path of the cans.

The new worktable design allows operators to push cans horizontally to the capping machine after filling them, without any lifting and laying activities (Fig. 7). After workstation adjustment, the weight scale was elevated to a level that allows the operator to read its value comfortably. Furthermore, commercially available tilting devices, located at both ends of the worktable, could be used to raise and tilt a container or paperboard box. This arrangement could facilitate operators taking out empty cans from the container or putting filled cans into the paperboard boxes and sealing them (Fig. 8).

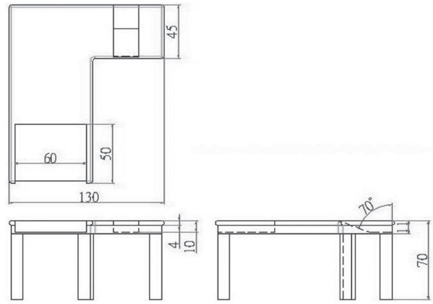


Figure 5. The new worktable design.

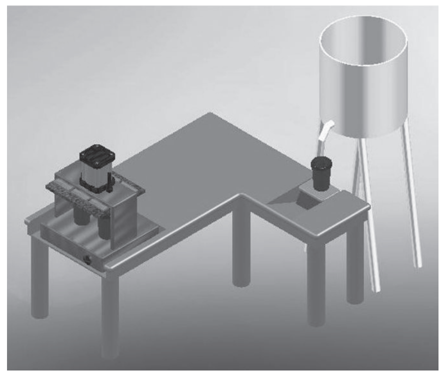


Figure 6. Fillister design for the weight scale and capping machine.

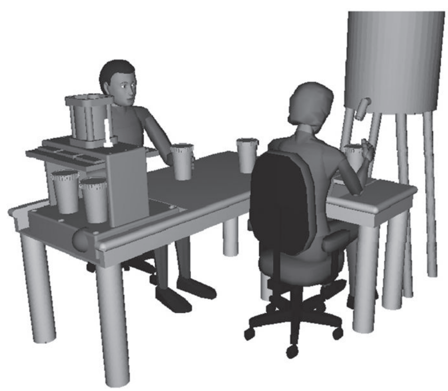


Figure 7. Workstation redesign and suggested adjustments to packing tasks.

3.4 *Efficacy of the ergonomics interventions*

The proposed workstation redesign can improve work postures for the filling, weighing, and packing operators. The use of tilting devices can further prevent operators from bending their trunks and picking up empty cans from floor level. In addition, the muscle strain in their waists and upper extremities can be reduced. According to BRIEF analysis, regional risk scores can be reduced from 2 to 1 for the wrists, 3 to 1 for the shoulders, 3 to 0 for the neck, and 3 to 0 for the back (Fig. 9). Therefore, no

Figure 8. Commercially available tilting devices.

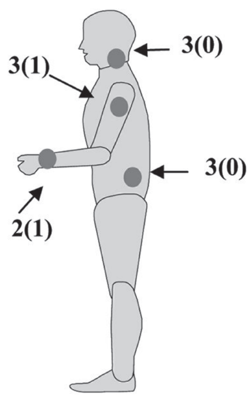


Figure 9. BRIEF analysis before (after) suggested changes.

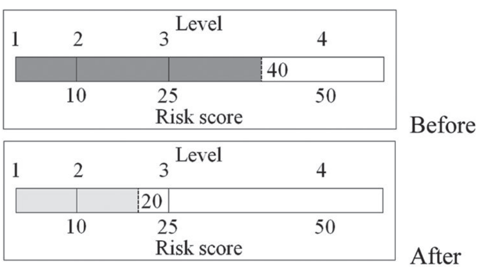


Figure 10. KIM analysis before and after suggested changes.

awkward operator postures are expected to develop after the suggested adjustments are implemented.

For the filling task, KIM analysis shows that the rating points for posture can be reduced from 2 to 1, and the rating points for work environment can be lowered to 0 by means of the suggested workspace rearrangement. In addition, KIM assessment for the filling task after amendment falls within risk level 2, indicating physical overload is unlikely for healthy adults (Fig. 10).

4 CONCLUSIONS

This study conducted an ergonomic evaluation to reduce the identified risk factors associated with the developing of MSDs in a printing ink manufacturing plant. Based on the results of our ergonomic assessment, the worktable was redesigned to improve work postures for the filling, weighing, and packing operations. The new design can also reduce the repetition of can lifting for the filling operator. In addition, the design of the two fillisters will allow the operator to check the weight of ink in a neutral posture and to push the cans easily to the capping machine without extra effort. The use of tilting devices can further prevent operators from bending their trunks to pick up empty cans from floor level. In addition, the muscle strain in their waists and upper extremities can be reduced. Implementing these ergonomic interventions can not only improve work posture and increase worker comfort, but also reduce the risk of MSDs and potentially increase the productivity of this ink manufacturing factory.

ACKNOWLEDGEMENTS

The authors would like to thank the Institute of Occupational Safety, Taiwan for financially supporting this research under Contract No. 973029 and 983067.

REFERENCES

Humantech Inc. 1993. *Applied Ergonomic Training Manual*. South Padre Island, Texas.

Putz-Anderson, V. 1988. *Cumulative trauma disorders: A manual for musculoskeletal diseases of the upper limbs*. USA: Taylor & Francis.

SLIC European inspection and communication campaign 2010. Lighten the load. *http://www.handlingloads.eu/en/site/18/19*

Taiwan IOSH 2005, Study of upper-extremity mechanical exposures in manufacturing industry. *Research report IOSH93-H103. Institute of Occupational Safety & Health, Taiwan (IOSH).*

Taiwan IOSH 2010, A study of effectiveness in prevention of ergonomic hazards in workplaces. *Research report IOSH98-H315. Institute of Occupational Safety & Health, Taiwan (IOSH).*

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A field study on ergonomic interventions of WMSDs prevention in a chemical company

Hsin-Hung Tu
Department of Industrial Engineering and Management, Hsiuping Institute of Technology, Taiwan

Chi-Yuang Yu
Department of Industrial Engineering and Engineering Management, National Tsing-Hua University, Taiwan

Yi-Tsong Pan
Institute of Occupational Safety and Health, Council of Labor Affairs, Taiwan

ABSTRACT: In Taiwan, Work-related MusculoSkeletal Disorders (WMSDs) in strenuous works are becoming a serious occupational safety and health problem. The strenuous tasks are defined as the work which is over-exertion or intense work with risk of WMSDs incidence. The primary reason to WMSDs is primarily due to inappropriate workplace design, which enforces the worker to adopt awkward working posture, and this in turn causes fatigue and results in musculoskeletal disorders. The purpose of this study is to apply ergonomic principle on workplaces improvement for workers to prevent WMSDs. An atlas was used to help improvement, and Key Indicators Method (KIM) was used to evaluate the goodness of the improvement. In this article, a case of chemical company in field study was studied, and three stations in it were improved to show results of the intervention. The results showed that the ratings by KIM of three stations were reduced.

Keywords: work-related musculoskeletal disorders (WMSDs), ergonomic principle, KIM

1 INTRODUCTION

Work-related MusculoSkeletal Disorders (WMSDs) in strenuous works are becoming a serious occupational safety and health problem in many countries (Liang 2003). The strenuous tasks are defined as the work which is over-exertion or intense work with risk of WMSDs incidence. In epidemiologic survey of occupational injuries of many countries, there are high ratios of the WMSDs in occupational injuries. For example, in U.S., the labor loss by WMSDs is about 33.3%, and it costs about 13–20 millions US$ every year (committee on Human Factors 1998). In fact, that is almost about 1–2% GDP of U.S. (Ahrens & Pigeot 2004). In Japan, the ratio of WMSDs to occupational injuries is about 41.2% (Chang et al. 2007), and 38% in EU (ERUOGIP 2007). In Taiwan, the ratio of WMSDs to occupational injuries is increasing as well as the payment for WMSDs in the labor insurance, which is about 80% (BLI 2010). Either in Taiwan or in other countries, WMSDs have become the most important problems in the labors.

WMSDs increase the fatigue in work and the harm in health, and decrease the effectiveness and the quality in production, that would be against both the labor and capital. For the benefit of both the labor and capital, WMSDs must be decreased or eliminated. To decrease or eliminate WMSDs, the factors inducing WMSDs have to be recognized. Five most-common-agreed major factors recognized include exertion, repetitiveness, vibration, temperature, and awkward posture (Lee 1999, Sillanpakak et al. 1999, Ulin et al. 2004).

To decrease or eliminate these five factors, many institutes or researchers are using ergonomic techniques to intervene the workplaces. Easy Ergonomics (Easy Ergonomics 1999), Marras and Karwowski (Marras & Karwowski 2006), Safety GRANT$ Project (Ohio Bureau of Workers' Compensation-Safety 2008) are for instances. European Agency for Safety and Health at Work are campaigning for the occupational safety and health, and awarded the good practice every year (European Agency for Safety and Health at Work 2007).

Although the researches or projects tried to improve five factors in workplace, most of them worked on awkward posture problem. It is due to the essential of the task and the cost of improvement. It is usually hard to change the repetitive wrist twisting in bearing work station, or the standard 20 kg weight of customer package in airline.

The purpose of this study is to improve and prevent WMSDs with ergonomic intervention in workplace. The case study in this paper is a chemical company, and three work stations in it were improved and evaluated by Key Indicators Method (KIM) (The KIM Tool 2010).

2 METHOD

2.1 *Altas "Collection of Ergonomical Working Postures"*

An atlas developed by Institute of Occupational Safety and Health (IOSH), entitled "Collection of Ergonomical Working Postures", was used to help the improvement of the workplace in this study (Chang et al. 2007). It was developed as a quick reference for the safety engineer or work station designer in workplace since 2005. Measurements and layout of 100 different workplaces were collected as well as a standard template three dimensional human model with proper working posture in it. As in Figure 1, four difference views (front, lateral, superior, and perspective) of the same workplace design were shown in one page and illustrated for the engineer or designer.

2.2 *Key Indicators Method (KIM)*

Key Indicators Method (KIM) has been developed by the Federal Institute for Occupational Safety and Health (BAuA) and Committee of the Laender for Occupational Safety and Health (LASI). Its blueprint was published in 1996, and final versions were published in 2001 and 2002 (The KIM Tool 2010). For KIM manual handling tasks, three steps are necessary in the assessment: the first step is determination of time rating point, that is to determine the rating point by frequency (time) and three possible forms of loading handling. The second step is determination of the rating points of load, posture and working conditions. The third step is evaluation, that is to calculate the activity-related risk score by addition of the rating points of the key indicators and multiplication with the time rating point. Base on the risk score, some design needs and approaches are suggested. This method was used in this study to evaluate the before-improvement and after-improvement working conditions. These assessment steps are designed in a worksheet, which could be easily accessed from the internet (The KIM Tool 2010) and not be shown in this paper.

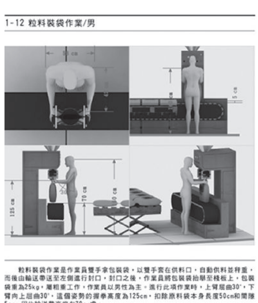


Figure 1. A sample page of the "Collection of Ergonomical Working Postures" atlas: The descriptions in it are in Chinese.

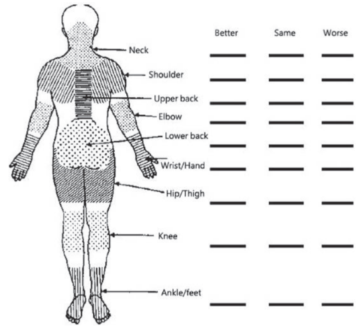


Figure 2. An illustration of a revision NMQ used in this study, which quantifies musculoskeletal pain and activity prevention in 9 body regions.

2.3 *Nordic Musculoskeletal Questionnaire (NMQ)*

The Nordic Musculoskeletal Questionnaire is a screening instrument to quantify musculoskeletal pain and related activity prevention (Kaewboonchoo 1998). It quantifies musculoskeletal pain and activity prevention in 9 body regions, illustrated as Figure 2 above. A revision of NMQ was used in this study to evaluate the subjective rating of the workplace before and after improvement.

3 RESULTS

Three work stations were improved and evaluated by KIM, the first one is a trimming task work station, the second one a plastic cleaning task work station, and the third one a razor-assemble task work station.

3.1 *Trimming task*

In the trimming task, as shown in Figure 3(a), an operator trimmed the ragged edge of the product item on a desk aside a conveyor. The operator bent forward her upper body to near the product item on the desk. She processed each item in 4 minutes, that is, 120 items one working day. The product item weight less than 1 kilogram.

After improvement, as shown in Figure 3(b)(c), the operator bent less her upper body, and could decrease the risk of WMSDs. The risk score of this task changed from 20 to 12, and the risk range stays in level 2, as shown in Table 1.

The subjective rating of the operator shows that the neck, the shoulder, and the lower back are better after the workplace improvement.

3.2 *Plastic cleaning task*

In the plastic cleaning task, as shown in Figure 4(a), an operator cleaned the residuals of thermosetting plastic on a steel cylinder with a hammer weighting about 5 kilograms. The operator bent his upper body slightly, and waved the hammer about 40~50 times in each cleaning task, and 4 cylinders have to be cleaned one day. That is, less than 200 times of waving one working day.

After improvement, as shown in Figure 4(b)(c), the upper body of the operator could be upright, and he could decrease the risk of WMSDs. The risk score of this task changed from 12 to 8, and

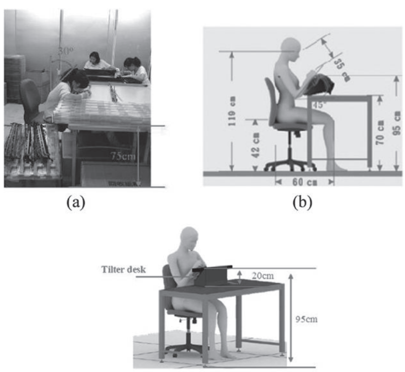


Figure 3. Trimming task (a) the workplace before improvement (b) a improving reference in "Collection of Ergonomical Working Postures" atlas (c) the workplace after improvement.

Table 1. The KIM evaluation of trimming task.

Status	Time rating point	Load rating point	Posture rating point	Working conditions	Risk score	Risk range
Before	4	1	2	0	12	2
After	4	1	1	0	8	1

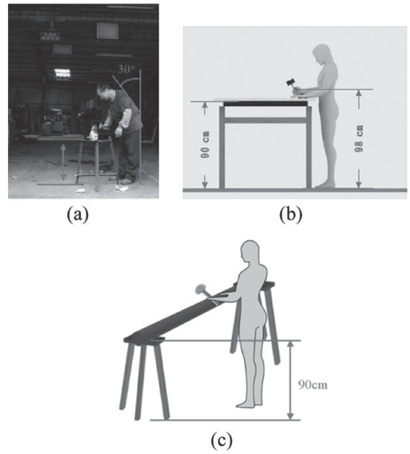


Figure 4. Plastic cleaning task (a) the workplace before improvement (b) a improving reference in "Collection of Ergonomical Working Postures" atlas (c) the workplace after improvement.

Table 2. The KIM evaluation of the plastic cleaning task.

Status	Time rating point	Load rating point	Posture rating point	Working conditions	Risk score	Risk range
Before	4	1	4	0	20	2
After	4	1	2	0	12	2

the risk range decreased from level 2 to level 1, as shown in Table 2.

The subjective rating of the operator shows that the neck, the shoulder, and the lower back are better after the workplace improvement.

3.3 *Razor-assemble task*

In the razor-assemble task, as shown in Figure 5(a), an operator had to assemble 22 razor blades into a cylinder with a squat posture. Each razor blade weights less than 1 kilogram, and the operator had to assemble and dissemble the blades 7 times in a working day, that is, 308 times one working day.

After improvement, as shown in Figure 5(b)(c), the operator could use a normal standing posture to assemble and dissemble the razor blades, and could decrease the risk of WMSDs. The risk score of this task changed from 30 to12, and the risk range decreased from level 3 to level 2, as shown in Table 3.

The subjective rating of the operator shows that the neck, the shoulder, the lower back, the thigh, and the knee are better after the workplace improvement.

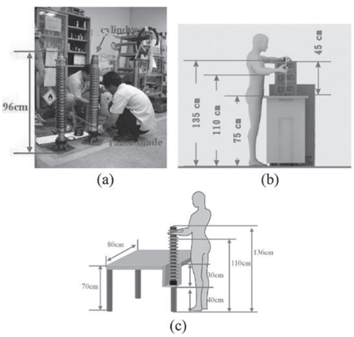


Figure 5. Razor-assemble task (a) the workplace before improvement (b) a improving reference in "Collection of Ergonomical Working Postures" atlas (c) the workplace after improvement.

Table 3. The KIM evaluation of the razor-assemble task.

Status	Time rating point	Load rating point	Posture rating point	Working conditions	Risk score	Risk range
Before	6	1	4	0	30	3
After	6	1	1	0	12	2

4 CONCLUSION AND DISCUSSION

The purpose of this study is to apply ergonomic principle on workplaces improvement for workers to prevent WMSDs. An atlas, "Collection of Ergonomical Working Postures", was used to help improvement, and Key Indicators Method (KIM) was used to evaluate the goodness of the improvement. Three work stations in a chemical company were improved to show results of the intervention. The results showed that the KIM risk scores of all three stations were reduced.

The rating scores corresponded to the subjective rating. Although the rating scores of all three work stations are reduced after improvement, only two of their risk ranges are leveled down. However, the subjective rating of the operators in all three work stations showed that they felt better in many important body regions. This shows the correspondence between rating scores of KIM and subjective rating.

A sitting posture was evaluated with definition not illustration reference of KIM worksheet in this study. In trimming task (see also Figure 3), the KIM method seems not to have proper illustrative postures to be referred for the sitting posture. Instead of the direct reference of illustrative postures in KIM worksheet, the definition of posture in KIM worksheet—"Low bending or far bending forward"—was used to determine the rating of sitting posture before improvement. And for the posture after improved, the definition of posture in KIM worksheet—"Slightly bending forward or twisting the trunk"—was used. However, it is only a trial in this study, the correctness of using KIM to evaluation the sitting posture with prolonged workload should be further studied.

REFERENCES

Ahrens, W. & Pigeot, I. 2004. *Handbook of epidemiology*: 1444. Springer.

BLI 2010. 2005–2009. *Annual reports*. Bureau of Labor Insurance, Council of Labor Affairs, ROC.(Taiwan).

BLS 2007. Nonfatal occupational injuries and illnesses requiring days away from work (2002~2006). Bureau of Labor Statistics, USA.

Chang, C.P., Chen, C.Y., Yu, C.Y., Tu, H.H. & Chuang, H.F. 2007. *Collection of ergonomical working postures*, Institute of Occupational Safety and Health, ROC. (Taiwan).

Committee on Human Factors. 1998. Work-Related Musculoskeletal Disorders: A Review of the Evidence, National Research Council, National Academy Press, Washington, DC.

Easy Ergonomics 1999. A Practical Approach for Improving the Workplace, Cal/OSHA Consultation Service, California Department of Industrial Relations.

EUROGIP 2007. Musculoskeletal disorders in Europe-Definitions and statistics, EUROGIP.

European Agency for Safety and Health at Work. 2007. Lighten the Loard. *Safety and Health at Work European Good Practice Awards,* European Agency for Safety and Health at Work.

Kaewboonchoo, O., Yamamoto, H., Miyai, N., Mirbod, S.M., Morioka, I. & Miyashita, K. 1998. The Standardized Nordic Questionnaire Applied to Workers Exposed to Hand-Arm Vibration. *J Occup Health* 40: 218–222.

Lee, C.L. 1999. MSDs of construction labourer, *Labor Safe and Hygiene News in Brief* 39: 4.

Liang, H.W. 2003. Labor investigate yearly report, A brief of WMSDs. *Industrial safety scientific and technological quarterly* 50: 2–6, Council of Labor Affairs Executive Yuan, ROC (Taiwan).

Marras, W.S. & Karwowski, W. 2006. Interventions, Controls, and Applications in Occupational Ergonomics, CRC Press.

Ohio Bureau of Workers' Compensation-Safety GRANT$, http://www.ohiobwc.com/employer/programs/safety/EmpGrants.asp, visited on 2008.10

Sillanpakak, J., Lappalainen, J., Kaukiainen, A., Viljanen, M. & Laippala, P. 1999. Decreasing the physical workload of construction work with the use of four auxiliary handling devices. *International Journal of Industrial Ergonomics* 24: 211–222.

Suka, M. & Yoshida, K. 2005. Musculoskeletal pain in Japan : prevalence and interference with daily activities. *Modern Rheumatology* 15(1): 41–47.

The KIM Tool, http://www.handlingloads.eu/en/site/18/19, visited on 2010.05

Ulin, S.S. & Keyserling, W.M. 2004. Case Studies of Ergonomic Interventions in Automotive Parts Distribution Operations. *Journal of Occupational Rehabilitation* 14(4): 307–326.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The promotion of WMSD prevention in Taiwan

ChihYong Chen, ChengPing Chang, WenYi Yeh & Yi-Tsong Pan
Institute of Occupational Safety and Health, Council of Labor Affairs, Taiwan, ROC

Chi-Yuang Yu
Department of Industrial Engineering and System Management, Chung Hwa University, Taiwan, ROC

ABSTRACT: In Taiwan, the statistics of labor insurance showed that the compensated WMSD cases accounted for 68% of all occupational illness in 2008. To address musculoskeletal disorders in the workplace, we provide assistance to the small and medium business, including the free on-site ergonomic consultation and intervention, easy-to-follow intervention guides. To date, the project has counseled 200 factories, more than 800 workplace cases. Results of on-site workplace improvement is highly recognized. Complains of pain and discomforts of workers drops and work performance increases significantly. After intervention, the lumbar compression force reduces about 20%–73%. The cost of each intervention for most of the cases were less than 50 thousands NT dollars.

Keywords: musculoskeletal disorder, ergonomic intervention, anthropometry

1 INTRODUCTION

Work-related MusculoSkeletal Disorders (WMSD), often referred to as ergonomic injuries, are injuries or illnesses affecting the connective tissues of the body such as muscles, nerves, tendons, joints, cartilage, or spinal discs. There is debate concerning sources of risk, and mechanisms of injury. The complexity of the problem is further increased because the related risk factors interact and vary over time. Research is needed to clarify these risk factors, but research is complicated by the fact that estimates of incidence in the general population, as contrasted with the working population, are unreliable because the two overlap (NRC, 2001).

Though, the mechanism of injury is elusive and further research is needed, WMSD has been recognized as one of the most prevalent occupational ailment affecting health care expenditure and work force resources. In USA, WMSD accounted for 29 percent of all workplace injuries requiring time away from work in 2007, compared to 30 percent of total days-away-from-work cases in 2006. (DOL, 2008). In Taiwan, the statistics of labor insurance showed that the compensated WMSD cases accounted for 68% of all occupational illness in 2008 (CLA, 2008). Some of WMSD cases do return to work after medical treatment, but many will face recurrence and worsen health conditions.

To address musculoskeletal disorders in the workplace, OSHA developed a four-pronged ergonomics strategy through a combination of industrial-specific and task-specific guidelines, outreach, enforcement, and research (Figure 1). In Taiwan, 97% of the enterprises are small and medium business, in which about 76.58% of all labors are hired. Employees are less than 200 labors in every individual business.(MEA, 2009). On-site safety & health stuff do not have the abilities to evaluate the ergonomic risk and to implement ergonomic intervention. Furthermore, the exposure-disorder relation of MSD is not very clear. Rashly setting an ergonomic duty clause would face tough challenge. Though, none of the common musculoskeletal disorder is uniquely caused by work exposures, MSD is frequently associated with work environment. The intervention strategies to reduce these risks should be encouraged and extended. The proactive strategy has the potential of benefiting not just to improve quality of workers' life but also the productivity of the industry. For that majority of smaller employers facing the challenging of economic recession, government should provide more assistance to them, including the free on-site ergonomic consultation and intervention, easy-to-follow intervention guides.

2 METHOD AND MATERIAL

According to a nation-wide survey, 51.7% of labors considered that their work involved with

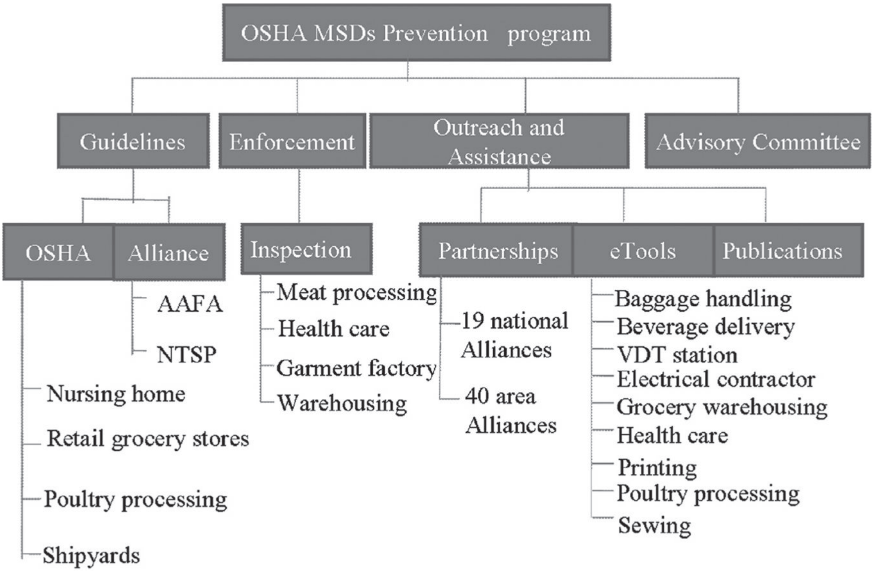


Figure 1. OSHA MSDs prevention program.

awkward postures, and 60% with repetitive motions (IOSH, 2004). Among thee MSD risk factors, such as awkward posture, forceful exertion, repetitive motion, vibration, and extreme temperature, awkward posture is much easier to recognized and evaluated by workers, e.g. elevation of shoulder, bend back. Prolonged work in poor posture is one of the leading causes of WMSD, which results in fatigue, pain, or disease in the neck, shoulder, wrist, back, or other parts of body. To some extent, poor working postures are related with unsuitable design of workstations, which do not fit the labors' anthropometry.

To help labors further understand the relation between poor postures and anthropometry, first IOSH (Institute of Occupational Safety and Health) and National Tsing Hua University had had undergone a series of 1D and 3D anthropometrical database survey, which completed in 2004. The evolution of MSD prevention in Taiwan is depicted as following (Figure 2).

The above database was used to compile a booklet "Collection of Ergonomic Working Postures" which contains 100 typical workplace designs based on the concept of "functional working posture" in 2005. A functional working posture is one that is natural and energy saving for task performance. The basic idea is to keep the trunk and neck upright as much as possible, because bending of these two segments will result in excessive stress on the lumbar and neck. In addition, keep the hands as close to the body as possible to reduce stress.

In 2007, a pilot study on intervention was conducted and a second booklet "Method of Ergonomic Workplace Improvement" was compiled to instruct on-site S & H stuffs how to utilize the booklet "Collection of Ergonomic Working Postures" and understanding the concept of "functional working posture". First, through the posture comparison, such as standing or sitting, and heavy or light work, they can choose an appropriate typical workplace design from the booklet. Then adjust the size in the typical workplace design to fit the on-site situation and workers' anthropometry.

Based on intervention cases of pilot study, we put together these preliminary techniques of ergonomic intervention for experts. This preliminary techniques consist of a standardized chart of intervention process and a set of work sheets of standard procedure. To widely diffuse the ergonomic intervention, an easy-to-follow intervention guides was made. This rule-based technique, checking-typed process chart and diagrammatic Standard Operation Procedure (SOP), in reducing the dependency on ergonomics and engineering knowledge, so any personnel with basic training in ergonomics is able to improve workplace by him/herself based on this SOP.

With all these booklets served as guiding reference, we organized a consulting team to promote the MSD prevention in 2009. To achieve this objective, the following works were devised: (1) world-wide literature review on promotion of ergonomic hazard prevention, (2) design and print colored promotion leaflets on promotion of ergonomic hazard prevention, (3) held workshops on promotion of ergonomic hazard prevention, (4) held

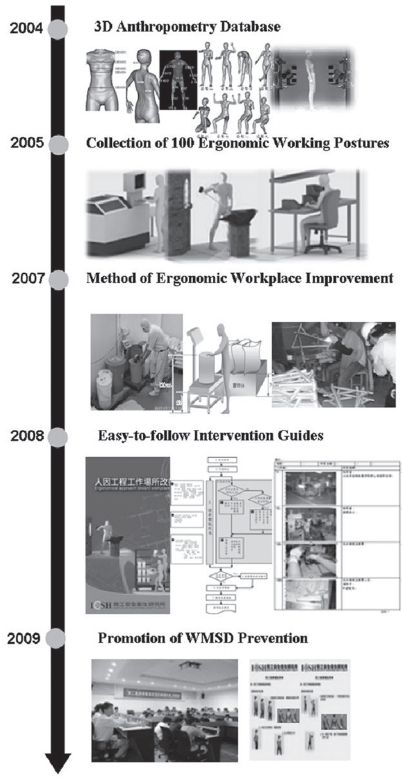


Figure 2. The evolution of WMSD prevention program.

colloquiums on promotion of ergonomic hazard prevention (for recruiting participating factories), (5) conduct the on-site ergonomic intervention.

3 RESULTS AND DISCUSSION

To date, the project has counseled 200 factories, more than 800 workplace cases. Results of on-site workplace improvement is highly recognized. Complains of pain and discomforts of workers drops and work performance increases significantly (Figure 3). The cost of each intervention for most of the cases were less than 50 thousands NT dollars (Figure 4).

This prevention project primarily focused on working posture improvement, through workstation redesign and use of easy supporting equipment. To reduce repetitive hand/wrist motion, it is sometimes involved with the use of machine or automation. The cost would be higher than that of the posture improvement and this higher cost investment may impede employer's attempt. We hope to acquire some more experience on improvement of repetitive hand/wrist motion and conquer this problem in the future.

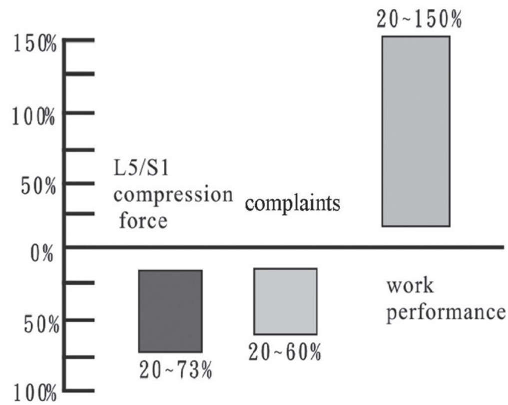


Figure 3. The percentage of reduction of L5/S1 compression force, complaints, and increase of performance after intervention.

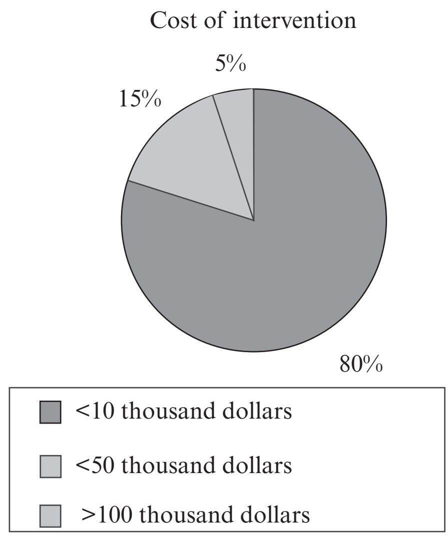


Figure 4. The cost distribution of intervention.

REFERENCES

Council of Labor Affairs 2008. *Yearbook of labor insurance statistics*. Taiwan, ROC.

IOSH 2005. *Survey of perceptions of safety and health in the work environment in 2004*. Institute of Occupational Safety and Health, Taiwan, ROC.

MEA 2009. *White paper on small and medium enterprises*. Ministry of Economic Affairs, Taiwan, ROC.

National Research Council & Institute of Medicine 2001. *Musculoskeletal disorders and the workplace*. Washington DC, USA: National Academy Press.

United States Department of labor 2008. *Nonfatal occupational injuries and illnesses requiring days away from work 2007*, USA.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Field studies on work-related musculoskeletal disorders

Shu-Zon Lou
School of Occupational Therapy, Chung Shan Medical University, Taichung, Taiwan, ROC

Chun-Chieh Chen
Faculty of Medicine, School of Medicine, Chung Shan Medical University, Taichung, Taiwan, ROC

Wen-Yu Ye
Labor Safety and Health Management Society, ROC

Chi-Yuang Yu
Department of Industrial Engineering, National Tsing Hua University, Taiwan, ROC

Cheng-Lung Lee
Industrial Engineering and Management, Chaoyang University of Technology, Taiwan, ROC

ABSTRACT: Taiwan was once a manufacturing country. Even now there are still many traditional factories. Occupational diseases from Work-related MusculoSkeletal Disorders (WMSD) such as low back pain, carpal tunnel syndrome, are commonly seen in Taiwan because of the lack of concept of ergonomics to prevent. The purpose of Field studies on Interventions of WMSD Prevention for the factories of Taiwan is to explore the WMSD on workplace and to improve by redesigning the workplace. The joint loads on L4/L5 were used for analysis and assessment with JACK software.

Keywords: WMSD, occupational disease, JACK software

1 INTRODUCTION

Musculoskeletal disorders are reported to occur in certain industries and occupations with rates up to three or four times higher than the average rate cross all industries (Punnett et al. 2004, Spallek et al. 2010). Taiwan was once a manufacturing country. Even now there are still many traditional factories. Occupational diseases from Work-related MusculoSkeletal Disorders (WMSD) such as low back pain, carpal tunnel syndrome, are commonly seen in Taiwan because of the lack of concept of ergonomics to prevent.

A number of intrinsic and extrinsic factors have been implicated in the aetiology of WMSDs (Punnett et al. 2004, Tinubu et al. 2010). Silverstein et al. (Silverstein et al. 2007) reported repetitious movement, awkward postures, and high force levels as the three primary risk factors that have been associated with WMSDs.

Accurate reporting of work-related conditions is necessary to monitor workplace health and safety and to identify the interventions that are most needed (Hoonakker et al. 2010).

The aim of this field study was to explore the effects of the loads, posture and working type on WMSD especially for lumbar spinal disorders on Taiwan.

2 METHOD

2.1 *Subjects*

This study was performed in the Chung Shan Medical Hospital, Chung Shan Medical University. Ten patients seeking medical care for pain associated with clinically and radiologically verified Herniated Intervertebral Disc (HIVD) or symptomatic lumbar disc spondylolisthesis and applying for worker's compensation from 2008 to 2010 were surveyed and analyzed. Field study was then done in the worksite. They are from meat processing worker (Meat), worker in injection molding equipment (Injection), worker in mold factory (Mold), four construction workers (Construction), two package delivery workers (delivery), and one power pole worker (Power). Basic demographic data was

Table 1. Basic demographic data of patients.

	Age (yrs)	Gender	Years of experience	Loading hours per day	Injury
Meat	60	M	39	1	SP
Injection	51	F	19	3	HIVD
Mold	41	F	18	5	SP
Construction1	55	F	20	2	HIVD
Construction2	48	F	32	2	HIVD
Construction3	56	F	25	2	HIVD
Construction4	52	F	20	2	HIVD
Delivery1	28	F	5	3	HIVD
Delivery2	51	F	26	3	HIVD
Power	34	F	11	2	HIVD

SP: Spondylolisthesis, HIVD: Herniated Intervertebral Disc.

listed on Table 1. The meat processing worker dealt with pig internal organ. The job included the packing the internal organ and put them to a barrel. The job in injection molding machine included lifting the raw materials of the ground to dry machine and pour the dried materials to tank of injection molding machine. The job in the construction included the lifting the steel and wood plate for building. The package delivery worker processed the packages passing to customs. Their job included the lifting and carrying the packages. Power pole worker processed the power wire connection. Their job included the lifting materials to electrical pole and processed the connection of wire.

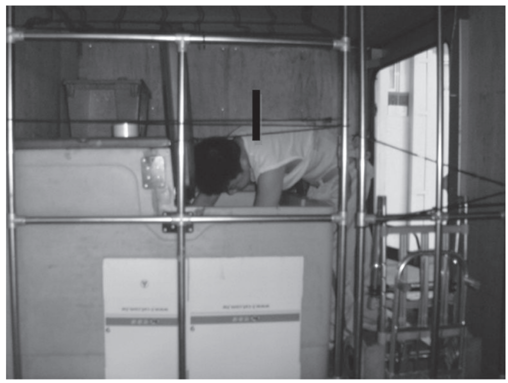


Figure 1. A typical picture from workplace.

2.2 *Data process*

The main lifting tasks were recorded by digit camera (fig. 1). There are some parameters calculated. Peak load is the maximum weight of lifting or carrying. Total loads is the total amount of load being lifted. Number of bending is the total number of bending body forward while lifting the weight of objects over 5 Kg. Peak loads in lumber is the vertical force exerted on spine, calculated from the Jack software (fig. 2). Average loading is the mean of the loads on lumbar spine in the period of lifting. Cumulative exposure was sum of loading exerted in lumbar in a year of experience as follows (Seidler et al. 2010).

$$\textbf{Sum dose of year} = \textbf{Days x} \left(\textbf{8h x} \left[\sum \textbf{F}_i^2 \textbf{ x t}_i\right]\right)^{1/2}$$

$$\textbf{Cumulative exposure} = \textbf{work year} \times \textbf{sum dose per year}$$

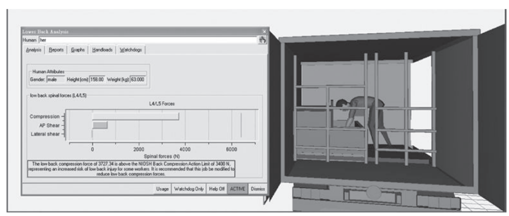


Figure 2. Loading analysis on L4/L5 with JACK software.

3 RESULTS

Peak load total loads and number of bending of Table 2 is information of the interview with the patients on his or her worksite. Peak load in lumber and average load is calculated from Jack software by the measurement of workplace environment and the height of subjects.

3.1 *Age*

The youngest patient was 28 from package delivery worker, and the oldest was 60 from meat processing worker.

Table 2. Loading data on patients.

	Peak loads	Total loads	Number of bending	Peak loads in lumber	Average loading	Cumulative exposure
Meat	30	1500	100	3400	2200	$55*10^6$
Injection	30	1200	200	4400	2500	$61*10^6$
Mold	40	4134	925	4400	3000	$36*10^6$
Construction1	60	2300	200	4200	2500	$67*10^6$
Construction2	50	7200	510	4408	3120	$111*10^6$
Construction3	50	3200	425	4105	3050	$85*10^6$
Construction4	50	3200	225	4215	2875	$64*10^6$
Delivery1	25	9100	1300	4312	2757	$18*10^6$
Delivery2	35	3000	800	4415	2845	$101*10^6$
Power	30	3120	110	4240	2750	$49*10^6$

3.2 *The number of bending*

The most number of bending occurred on Delivery1 to 1300 times per day. On the other hand, the least number of bending was Meat to 100 times per day.

3.3 *Peak load in lumber*

The least peak load in lumber was 3400 N at Meat. The others in peak load were very close to 4300 N.

4 DISCUSSION

The field studies explored the effects of the working posture and the weight of object on lumbar force and lumbar spinal disorders. Calculation of the total loads, number of bending, peak loads in lumbar and average load in lumbar was based on self reported data on the weight of lifted or carried objects, the duration of lifting, and harmful working posture. With the on-site interview and estimation, the self reported data could be more reasonable.

According worker compensation of occupational disease in lumbar spine on Taiwan, there are three adequate occupational exposures. Firstly, workday is over 220 days per year and the maintaining 8 to 10 years. Secondly, cumulative exposure is greater than $25*10^6$ Nh. Thirdly, the compressive force on lumbar spine is greater than 3400 N. In this study, ten patients reached the criteria as the occupational disease of lumbar spine. The case Delivery1 should pay more attention. The weight of objects is less and the getting the occupational disease is the faster. The is possible from the larger number of bending that leaded the cumulative exposure to the limited.

We notice that the heavier loads in lumbar spine come from lifting objects from the ground especially for the package delivery. It is suggested that the height of the object should move to the height of the work's elbow about 75 cm with a table, which can reduce the bending and the loads on lumbar. This is a easy prevention method on WMSD but the stack space would decrease or limit the walkway.

5 CONCLUSIONS

The larger number of bending will decrease the weight that worker could lift and shorten duration of being occupational disease even while workplaces are tending to change towards decreasing physical work load overtime.

REFERENCES

Hoonakker, P. 2010. Monitoring working conditions and health of older workers in Dutch construction industry. *Am J Ind Med*. 53(6): 641–53.

Punnett, L. 2004. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *J Electromyogr Kinesiol.* 14(1): 13–23.

Seidler, A. 2001. The role of cumulative physical work load in lumbar spine disease: risk factors for lumbar osteochondrosis and spondylosis associated with chronic complaints. *British Medical Journal* 58(11): 735.

Silverstein, B.L. 2007. Occupational factors and carpal tunnel syndrome. *American Journal of Industrial Medicine* 11(3): 343–358.

Spallek, M. 2010. Work-related musculoskeletal disorders in the automotive industry due to repetitive work—implications for rehabilitation. *J Occup Med Toxicol.* 5: 6.

Tinubu, B.M. 2010. Work-related musculoskeletal disorders among nurses in Ibadan, South-west Nigeria: a cross-sectional survey. *BMC Musculoskelet Disord* 11: 12.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Application of the key indicators method in ergonomic interventions

Hsieh-Ching Chen, Feng-Ying Lin & Cheng-Lung Lee
Chaoyang University of Technology, Wufong, Taichung County, Taiwan

Yi-Tsong Pan & Cheng-Ping Chang
Institute of Occupational Safety and Health, Xizhi, Taipei County, Taiwan

ABSTRACT: Ergonomic interventions to reduce risk of Work-related Musculoskeletal Disorders (WMSDs) typically require tools to help investigators identify ergonomic risk factors. This study evaluates the applicability of Key Indicator Method (KIM) tools to Taiwan industries using ergonomics data collected by a project conducted by the Institute of Occupational Safety and Health (IOSH), Taiwan, during 2008–2009. The dataset includes inspection reports and advice for 379 tasks for 102 companies conducted by 10 ergonomists. Risk score and risk level of each task before and after suggested ergonomic changes were evaluated using the KIM. Statistical results show that 77% of all tasks were associated with manual load handling, to which the KIM is applicable. Unfortunately, about two-thirds of these tasks were not evaluated by KIM due to missing rating data for "time" (97%), "load" (36%), "posture" (2%), and "environment" (1%). Only 98 inspection tasks with complete rating data were assessed. Among these tasks, none had an increased risk score after a suggested amendment and 72 cases (74%) had a reduced risk level according to ranges of risk scores by the KIM. The overall analytical result suggests a good agreement between inspector experience and the KIM scoring system. This study demonstrates that the KIM tools are appropriate for Taiwan's manufacturing industry for onsite screening physical workloads.

Keywords: physical workload, musculoskeletal disorder, ergonomic intervention

1 BACKGROUND

MusculoSkeletal Disorders (MSDs) are the most common work-related injury in numerous countries (Punnett & Wegman 2004). Work-related injuries and disability are potentially preventable and identifying suitable interventions that reduce work-related MSDs (WMSDs) is important. A combination of physical, psychological, and psychophysical workplace risk factors have been documented (Punnett & Wegman 2004). Physical risk factors, such as high forces, high repetition, working in awkward postures, long-term static postures, local contact forces, and vibrations are common. However, the complex causes of MSDs pose a significant obstacle to their control. Consequently, ergonomic risk assessment methods, which consider multiple physical exposures in an integrated model of risk prediction, are frequently utilized to direct industrial prevention initiatives (Jones & Kumar 2007).

Chris Hamrick (2006) proposed that the ideal field assessment tool should have the following characteristics:

– Predictability—the tool should provide a valid predictive measure of risk of musculoskeletal injury that would occur for a population performing the assessed task.
– Robustness—can be used in any work situation.
– Inexpensiveness—the tool must be available at minimal monetary cost.
– Non-invasion—the assessment does not alter the way a worker performs a job, nor alter the process workflow or work quality.
– Quickness—assessment and analysis can be performed quickly.
– Ease of use—the tool can be used with minimal training.

Various assessment tools have been developed with these characteristics. For instance, the revised National Institute for Occupational Safety and Health (NIOSH) lifting equation (Waters et al. 1993, NIOSH 1994) and the lifting Threshold Limit Value (TLV) of American Conference of Industrial Hygienists (ACGIH 2003) were developed for material handling assessments. The Strain Index (Moore & Garg 1995), the ACGIH Hand

Activity Level (HAL) TLV (ACGIH 2003), and the concise exposure index (OCRA) (Occhipini 1998) were developed for upper extremity assessments. The Rapid Upper Limb Assessment (RULA) (McAtamney & Corlett 1993) and Rapid Entire Body Assessment (REBA) (Hignett & McAtamney 2000) were developed for entire body assessments. Although these assessment tools were developed and validated for identifying potential MSD risk factors in particular instances, an assessment tool with few limitations in its application to general on-site tasks is preferable.

The Key Indicator Method (KIM), developed by the Federal Institute for Occupational Safety and Health, German (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, BAuA) and Committee of the Laender for Occupational Safety and Health, German (Länderausschuss für Arbeitsschutz und Sicherheitstechnik, LASI), was published in 2001–2002 and has been adopted by several EU countries for risk assessment on the screening level for manual load handling (SLIC 2010, Carlsson 2009). The KIM is intended only for identification of bottlenecks and need for action. The KIM contains descriptions of relevant activity indicators, the so-called key indicators. Indicator selection is primary related to their influence on cause-and-effect relationships. The indicators are duration, frequency, and distance with a multiplier for weight, posture, and working conditions. The KIM is a simple diagnostic and/or assessment tool designed for employers, inspectors, technologists, ergonomists, safety officers, company physicians, and even employees.

Although the KIM is easily employed (Carlsson 2009), no Taiwanese study has used the KIM as an assessment tool. The purpose of this preliminary study is to survey the coverage of the KIM to Taiwan industries by reviewing data from previous ergonomic inspections.

2 METHOD

2.1 *Data resource*

The data analyzed were obtained for 379 tasks at 102 industrial companies by 10 inspectors participating in a project by the Institute of Occupational Safety and Health (IOSH), Taiwan, during 2008–2009. The purpose of this IOSH project was to apply ergonomic interventions in workplaces to eliminate or reduce the incidence of WMDs. Ninety percent of inspected companies were in the manufacturing industry and were located in Taiwan's industrial or science parks. The other companies inspected were in the healthcare or social-welfare services industries (4%), transportation and storage (2%), mining and quarrying (1%), animal husbandry (1%) and other services (2%). All inspected companies voluntarily participated in this study after attending symposiums held by IOSH. All inspectors in this project were university faculty specialized in ergonomics. Their average length of experience in workplace inspections was 2.5 years (range, 1–10 years).

For each inspected task, the factors that were assessed, and information gathered by all inspectors comprised an overview of a task, key problems raised by a company, a workplace tour, analyses of specific tasks using such tools as the NIOSH lifting equation, photography and videos of operators performing tasks, and interviews with individual operators. Photographs, physical dimensions of a workplace and assessment tools that enabled quantification of information (e.g., RULA, Ovako Working Posture Analysis System (OWAS), HAL TLV, biomechanics estimation, or the NIOSH lifting equation) were used to explain problems to management. When recommending task changes, inspectors usually indicated that changes were required, such as modifications to equipment, workspace layout and/or working practices. Diagrams or sketches for tool and workspace changes were typically provided in the inspection report.

2.2 *Analyses procedure*

Two investigators in this study reviewed each inspection report and recommended task changes. All task changes were divided into groups of manually handling loads and others. For each task involving manual load handling, investigators read the inspection report to identify the indicator information of handling duration, frequency or distance, and handling weight (load), posture, and working conditions. Inspected tasks with incomplete information were identified and the types of missing data were registered. For tasks with complete information, risk scores and risk levels before and after suggested ergonomic changes were evaluated using the KIM. Risk scores and risk levels, before and after changes, of each evaluated task were then compared to assess whether adequate improvement can be achieved by the recommended change.

A reduction from a risk score of ≥25 to <25 or from a risk level of ≥3 to ≤2 is considered a beneficial change. A change is considered ineffective if it has a risk score of ≥25 or a risk level of ≥3 after a change. Relatively, recommended changes to a task with an original risk score of <25 or with a risk level of ≤2 is considered inadvisable. Descriptive statistics were used when analyzing quantitative data.

3 RESULTS

Evaluation results indicate that 292 (77%) of all inspected tasks involved manual load handling. Eighty-seven (23%) recommendations not associated with manual load handling included concerns and suggestions for vision (13.8%), the work environment (16.1%), prolonged static posture (59.8%), and safety (10.3%). Unfortunately, 194 tasks involving manual load handling were not evaluated by the KIM due to missing key indicator data. Among these tasks, missing task information for "duration/frequency/distance," "load," "posture," and "working condition" accounted for 97%, 36%, 2% and 1% of causes, respectively.

Only 98 tasks with complete information were evaluated using the KIM. The percentages of these tasks in manufacturing, health-care or social-welfare services industries, transportation, and storage and other services were 92%, 4%, 1%, 3%, respectively. Among these tasks, none had an increased risk score after implementing recommended changes. Figure 1 and Table 1 show a scatter diagram of risk scores and the count of tasks in each risk level, respectively. Eighty-seven (88.8%) tasks had a lower risk score after the suggested change than before the change. According to risk ranges defined by the KIM, 73 tasks (74.5%) had a reduced risk level after suggested changes. Forty-six tasks (46.9%) originally had a risk level of ≥3 and reached a risk level of ≤2 after suggested changes. Nevertheless, 24 (24.5%) of 98 evaluated tasks still had a risk level ≥3, even though 17 tasks had a risk level reduced from 4 to 3, indicating physical overload is possible or likely for a normal worker. Another 28 (28.6%) evaluated tasks had risk levels ≤2 (risk scores <25) before implementing recommended changes, indicating that physical overload is unlikely for a normal worker.

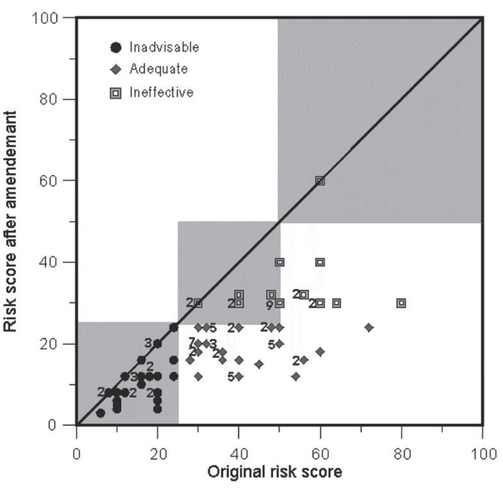


Figure 1. Risk scores of tasks evaluated using the KIM before and after ergonomic changes. Numbers in the plot indicate multiple data. The three grey zones at lower-left, center and upper-right, indicate risk levels of ≤2, 3, and 4, respectively.

Table 1. Number of tasks in each KIM risk level before and after suggested changes (N = 98).

		Before suggested changes			
Risk level		1	2	3	4
After suggested changes	4	0	0	0	1
	3	0	0	6	17
	2	0	15	35	11
	1	3	10	0	0

4 DISCUSSION

According to inspection reports from the IOSH project, significantly more companies in the manufacturing industry than in other industries have a pressing need to identify key factors and reduce risk of MSDs. Those inspectors focused primarily on the physical aspects of tasks. Of all inspected tasks, 77% were associated with manual load handling tasks and most of these tasks required reductions to physical workloads. This demonstrates the potential use and coverage of the KIM tool for Taiwan's manufacturing industry. Nevertheless, nearly 66% of these inspection reports did not have "time-rating" data, the most important determinant for assessing physical workload in the KIM. Investigators inferred the likely causes when time-rating data were missing due to inconsistent task durations or large predicted spinal compression forces exceeding the NIOSH action limit (3400 N). However, missing time information also suggests that operational parameters, such as frequency, duration, overall loading time, or carrying distance, were likely overlooked during ergonomic inspections. The use of the KIM can therefore prompt inspectors to collect decisive parameters.

According to KIM evaluations, all inspector recommendations, except for 11, have a positive effect on reducing manual handling loads. The overall analytical result indicates good agreement between inspector experience and the scoring system in the KIM. Those 11 tasks without reduced risk scores after suggested changes were implemented made use of a floor mat (2), protection or anti-vibration gloves (5), a high chair (1), and an altered equipment arrangement (1) or increased pause duration (2) instead of "key indicators." Among these tasks, 8 had low KIM risk levels (≤2), partly explaining why inspectors suggested changes based on risk factors other than the key indicators in the KIM.

About one-fourth of evaluated tasks did not reach a low KIM risk level (≤2), indicating a possibility of physical overload, even after implementing inspector suggestions. This observation suggests that use of the KIM can help inspectors identify bottlenecks, tasks requiring action, and predict the effectiveness of recommendations. With the KIM, if an inspector determines that a planned change cannot reduce risk level of a task to ≤2, additional action or alternative changes should be implemented. Conversely, if a task already has a risk level ≤2, actions to lower the key indicator rating would be inadvisable to an employer.

To date, the KIM consists of two evaluation tools, one for evaluating tasks with lifting, holding, or carrying (KIM-LHC) operations, and the other for tasks with pushing or pulling (KIM-PP) operations. Investigators identified over 98% of tasks using the KIM-LHC tool. Although the KIM-LHC tool is not applicable to all tasks, it has good coverage and is an effective screening tool for further actions. This KIM tool is easily employed and can be helpful in dialogue with employers about assessing the risk for MSDs in workplaces. The KIM tools can be promising for onsite screening of physical workloads if future development further covers tasks of repetitive manual processing, high action forces, constrained postures and body movement without manual load handling.

ACKNOWLEDGEMENT

The authors would like to thank the Institute of Occupational Safety, Taiwan for financially supporting this research under Contract No. 973029 and 983067.

REFERENCES

ACGIH 2003. *Threshold limit values for chemical substances and physical agents & biological exposure indices* (TLVs and BEIs), American Conference of Governmental Industrial Hygienists, Cincinnati, Ohio.

Carlsson, C. 2009. Presentation of the KIM-tool. In *Proc. 17th Triennial Congress of the International Ergonomics Association (IEA); Beijing, 9–14 August 2009.*

Hamrick, C. 2006. Overview of ergonomic assessment. In W.S. Marras & W. Karwowski (eds.), *Fundamentals and assessment tools for occupational ergonomics*: 34.1–11. Taylor & Francis.

Hignett, S. & McAtamney, L. 2000. Rapid Entire Body Assessment (REBA). *Applied Ergonomics* 31: 201–205.

Jones, T. & Kumar, S. 2007. Comparison of ergonomic risk assessments in a repetitive high-risk sawmill occupation: Saw-filer. *International Journal of Industrial Ergonomics* 37(9–10): 744–753.

McAtamney, L. & Corlett, E.N. 1993. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics* 24(2): 91–99.

Moore, J.S. & Garg, A. 1995. The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. *American Industrial Hygiene Association Journal* 56(5): 443–458.

NIOSH 1994. *Applications manual for the revised NIOSH lifting equation*. Cincinnati, U.S. Department of Health and Human Services (DHHS) Publication No. 94–110.

Norman, R. & Wells, R. 1998. Ergonomic interventions for reducing musculoskeletal disorders: an overview, related issues and future directions. Retrieved from: *http://www.qp.gov.bc.ca/rcwc/research/norman-wells*

Occhipinti, E. 1998. OCRA: a concise index for the assessment of exposure to repetitive movements of the upper limbs. *Ergonomics* 41: 1290–1311.

Punnett, L. & Wegman, D.H. 2004. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology* 14: 13–23.

SLIC European inspection and communication campaign 2010. Lighten the load. Retrieved from: *http://www.handlingloads.eu/en/site/18/19*

Waters, T.R., Putz-Anderson, V., Garg, A. & Fine, L.J. 1993. Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics* 36(7): 749–776.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

An ergonomic intervention of the blood testing task for the prevention of WMSDs

Peng-Cheng Sung
Department of Industrial Engineering & Management, Chaoyang University of Technology, Wufong, Taichung County, Taiwan

Hsiang-Yi Lee & Yu-Ting Gong
Department of Information Management, Nanhua University, Chiayi, Taiwan

Chao-Ching Ong
Industrial design consultant, Ergonomic intervention on worker's WMSDs project, Taipei, Taiwan

Chi-Yong Chen
Institute of Occupational Safety and Health, Taipei, Taiwan

ABSTRACT: This study is to develop an ergonomic intervention proposal for the blood testing task in a blood donation center. When performing the task, the medical technologist must raise the test tube high above the head and visually check the react of blood at the bottom of test tube. The technologists complain about sore neck and shoulder. To develop the intervention, this study conducted a field investigation and RULA assessment. It is noted that the working posture involved shoulder elevation and neck extension and the resulting RULA score is 7, which requires intervention. Based on the findings in the field investigation, two proposals are made. One is to use armrest for the support of upper arm and the second one use mirror to reflect the image of test tube bottom. The latter prevents technologist from shoulder elevation and neck extension. The RULA score is 2, indicating that the intervention is helpful.

Keywords: ergonomic, intervention, WMSD, RULA, blood testing

1 INTRODUCTION

Work-related MusculoSkeletal Disorders (WMSDs) are an important issue in occupational safety and a common cause of health problem. (Bernard, 1997, European Agency for Safety and Health at Work, 2007). Many studies have been conducted to investigate the casues of WMSDs and the way to prevent it (Albert et al., 2005, Colombini & Occhipinti, 2006, David et al., 2008, Herbert et al., 2001, Melhorn, 1999). As the study indicated, the causes of WMSDs are very complicated. These factors range from workplace related, individual, and psychosocial (David et al., 2008, Menzel et al., 2004). As there are so many factors that could lead to WMSDs, it is important to identify the main courses for the development of an effective intervention. Among these factors, US occupational safety and health administration (OHSA) suggested that workplace risk factor along are necessary to produce WMSDs (US department of Labor, 2000). Workplace risk factors are the physical demands when performing the task. Some of the commonly known factors are posture adopted, force applied, frequency and repetition of movement exposed and vibration experienced … etc. (Burdorf and Sorrock, 1997, Kilbom, 1994, Melhorn, 1999). As the factor varies from task to task, it is important to develop customized ergonomic intervention base on the job characteristic of each task. However, not every task units is specialized in ergonomic and need consultation from ergonomic experts to develop their own intervention plan. Under this circumstance, the Institute of Occupational Safety and Health (IOSH) in Taiwan launches a multiple years filed ergonomic intervention project. The project sends out ergonomists to work with organizations that request for consultation to develop the ergonomic intervention proposal. The project then submits the ergonomic intervention proposal to the organization.

Taipei blood donation center, an organization that put emphasis on employee's occupational safety and health, filed a request to the IOSH for

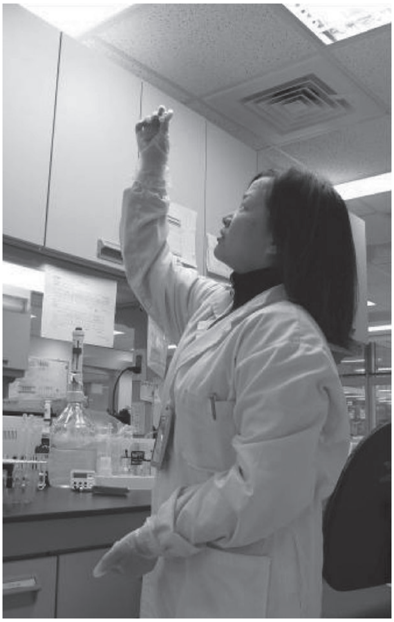


Figure 1. Photograph of the blood testing task.

ergonomic intervention consultation. The main target is on the blood testing task. The blood testing task is a key operation in blood donation center. It visually screens the donated blood and evaluate whether there is any certain effect exit preventing the use of donated blood on other patient.

To perform the blood testing task, the medical technologists first drop some chemical reagent drops into the test tube and then soon raise the test tube high above the head and slightly shake for few seconds. Then they visually check the reaction of blood clot at bottom of test tube and make the evaluation (Fig. 1). In order to check the result clearly, the technologists must raise the test tube high above the head and look and the bottom with a light from the opposite side of test tube. The process is quick, usually finishes within 1 min. The average frequency of the test is about 300–400 times everyday. There are some complains about sore neck and shoulders.

The purpose of this study is to propose an ergonomic intervention for this task to prevent the suffering of WMSDs. The field investigation and RULA evaluation were conducted in order to draw an ergonomic intervention proposal.

2 METHODS

Before the development of the ergonomic intervention, a field investigation was conducted to collected necessary information such as posture adopted, task frequency ... etc. This study also completed the Rapid Upper Limb Assessment (RULA) developed by McAtamney & Corlett (1993) to this task. First, a field investigate was conducted. The ergonomist went to the lab and watched how the test was conducted, record the layout, task frequency and took necessary pictures. The study also interviewed the technologist to understand some key points of the task. The study completed RULA after the investigation to check the risk score. The second step was to develop an ergonomic intervention proposal base on the findings of first two steps.

The development of ergonomic intervention includes three steps. The first step was to identify if there is an alternative for using mechanical tool to perform the task. If there is any alternative, then evaluate the suitable tool that could be recommended to the operator. If there is no alternative, the study then evaluates if there is any possibility to change the body parts that technologist used to apply force. For example, is it possible to use the leg to press the foot pedal instead of using hand to press the button, or is it possible to use trunk to carry the load instead of using hands to carry the load … etc. The ergonomist then review the working posture, try to suggest a functional posture that allow operators to perform the task easily, without bring extra burden onto human body. Figure 2 shows the evaluation flow chart that helps this study to develop the proposal.

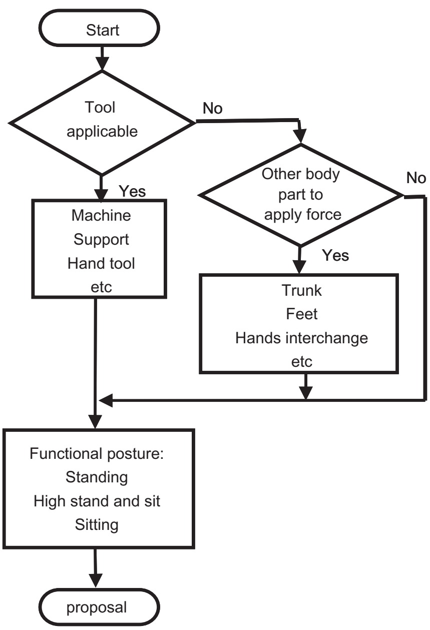


Figure 2. The evaluation flow chart.

3 RESULT

3.1 *Field investigation*

From the field observation, it is noted that some of the WMSDs risk factors may exist. Table 1 listed the risk factors noted from the field investigation. The study noted that the technologists are performing the task in an awkward body posture that may raise the WMSDs concern. The task is performed in the condition described below. The task is conducted in a lab and the samples and reagents are placed on the table. The light source is mounted on the ceiling. The task is performed in a standing posture. As the technologists raise the test tube higher than the head and visual check it, there are four points worth attention. First, as the upper arm is raised above shoulder, it means that the shoulder is elevated. The elevation of shoulder involves the movement of scapular and clavicle bone, which means many muscle are used to support the stability of the elevated shoulder. This may be the major cause of musculoskeletal discomfort. Second, the head titles back ward a little bit which imply a constricting neck muscle. This could be another source of discomfort. Third, the forward arm is also flexed in an upright position as technologist needs to bring the test tube I align the eyesight for a better visual acuity. This could also apply more loads to the arm. Fourth, the hand is flexed too to hold the test tube in a tilt position. Although the duration of each task is not long, only takes about 1 min, it repeats 300–400 times in a work shift per person. The frequency factor could also be another major cause of musculoskeletal discomfort complains.

During the interview, the technologists pointed out the important of clear visual acuity. They emphasizes that a slight difference in the clotting of blood would lead to very different evaluation result. Besides, the technologists also mentioned the important of testing time as the clotting of blood changes in a very short time. Therefore, the technologist must raise the test tube up and make the evaluation soon after the chemical reagent was dropped into the test tube. Therefore, they must be very concentrate when performing the test.

Table 1. Summarized WMSDs risk factors noted in the task.

Risk factor	Description
Posture	Shoulder elevation
Posture	Neck extension
Posture	Forward arm flexion
Posture	Hand flexion
Frequency	300–400 per work shift per person

The result of RULA assessment further supports the study's concern. The RULA score is 7, which means the task in action level 4 (AL 4). This means that immediate intervention is needed and must be implemented soon. Therefore, this study went further to develop the ergonomic intervention proposal based on the intervention flowchart and the finding of field evaluation.

3.2 *Ergonomic intervention proposal*

Following the intervention flowchart, this study first checked if there is an alternative to use mechanical tool to perform the test. However, raising and shaking the test tube is not a load demanding task, therefore, it is not necessary to find any tool that replace hand. Second, as the load is not demanding, there is no need to replace the hand with other body part to hold the test tube and shake it. Third, the posture adopted was checked by this study. Based on the finding from the field investigation, this study suggests that modifying the posture used in performing the task is needed. A first idea is to provide a support for the upper arm. This could help to relieve the load applied on the shoulder when raising the upper arm. Although this proposal could not solve the problem of constricting neck muscle in the same time, it is easy to implement, only need to install an arm support on the table (Fig. 3). Besides, it is suggested to further install a light source at sitting eye height. This helps to provide a better luminance and prevent technologist to raise the test tube too high above head just because the only light source is installed on the ceiling.

However, to propose a functional posture that fit the task without causing body load, this study reexamines the posture requirement of performing

Figure 3. The sketch of armrest proposal.

the test task. From the interview result, it is clear that the key point of this test is to check the bottom of the test tube as the clotting of blood sinks near the bottom of test tube. This is the major reason that leads to the elevation of shoulder and extension of neck. Therefore, the fundamental issue is to propose a posture that could easily check the bottom of the test tube. To solve the problem, a mirror is utilized in the proposal to reflect the image of test tube bottom. By checking the image in the mirror, the technologists do not need to raise the test tube high above the head. This eliminates the need to elevate the shoulder and extend the neck.

The proposed ergonomic intervention by introducing a mirror is summarized into four parts. First, the mirror is installed on the table at height of sitting eye height and slanted back little bit. The degree of slant angle and mirror height should be adjustable to match each technologist's dimension and preference. Second, a light source is installed at bottom and top of the mirror to provide enough luminance for visual check task. The technologists sit on the chair and take the test tube in front of the mirror and shake it. The technologists then look at the image of test tube in the mirror and make the evaluation. Under this circumstance, the technologists perform the task in a normal sitting posture with upper arm flex a little bit and forward arm flexed to bring the test tube in front of the mirror. The head only needs to tilt down a little bit to look at the image in the mirror. Figure 4 is a sketch that illustrates the idea of this intervention.

To further assure the effectiveness of this intervention, this study use RULA to assess the proposal. As there is no much elevation and flexion, the resulted RULA score is 2 and this indicate that the task is in action level 2 (score 1–2). This means the task is acceptable, no further intervention is needed.

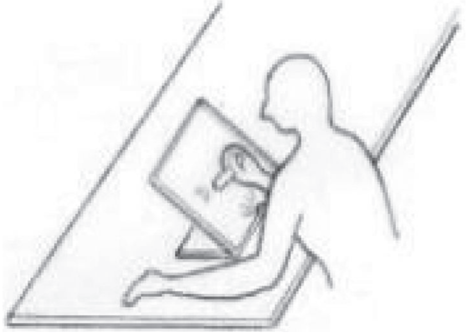


Figure 4. The sketch of ergonomic intervention by introducing the use of mirror.

4 DISCUSSION

The draft of ergonomic intervention was submitted to the blood testing lab for their opinion. The feedback is positive yet. The medical technologists accept these two ideas in the first glance. For the arm rest proposal, they are sure this provides some support to their arms and could relieve the workload on the shoulder. However, this still bring excessive workload on the neck muscle.

As to the mirror proposal, the technologists appreciate the new working posture that eliminates the need of shoulder elevation and neck flexion. However, the key point is the clearance of the image in the mirror. If it is not clear enough, it will force the technologist bend their head down close to the mirror for a better image. This will result in a sore neck and poor evaluation accuracy. In the mean time, the new proposal alters the way technologist used to perform the evaluation. Resistance of the new proposal from some of them is expected. The resistance will be higher when this proposal does not bring any benefit in the beginning and even bring adverse effect such as sore neck, poor eyesight and bad performance. At the end, the technologist will refuse to adopt the new proposal and the new proposal will be failed.

Therefore, to make sure the new proposal could be implemented successfully, following measures are needed. First, assure the grade of mirror installed is high enough to provide clear images. Second, the technologist should develop a SOP for the new proposal themselves so that no much conflict exists in using the new method. Third, there should be a training course for the new working procedure to make sure technologists are familiar with the new method and the performance is not affected. Fourth, conducts benchmarking to the original method to assure technologist that the performance will not be affected by the new method. It is believed the new proposal could be accepted by the medical technologist and reduces their musculoskeletal discomfort after these measures are implemented.

5 CONCLUSION

Two ergonomics intervention are proposed in this study. The armrest proposal is easy to implement but with limited effectiveness. The mirror proposal substantially changes the working posture and shows greater benefit based on RULA result. While the cost of mirror proposal is very few, it implies some ergonomic intervention could be very cost effective and revision of working posture or provide some support is worth study. The result also support the use of intervention flowchart developed in this study. It helps to derive a useful intervention proposal following the logic of this flowchart. It is hoped that all these could encourage ergonomists to develop more ergonomics intervention in various industries for the prevention of WMSDs.

ACKNOWLEDGEMENT

This study is partial supported by the Institute of Occupational Safety and Health (IOSH), Taiwan. The authors would also show gratitude to Taipei blood donation center for their valuable assistance.

REFERENCES

Albers, J., Estill, C. & MacDonald, L. 2005. Identification of ergonomic interventions used to reduce musculoskeletal loading for building installation tasks. *Applied Ergonomics* 36: 427–439.

Bernard, B.P. (ed.) 1997. Muskeletal disorders and workplace factor. In *A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back*. National Institute for Occupational Safety and Health, Cincinnati, OH.

Burdorf, A. & Sorock, G. 1997. Positive and negative evidence of risk factors for backorders (review). *Scand. J. Work Environ. Health* 23, 243–256.

Colombini, D. & Occhipinti, E. 2006. Preventing upper limb work-related musculoskeletal disorders (UL-WMSDS): New approaches in job (re)design and current trends in standardization. *Applied Ergonomics* 37: 441–450.

David, G., Woods, V., Li, G. & Buckle, P. 2008. The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal to risk factor. *Applied Ergonomics* 39: 57–69.

European Agency for Safety and Health at Work. 2007. *A European campaign on musculoskeletal disorder*. Office for Official publications of the European Communities, Luxembourg.

Herbert, R., Dropkin, J., Warren, N., Sivin, D., Doucette, J., Kellogg, L., Bardin, J., Kass, D. & Zoloth, S. 2001. Impact of a joint labor-management ergonomics program on upper extremity musculoskeletal symptoms among garment workers. *Applied Ergonomics* 32: 453–460.

Kilbom, A. 1994. Assessment of physical exposure in relation to work-related musculoskeletal disorder-what information can be obtained from systematic observation? *Scand. J. Work Environ. Health* 20: 30–45.

McAtamney, L. & Corlett, E.N. 1993. RULA: A survey method for investigation of work-related upper limb disorders. *Applied Ergonomics* 24(2): 91–99.

Melhorn, J.M. 1990. The impact of workplace screening on the occurrence of cumulative trauma disorders and workers' compensation claims. *J. Occup. Environ. Med.* 41, 84–92.

Menzel, N.N., Brooks, S.M., Bernard, T.E. & Nelson, A. 2004. The physical workload of nursing personnel: association with musculoskeletal discomfort. *International Journal of Nursing studies* 41: 859–867.

US Department of Labor Occupational Safety and Health Administration. *Ergonomic program: Final Rule*. 29 CFR part 1910. 65(20).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Ergonomic intervention technique for WMSDs prevention

Chi-Yuang Yu
Department of Industrial Engineering and Engineering Management, National Tsing-Hua University, Taiwan

Dong-Sheng Shi, Cheng-Ping Chang & Sweo-Chung Yang
Institute of Occupational Safety and Health, Council of Labor Affairs, Taiwan

Wen-Yu Ye
Industrial Safety and Health Association of ROC, Taiwan

ABSTRACT: The objective of this paper is to demonstrate an ergonomic intervention technique for preventing Work-related MusculoSkeletal Disorders (WMSDs). WMSDs are a serious occupational safety and hygiene problem in industrial countries, and generally agreed that the occupational risk factors of WMSDs are over-exertion, high repetition, unfavorable temperature, vibration, and awkward working posture. A protocol technique is developed and to seek a design improvement to eliminate or reduce these risk factors. This technique consists of an ergonomic intervention process chart, and 3 standard operation work sheets. A driver's seat final assembly operation is used as an example to shows that ergonomic intervention can be done, step by step, by following the chart and work sheets, and the result evaluation shows that musculoskeletal stress is reduced, while subjective rating and productivity are increased. Comparative studies indicate that this technique is applicable to field studies in Taiwan, and most ergonomic intervention cases illustrated in the literature as well.

Keywords: work related musculoskeletal disorders, WMSD, WMSD prevention, ergonomic intervention

1 INTRODUCTION

Work-related MusculoSkeletal Disorders (WMSD) are a serious occupational safety and hygiene problem in industrial countries, affecting not only the health of workers but also causing tremendous cost to industry due to work absenteeism, medical expenses, and productivity loss (Liang 2003). According to EU, US, and Japanese statistics, the percentage of occupational injury lost work days that are attributable to WMSD is consistently high: EU 38% (ERUOGIP 2007), US 32% (BSL 2007), and Japan 41.2% (Suka & Yoshida 2005). The total loss due to WMSD in the EU is as high as $216 billion US, and in the US itself $168 billion. In Taiwan, the percentage is about 33%, resulting in a loss of $2 billion NT each year (BLI 2010).

Although the causes of WMSD include diverse generic and occupation specific factors, nevertheless it is generally agreed that the leading occupational causes of WMSD are over-exertion, repetitive motion, vibration, unfavorable temperature, and awkward working posture (Lee 1999).

To prevent WMSD, the approach of this study is to identify key risk factors of WMSD, and then for each risk factor, to design a new work method and/or workplace that eliminates or lowers the severity of these factors, so that the musculoskeletal stress of the worker is lowered and WMSD are prevented.

2 METHOD

Aiming to eliminate or reduce the risk factors of WMSD, we have developed a protocol technique for WMSD prevention: an ergonomic intervention process chart (see Figure 1 below) and 3 worksheets (see Figure 2 below). The process chart consists of 5 phases: "work observation," "risk factor diagnosis," "design improvement," "result evaluation," and "other intervention"; the 3 worksheets are used to assist information recording and data analysis for intervention.

"Work observation" is used to record the current work situation, such as workplace layout, equipment and tools, working parts and materials, worker characteristics, as well as task analysis of the work performance. This phase is meant to collect background information for "Risk factor diagnosis."

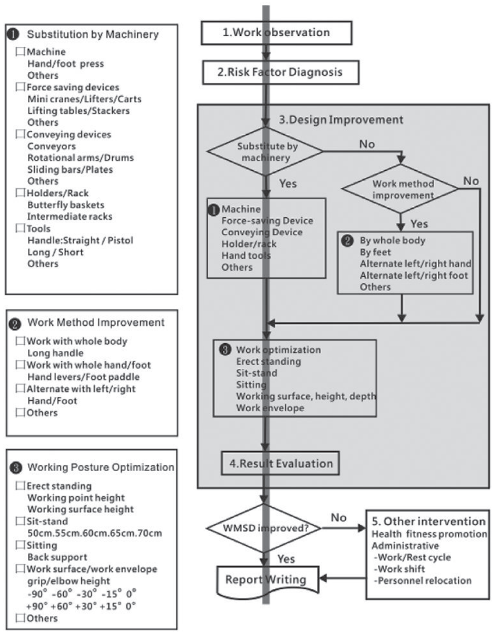


Figure 1. Ergonomic intervention process chart for the driver's seat final assembly.

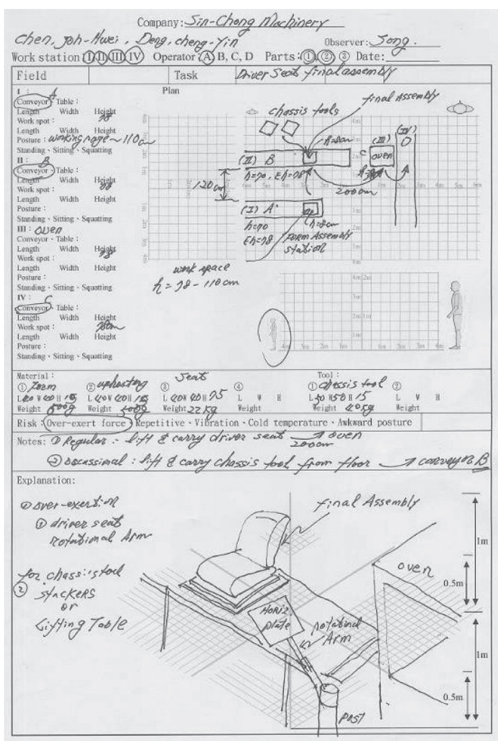


Figure 2. SOP work sheet #1 for the driver's seat final assembly.

"Risk factor diagnosis" analyzes and identifies the key risk factors (e.g. over-exertion, high repetition, unfavorable temperature, vibration, and awkward posture) that lead to musculoskeletal stress during task performance. This is the core process in ergonomic intervention.

"Design improvement" is seeking a better work method and/or workplace design via "machine substitution," "work method improvement," or "work posture improvement." This is the idea development stage in which factors identified by "risk factor diagnosis" are reduced or eliminated.

"Result evaluation" gauges the effectiveness of the design improvement, e.g. reduced musculoskeletal stress and increased satisfaction rating and work efficiency.

If the result is not entirely effective, then "other interventions," such as "health fitness promotion" or "work-rest scheduling" are recommended.

3 SOP worksheets are used in recording information and data analysis for intervention (Figure 2). The first sheet is used to record and draw information about the current work situation, to analyze risk factors, and to sketch design improvements. The second sheet is used for photo bookkeeping. The third sheet is used for risk factor analysis, such as KIM or biomechanical analysis of joint stress.

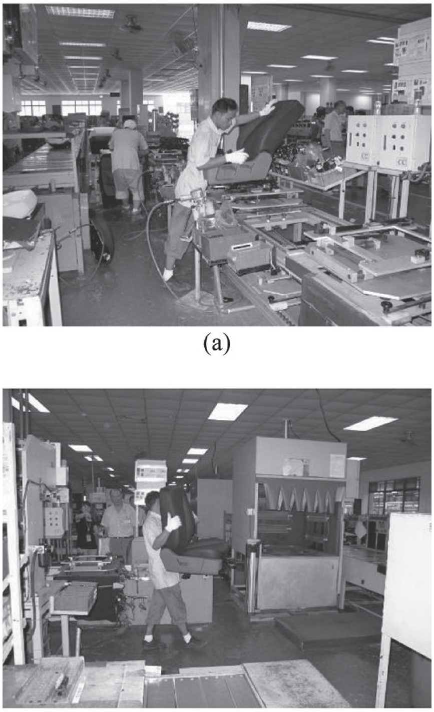
This intervention technique will be demonstrated using the example of a "driver's seat final assembly" operation in the following Section.

3 RESULTS

The intervention process of this driver's seat final assembly is marked with a red line in Figure 1. The logic of the process is described as follows.

3.1 *Work observation*

Driver's seat final assembly means fitting the inner form onto the seat chassis and covering it with external upholstery. Upon completion of assembly, the worker unloads the seat and carries it to the oven for temperature setting. This final assembly involves 3 workstations: the inner form assembly, the final assembly, and the temperature setting oven, as illustrated in Figure 2 on SOP work sheet #1 and Figure 3 below. The inner form assembly workstation is the last workstation of conveyor A; the final assembly workstation is likewise the last workstation of conveyor B. These 2 conveyors are parallel and 120 cm apart. The oven is 2 meters away from the end of conveyor B. The working surfaces of all 3 of these workstations are 78 cm high and the working height ranges between 78 cm and 110 cm. The size of the inner form is about 40 cm × 40 cm × 15 cm, but only



(a)

(b)

Figure 3. (a) The final assembly of covering the form material with external vinyl upholstery. (b) The worker carries the final assembly of the driver's seat to the oven for temperature setting.

weighs 600 g. The external upholstery is the vinyl cover of the inner form, weighing 500 g. The final assembly of the driver's seat has a size of 40 cm × 40 cm × 75 cm, and weighs 22 kg. The worker is a 27-year-old male of 160 cm height and 60 kg weight.

First, the worker grabs the form materials from conveyer A, assembles them with glue, and compresses them with the power press on conveyor A for 5 seconds. Then he takes the form assembly and turns to the final assembly workstation of conveyer B. He fits the form assembly onto the seat-pan chassis, and then grabs the external upholstery from conveyor B, wraps it over the form assembly, and pulls the rim of upholstery beneath the chassis. Then he tilts the seat 90 degrees to reveal the button of the chassis, whereupon he fastens the upholstery with a stapling gun at 20 sites. After completing this final assembly, the worker then unloads the seat and carries it 2 meters to the oven for temperature setting.

The work cycle is 3 minutes; the worker must complete 160 repetitions in an 8-hour work shift.

3.2 *Risk factor diagnosis*

The key risk factor in this work activity is over-exertion. The worker has to lift the completed driver's seat, hold it against his chest, and carry it 2 meters to the oven every 3 minutes. The driver's seat weighs 22 kg. The Key Indicator assessment Method reveals a load rating point of 4, a posture rating point of 4 (because of the bulky load and the great horizontal distance), and a working condition of 0. Multiplied by a time rating point of 4, the total risk score is 32. This score indicates a physical overload for the average worker, thus redesign of the workplace is recommended.

The analysis of biomechanical load moment on the shoulder is 28 N-m, on the elbow 25 N-m, with a compression force on L5/S1 of 1585 N. This compression force is 46.7% of the action limit (3400 N). It can be considered a heavy workload since the worker must perform 160 repetitions a day.

3.3 *Design improvement*

To develop a feasible design for reducing over-exertion in this work activity, various idea alternatives are checked, as depicted on the intervention process chart (see Figure 1). First, "machine substitution" is checked, such as machines, force-saving devices, conveying devices, tools, holders or racks, and so on, as listed on the left block diagram of the chart. Fortunately, a rotational arm seems to be a feasible substitution for the manual force of the worker.

The rotational arm consists of a main post of 10 cm diameter and 90 cm high. On the top of the post is a bearing cap. The cap is fabricated with an extendable horizontal bar. On the end of the bar there is a horizontal plate for holding the driver's seat. (see Figure 4.)

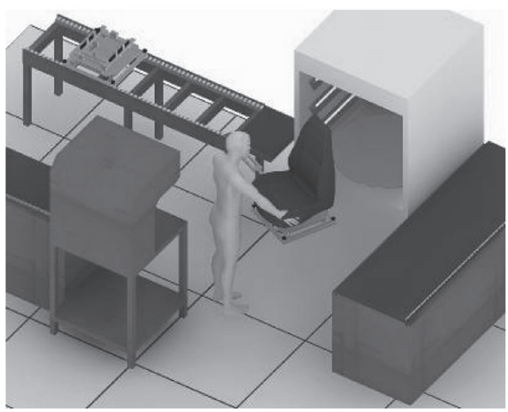


Figure 4. Idea sketch of rotational arm for carrying the driver's seat to the oven.

The rotational arm is installed in between the final assembly workstation and the oven. The worker is now tasked to direct and insert the horizontal plate under the driver's seat. The rotational arm carries the driver's seat, so the worker need only swing it to the oven, and push it inside. By this method, the worker's load is significantly reduced.

Since "machine substitution" is feasible ("Yes" on the ergonomic intervention chart), the process moves forward to "work posture improvement." However, the working posture in this case is standing erect. This is a good working posture, as listed in the block, so the design improvement is complete.

3.4 *Result evaluation*

Based on the improved design, worker overexertion is significantly reduced, and work efficiency increased. Analysis by Key Indicator Method reveals that the load rating point is now 1(less than 1 kg), the posture rating point is now 1 (erect standing), and the working condition is 0. Multiplied by a time rating point of 1 (as no lifting is required), the total risk score has fallen to 2. This indicates that physical overload has been almost completely reduced.

Biomechanical analysis shows that after the improvement, load moment on the shoulder has been reduced from 28 N-m to 3 N-m, on the elbow from 25 N-m to 2 N-m, and the compression force on L5/S1 has been reduced from 1585 N to 392 N, which is just 11.5% of the action limit (3400 N).

3.5 *Other intervention*

Since the workplace and work method are improved, and the result evaluation shows that the musculoskeletal stress of the work can be effectively reduced, no other intervention is needed.

4 CONCLUSION AND RECOMMENDATIONS

Our technique is the result of a long-term evolution. To date, we have intervened in some 800 WMSD cases, so it gradually developed through trial and error. In general, this intervention technique is effectively applicable to 97% of cases. Because of its general application, it can be used as a DIY directive in guiding ergonomists, industrial hygienists, and field supervisors to improve workplaces and/or work methods toward the prevention of WMSD.

We have also applied this technique to cases that appear in the literature, such as Ergonomics in Action, Easy Ergonomics, etc. Most of the cases can be assessed by KIM and biomechanical analysis, as illustrated in the Result Section, and almost all improvement methods or tools used in those cases, such as working posture improvement, conveyors, stackers, lifting tables, tilters, etc. can be categorized by either "machine substitution," "work method improvement," or "work posture improvement." These results show that our protocol technique is also applicable for WMSD prevention interventions undertaken in foreign countries.

Nevertheless, this is still a knowledge-based technique, as its application requires a high degree of ergonomics knowledge and engineering experience. Based on this groundwork, it appears that the technique should go a step further: to evolve into a rule-based technique. A checking-typed process chart and diagrammatic SOP worksheets with figures or icons should be developed. The rule-based technique is aimed at reducing the dependency on ergonomics and engineering knowledge, so any personnel with basic training in ergonomics and/or safety and hygiene engineering would be able to improve workplace by DIY.

REFERENCES

BLI 2010. *2005–2009. Annual reports*. Bureau of Labor Insurance, Council of Labor Affairs, Taiwan, ROC.

BLS 2007. *Nonfatal occupational injuries and illnesses requiring days away from work (2002–2006)*, Bureau of Labor Statistics, USA.

EUROGIP 2007. *Musculoskeletal disorders in Europe-definitions and statistics*, EUROGIP.

Lee, C.L. 1999. MSDs of construction labourer, *Labor Safe and Hygiene News in Brief* 39(4): 10.

Liang, H.W. 2003. Labor investigate yearly report: A brief of WMSDs. *Industrial safety scientific and technological quarterly* 50: 2–6, Council of Labor Affairs Executive Yuan, Taiwan, ROC.

Suka, M. & Yoshida, K. 2005. *Musculoskeletal pain in Japan: prevalence and interference with daily activities*, Japan College of Rheumatology and Springer-Verlag Tokyo.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Prevalence of cumulative trauma disorders among office workers in a semiconductor industry

Rosnah Mohd Yusuff & Airing Zairose Abdullah
Department of Mechanical and Manufacturing Engineering, University Putra Malaysia, Serdang, Selangor, Malaysia

ABSTRACT: The health records of the employees in a semiconductor industry showed a drastic increase of musculoskeletal disorders for 2007 compared to 2005 and 2006. This may have a significant effect not only to the employees but also to the organization. If these problems remain unsolved, this may lead to poor work performance, increase medical leave, lower productivity and increase of health compensation claims. A study was conducted to determine the existence of cumulative trauma disorders among the office employees, identify level of severity and what are the probable factors that contribute to cumulative trauma disorders. A survey questionnaire, interviews and observation methods were used to identify the problems and to provide recommendations to prevent or rapidly response to work related upper limb disorders. Based on these probable factors, some suggestions and recommendations were proposed to reduce cumulative trauma disorders risks.

Keywords: workers, cumulative trauma disorders, severity level

1 INTRODUCTION

There are many studies that have investigated the occurrence of Cumulative Trauma Disorders (CTDs) among office workers. In office work, most of the back pains were caused by prolonged sedentary work, rather than the manual handling of heavy materials. Improper workplace design, where workers have to exert themselves, perform work in awkward positions and repetition have led to strains on the various body parts. Workplace physical factors affecting musculoskeletal disorders or CTDs include heavy physical work, lifting and forceful movements, awkward postures, whole-body vibration, and static work postures. Static work postures of prolonged standing, sitting, and sedentary work are isometric positions where very little movement takes place. These postures are typically cramped or inactive and cause static loading on the muscles (NIOSH, 1997).

Studies have shown that the extensive use of computer has a consequent effect on musculoskeletal discomforts and disorders (Punnett and Bergqvist, 1997), due to the task demand and non-neutral work.

Other physical risk factors associated with the keyboard usage include the design of the keyboard, the repetitiveness and the force applied in keyboarding task (Amell and Kumar, 2000). Studies (Aaras et al., 2001; Cook and Burgess-Limerick, 2001, 2002) indicated that supporting the arms (forearm and/or wrists) during keyboard and mouse use was a preferable posture for most computer users. Westgaard and Aaras (1985) described that the supporting of the arms was effective in reducing muscle activity; on the other hand, Hedge and Powers (1995) reported that these supports did not significantly change working posture or reduce neck/shoulder muscle activity. Albin (1997) viewed that the wrist supports on the work surface might reduce the wrist flexion/extension, resulting in more neutral wrist postures. Bendix and Jessen (1986) as cited in Nag et al., (2009) viewed that the wrist support might increase neck and shoulder muscle load.

A study was conducted to identify the risk factors of musculoskeletal diseases in a semiconductor industry that showed an increase trend of shoulder, elbow, arm, wrist, hand, finger and muscle disorders reported by the employees. The health records showed that there is a drastic increase in back pain for 2007 compared to 2005 and 2006 as shown in Figure 1. This work related upper limb disorders may have a significant effect not only to the employees but also to the organization. If these problems remain unsolved, this may lead to poor work performance, increase of medical leave, lower productivity and increase of health compensation claims.

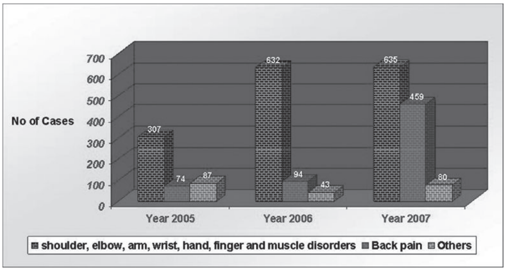


Figure 1. Reported cases of ctds in the company.

The prevention of musculoskeletal disorders is achieved by interventions, which reduce the probability and severity of injuries. It is estimated that, through ergonomic design up to one-third of compensable low-back pain in industry can be reduced (Levy and Bergman, 1995).

2 METHODOLOGY

Data was collected from medical claim records history to identify ergonomics problems faced by the employees. Since the most reported CTDs were from the office workers, a survey was conducted among the office employees, based on a ninety-five percent confidence level. By using random sampling, only thirty percent of 284 total office employees were surveyed. The study focused on computer users in the company. This group of employees was using the same type of workstation, laptop and desktop. The questionnaires were distributed according to the sampling size predetermined earlier.

The questionnaire was developed to evaluate individuals' exposure to postural, force and muscle activities that can contribute to cumulative trauma disorders. The questionnaires were adapted from Rapid Upper Limb Assessment (RULA), body part symptom survey and Occupational Safety and Health. A level of severity has been used and defined into four: level one is defined as an existence of work-related pain without medication; level two is defined as pain persists even though at home and/or affects the sleep of sufferers; level three is defined as pain persists which requires a visit and advice from doctors twice or more times for work-related pain; level four is defined as the pain persists and requires medical leave(s) for resting and recovering.

Workstation comfort survey was also developed to strengthen the outcomes from body symptoms survey. Interviews among the sampled workforces to get feedback related to personal health, acceptance of workplace design, the level of comfort experienced during working were also conducted. Observation was also performed and photographs taken. The photographs were reviewed and discussed in order to strengthen the findings from the survey. Observations were made on body posture during sitting and typing and used for REBA analysis.

3 RESULTS

3.1 *Level of severity*

Out of 284 workers surveyed, only 71 or about 30% of the workers responded to the survey. Figures 2–4 showed the occurrence of CTDs for various levels of severity with lower back, neck, shoulder and upper back affecting more respondents at severity level 1 and 2. The wrist was reported as having severity level 1 and 2 by all the respondents. Only severity level 3 affecting the lower back was reported by 10% of the respondents and the rest of the body parts by less than 10%. However, it can be clearly seen that the upper limbs are more affected at all three severity level compared to the lower limbs. There were six cases of severity level 4 as shown in Table 1.

The severity of the injury was assessed using the Rapid Entire Body Assessment (REBA) score sheet (Hignett and Lynn, 2000), for a typical posture. Based on observations and pictures taken, the various scores for the body parts were assessed following the procedures in the score sheet.

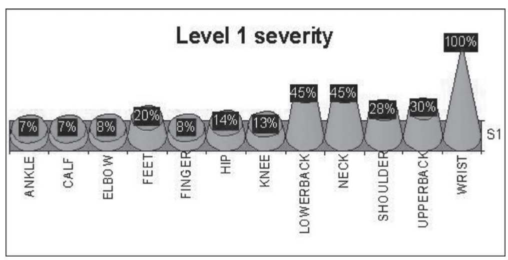


Figure 2. Body parts with level 1 severity.

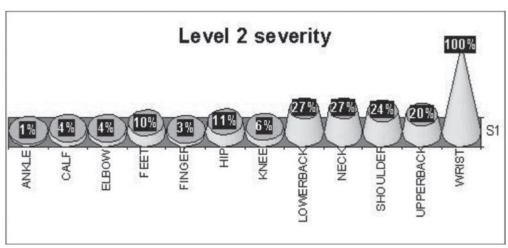


Figure 3. Body parts with level 2 severity.

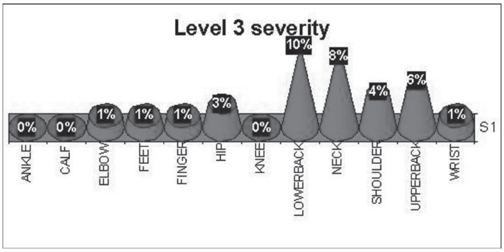


Figure 4. Body parts with level 3 severity.

Table 1. Level 4 severity cases.

Type of Medication	Number of medical leave
Employee was given muscle relaxation	6
Employee was given painkiller	2
Employee required undergoing X-Ray for wrist	1
Employee required undergoing surgery for wrist and hospitalised for wrist and back pain physiotheraphy.	34
Employee visited panel clinic for neck pain; given medical leave for resting	1
Employee had udergone operation for right finger	14

The posture and score of each of the body parts were estimated as follows:

- Trunk is flexed twenty degree (2).
- Neck is extended (2) sitting (1).

The employee has sitting balance, so the load/force score is zero (<5 kg exerted). The sub-total (5) and the load/force score (0) was added to get the Score A = 5 for this part.

For Group B, the score is:

- Upper arm flexed between forty-five degree and ninety degree (3).
- Lower arm is flexed less than sixty degree (2).
- Wrist is deviated fifteen degree flexion or extension (2).

The subtotal for this group as given in the score sheet is 5 and the coupling score (1) is added to get Score B (6). With Score A (5) and Score B (6), Score C (7) is obtained from the work sheet and the activity score (+1 added) as employees are doing repeated work with small range of action more than four times per minute. The total REBA score is 8, indicating high risk of injuries to employees and that action is necessary soon to reduce the risk level.

3.2 *Effect of age, height and weight*

Analysis was done based on age, height and weight in order to identify probable factors contributing to cumulative trauma disorders. In this study, each of the groups has been divided into sub-groups. The sub-groups are as follows:

1. Age was divided into two groups: 20 to 39 years, and 40 and above.
2. Height: less than 159 cm, and above 159 cm.
3. Weight: less than 59 kg, and above 59 kg and above.

For age, the younger group showed more effect of CTDs on the body parts compared to the older group. There were 25 cases of neck pain, 24 lower back, 18 upper back and 16 wrist pain in the lower age group for severity level 1 compared to only 5 neck and lower back pain each and 3 for upper back and wrist pain each in the higher age group. For severity level 2, the same trend was observed but higher reported cases of wrist pain at 51 in the younger group compared to only 12 in the higher group. For severity level 3, there was one case for wrist, 4 lower and 2 upper back pain in the older group.

Those with higher heights were more affected by CTDs compared to those with lower heights. There were 25 cases of lower back, 24 neck, 18 wrist, 16 upper back and 15 shoulder pain of severity level 1 for heights above 169 cm. There were 56 cases of wrist pain out of 68 cases affecting the taller group.

CTDs affected those who are heavier more. Most of severity level 1 and 2 affected those with weight of more than 59 kg. For level one severity, there were 25 cases of neck pains, followed by lower back (18) and wrist (15) for the heavier group. At severity level 2, there were 41 cases of wrist pain for the heavier group compared to only 18 for the lower weight group.

3.3 *Comfort of workstation design*

A comfort survey of the workstation design was conducted on 16 respondents. The respondents were asked regarding their work posture, desk, chair, footrest, keyboard and workspace/layout.

For adopted posture relative to the workstation design, all 16 respondents indicated discomfort of the neck, shoulder and elbow, followed by back pain (14) and wrist (12).

The respondents also indicated that the height of the chair gives inappropriate elbow height while typing (12). The chair is also unable to sufficiently close the gap between the chair and the desk or cubical to enable the elbow to be perpendicular to the shoulder. This can lead to shoulder and elbow discomfort. On top of that, the chair has insufficient back support to employees. These may cause neck and back pain or discomfort especially to those who are working for longer hours.

Table 2. Summary of number of cases for severity level 1, 2 and 3 for each body part.

Body parts	Level 1 number of employees affected	Level 2 number of employees affected	Level 3 number of employees affected
Ankle	5	1	0
Calf	5	3	0
Elbow	6	3	1
Feet	14	7	1
Finger	6	2	1
Hip	10	8	2
Knee	9	4	0
Lower Back	32	19	7
Neck	32	19	6
Shoulder	20	17	3
Upper Back	21	14	4
Wrist	71	71	1

All the respondents expressed discomfort with the position and design of the computer keyboard. The keyboard for laptop users is not tilted, the distance is not adequate for elbow comfort, no hand/forearm support and that the width of the keyboard is not appropriate. The current hard-surface arm rest available is not appropriate and also transfers heat to the wrist causing hurt to the nerve of the wrist.

Though most of the respondents indicated that they are able to rest their feet firmly, they also doubted that an appropriate footrest is available if needed. All the respondents expressed comfort with the space and room layout design and only about 3 expressed overall discomfort with the desk/cubicle design. The chairs, keyboards and laptop-stand design are inappropriate for longer working hours leading to neck, shoulder, elbow, arm and back discomfort or pain.

4 DISCUSSION

Based on surveys, interviews and observations, it is estimated that about seventy percent of the office employees were practising incorrect body posture during typing and sitting. A summary of the affected body parts as given in Table 2 clearly showed that some interventions are necessary to prevent them being aggravated further. The wrist, lower and upperback and the shoulder pains reported at severity level 1 and 2 may lead to severity level 3 and 4 over time. The office workers worked for long hours and repetitively. These are the risk factors for CTDs (Fernandez J.E, 1995).

For the workers with some awareness of the ergonomics risk factors, they had improvised ways of making themselves comfortable. As observed in the offices, paper files have been used to provide for a tiled angle for laptop, a box has been used as a footrest and cushions have been used to support the back as means of improving their work posture and greater comfort.

Though the study showed that there is a tendency for those who are taller and heavier reported more instances of CTDs on the various parts of the body, workstation design and working duration are the main contributing factors to CTDs. Sixty-three percent of the workers sat for 2 hours or more before taking a break, whereas the rest after an hour. The REBA score of 8, clearly indicated that immediate changes are necessary before the situation worsened as can be seen in cases involving severity level 4. Another cause of concern is the reported cases among the younger group. This again requires special attention before the pain cumulates in later years.

The study also indicated that the chair design posed some problems to the workers. Saporta (2000) suggested that the headrest, full-size backrest, and the seat pan should receive the weight of the head, the trunk, and the thighs. If the depth of the seat is beyond the buttock-popliteal length, the user will not be able to engage the backrest effectively without unacceptable pressure on the backs of the knees. Also, the deeper the seat, the greater the problems of standing up and sitting down will be (Pheasant, 2001).

5 CONCLUSIONS AND RECOMMENDATIONS

The study has shown that cumulative trauma disorder is prevalent among the office employees of the semiconductor industry. The office employees are experiencing cumulative trauma disorders starting from severity level one until the chronic level which is severity level four. The wrist has been shown as the most affected. Therefore, more investigations are required to improve the posture of the wrist by looking at the workstation design and its arrangement. Investigations have shown that the arrangement of the equipment (monitor, keyboard, document holder, etc.) not corresponding to the working task could lead to unfavorable postures (Benninghoven A, 2002). This result supported by the REBA score, which is 8 indicating high risk of injuries to employees and that action is necessary soon to further assess the workstation design to improve the work posture.

The contributing factors may lead to poor working performance and productivity because the higher the discomfort, the higher the risk for ergonomics problems to occur. The company has to look at these contributing factors and investigate them. Even though the occurrence of CTDs is currently at severity level one and two, over time, this may lead to more serious injury.

The chairs, keyboards and laptop-stand design are inappropriate for longer working hours which lead to neck, shoulder, elbow, arm and back discomfort or pain. When work is performed continuously and repetitively, it leads to cumulative discomfort or pain. Thus, it is suggested and recommended to provide suitable wrist support, back support and laptop tilters. Through careful design, meeting facilities, computers, storages and standard desk areas can be all located within one unit. VDUs can be mounted above the desk on a support or adjustable arm, creating an extra working area and minimizing unnecessary strain, thereby improving efficiency and enhancing the employee's job satisfaction (Brazier, 1993).

As observed, most workers have improvised their workstation design to increase their comfort. This is an indication to management that steps have to be taken to increase ergonomics awareness and provide training programmes to the workers. An increased knowledge on the right posture, ergonomics risk factors and the consequences can help the workers adjust their workstation to fit their needs, especially for those taller and heavier workers with more complaints of CTDs. Adjustable chairs and tables maybe appropriate to cater for the variations in the workers size.

Management must take serious heed on the occurrence of severity level one and two and reported CTDs among the younger age group. Epidemiologic studies of upper extremity disorders suggest that certain psychosocial factors (including intensified workload, monotonous work, and low levels of social support) have a positive association with these disorders. Lack of control over the job and job dissatisfaction also appear to be positively associated with upper extremity musculoskeletal disorders, although the data are not as supportive (Bernard et al., 1997). Whether these factors have an effect on the occurrence of CTDs among the workers need to be investigated also.

It is also recommended that a wider study to include the operators in the production floor and those involved in manual lifting be conducted. This can help the management evaluate the extent of ergonomics risk factors in the company and implement necessary actions. A comprehensive strategy on ergonomics may help the companies to increase productivity, reduce compensation claims due to illnesses, lost workdays, absenteeism and turnover. But most of all, it create a positive working environment where workers are satisfied in the knowledge that management cares.

REFERENCES

Aaras, A., Fostervold, K., Ro, O., Thoresen, M. & Larsen, S. 1997. Postural load during VDU work: a comparison between various work postures. *Ergonomics* 40(11): 1255–1268.

Aaras, A., Horgen, G., Bjorset, H.-H., Ro, O. & Walsoe, H. 2001. Musculoskeletal, visual and psychosocial stress in VDU operators before and after multidisciplinary ergonomic interventions: A 6 years prospective study. *Appl Ergon* 32: 559–571.

Albin, T. 1997. Effect of wrist rest use and keyboard tilt on wrist angle while keying. *The 13th Triennial Conference of the International Ergonomics Association*.

Amell, T. & Kumar, S. 2000. Cumulative trauma disorders and keyboarding work. *Int. J. Ind. Ergon.* 25(1): 69–78.

Benninghoven, A. 2002. Working conditions in call centers: Successfully and healthy. In *Proceedings of the Conference WWDU 2002 World Wide Work*.

Bernard, B.P. 1997. *A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back.* NIOSH.

Brazier, I. 1993. The ergonomic office – Fad or fiction? The facts. *Facilities* 11(5): 20–23.

Cook, C. & Burgess-Limerick, R. 2001. Forearm support and computer keyboard use. In M. Stevenson & J. Talbot (eds). *Proceedings of the 37th Annual Conference of the Ergonomics Society of Australia Inc.*: 101–106.

Cook, C. & Burgess-Limerick, R. 2002. Forearm support for intensive computer users: a field study. In *Proceedings of the 38th Annual Conference of the Ergonomics Society of Australia Inc.*

Fernandez, J.E. 1995. Ergonomics in the workplace. *Facilities* 13(4): 20–27.

Hedge, A. & Powers, J.R. 1995. Postures while keyboarding: Effects of a negative slope keyboard system and full motion forearm supports. *Ergonomics* 38: 508–517.

Hignett, S. & McAtamney, L. 2000. Rapid Entire Body Assessment (REBA). *Applied Ergonomics* 31: 201–205.

Levy, B.S. & Wegman, D.H. 1995. *Occupation health-recognizing and preventing work-related disease*. 3rd edition. Little, Brown and Company.

Nag, P.K., Pal, S., Nag, A. & Vyas, H. 2009. Influence of arm and wrist support on forearm and back muscle activity in computer keyboard operation. *Applied Ergonomics* 40(2): 286–291.

Pheasant, S. 2001. *Bodyspace: Anthropometry, ergonomics and the design of work*. Taylor & Francis Ltd.: London.

Punnett, L. & Bergqvist, U. 1997. *Visual display unit Work and upper extremity musculoskeletal disor ders: a review of epidemiological findings*. National Institute for Working Life—Ergonomic Expert Committee Document No. 1. Arbetslivsininstitutet, Solna, Sweden.

Saporta, H. 2000. *Durable ergonomic seating for urban bus operators*. Tri-Country Metropolitan Transportation District of Oregon.

Tittiranonda, P., Rempel, D., Armstrong, T. & Burastero, S. 1999. Effect of four computer keyboards in computer users with upper extremity musculoskeletal disorders. *Am. J. Ind. Med.* 35: 647–661.

Westgaard, R.H. & Aaras, A. 1985. The effect of improved workplace design on the development of work-related musculoskeletal illness. *Applied Ergonomics* 16(2): 91–97.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Musculoskeletal strain in Thai university hospital nurses: A pilot study

Pornkamon Thoicharoen & Pongjan Yoopat
Ergonomics Division, Faculty of Science, Rangsit University, Patumthani, Thailand

Kamiel Vanwonterghem
CERGO-MENSANA, Belgium

ABSTRACT: Musculoskeletal problems and work-related back pain are growing common risks among registered nurses. In Thailand, Low-Back Pain (LBP) is reported as one of the most frequent complaints among Thai nurses. The objective of this study was to assess musculoskeletal strain as pre-indicating risk for LBP in registered nurses at a Thai university hospital. The participants were twenty one female nurses. Their experienced discomfort was rated by means of the Subjective Workload Index (SWI) and Electro-MyoGraphy (EMG) recording the workload during patient lifting. As assessment criterion, the measured values were compared to the maximal capacity of the selected muscles, M. Trapezius and M. Erector Spinae recorded before and immediately after work. It was concluded that the total muscle strain in nurses has to be classified under "highly risky jobs" even lifting in a 4 nurse team. The imbalance between left and right the M. Erector Spinae is a worrying pre-sign for LBP.

Keywords: MSD in nursing, subjective workload estimation, electromyography

1 INTRODUCTION

Musculoskeletal problems and work-related back pain have become increasingly common risks among nurses. Various studies proved that the more frequent patient handling appears to correlate with the higher rates of low-back pain (Taptagaporn 1999, Hui et al. 2001, Hignett 2008).

Epidemiologic studies in several countries such as-England, United States, Denmark and Israel provide similar estimates of high frequent low back pain and diseases (Harber 1990). Moreover, a longitudinal study across eight years in 269 nurses employed by a large university in Switzerland found that over that period, almost all nurses indicated a similar degree of low back pain (Maul et al. 2003).

It is not surprising that also in Thailand low-back pain, reported as the most frequent complaint among Thai nurses (Taptagaporn 1999, Wongthanakit 2005), causes a considerable loss of valuable and qualified nurses. Wongthanakit (2005) studied the behavior to prevent low back pain as it is the predisposing factors among nursing personnel (231 registered nurses and technical nurses) in Governmental Hospitals using a self-administered questionnaire. The results revealed that 71% of the nurses had back pain within twelve months with a prevalence of 48% low back pain. The three most important ranked causes of back pain concern: 1) supporting patients to get up on bed, 2) lifting patient into the bed and 3) handling patient from one bed to another.

The objective of this study is to assess musculoskeletal strain in registered nurses at a Thai university hospital.

2 METHODS

2.1 *Participants*

Twenty one female nurses from a Thai university hospital aged 27 ± 6 years, with a Body Mass Index (BMI) of 20.9 ± 3.9 and 8 years experience volunteered to participate in the project. All the participants have formally agreed with the written and verbal information about the study procedures and methods.

2.2 *Measurements*

2.2.1 *Subjective Workload Index*

The participants rated their experienced discomfort using the Subjective Workload Index (SWI) questionnaire on a '0' to '10' scale, from no problem to extremely high (Yoopat 2004, Yoopat et al. 2006). The 6 stress factors (fatigue, risks for discomfort/injuries and diseases, concentration, complexity, work rhythm and responsibility) and 2 stress relieving factors (personal interest in the job and degree of autonomy) were averaged. A SWI score higher than 2.5 is interpreted as a warning level for psychosomatic strain and the higher SWI, the more critical the workload. The eventual participants' complaints were also recorded per body pain region.

The total SWI was calculated and evaluated as follows:

$$\mathrm{SWI} = [(\Sigma\ \mathbf{LF}) - (\Sigma\ \mathbf{CF})] / 8 \tag{1}$$

where, SWI = subjective workload index; Σ = summation; LF = load factors; CF = compensating factors.

SWI *evaluation scale*

≤ 2—Refers to negligible or no problem.

$1 \leq 2$—Minor problems, no action required.

$2 \leq 3$—Moderate problems: pass to detailed analysis and set priorities for specific problems on medium to long term.

$3 \leq 4$—Serious problems: calls for intervention for short term i.e. within one month.

$4 \leq 5$—Very serious problems: calls for intervention within one week.

≥ 5—Work should be interrupted and working conditions should be objectively analyzed.

From the level of SWI >2.5, a detailed analysis is advised.

2.2.2 *Electromyography*

Electromyographic measurements (EMG in micro-volts) were recorded during patient-lifting performed by a team of 4 nurses. Surface electrodes equipped with an amplifying system registered the electro-potentials of muscle bodies and stored data in a memory card (Portable EMG, ME 3000P, MEGA Instruments, Finland). The selection of studied muscle groups was made to the main muscles involved in patient lifting, M. Erector Spinae and M. Trapezius for the static and postural muscle load of this activity. The Maximal Voluntary Contraction (MVC) of both muscle groups, both left and right, was recorded before and immediately after work, and during simulated activities controlled by a neutral observer as shown in Figure 1. The duration of measurement was 20 min for each participant. The 20 min-EMG was compared with MVCs and the 15% MVC threshold (MSD risk on medium term) of the respective muscles. Dynamic activities of muscle fibres were characterized by alternation of contraction (high micro-volts) and relaxations (low micro-volts) with a certain amplitude. Static contraction of muscle fibres was almost permanent and varied around the average within small fluctuations. An eventual decline in muscle power (MVC) after work can be estimated as an indicator of muscles fatigue (Yoopat et al. 2006). An example of the EMG recording during lifting activities of the participants was shown in Figure 2.

Figure 1. During patient lifting activities of a 4 nurse team.

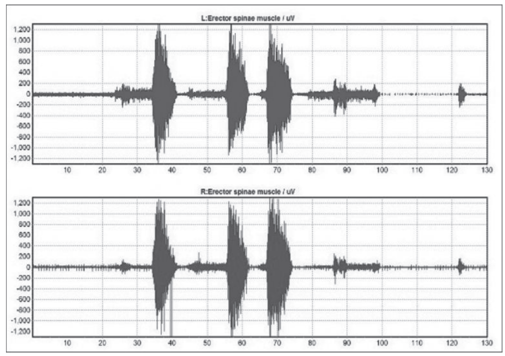


Figure 2. An example of the EMG recording during lifting activities.

An expression of the % used to interpret MVC has been applied as follows:

0% < MVC < 15%—No particular attention required.

15% < MVC < 30%—Analyze activities, reduce peaks.

30% < MVC—Represents respectively, the musculoskeletal disorders risks of low, moderate, and high.

2.3 *Statistical analyses*

The data were evaluated with descriptive statistics using means, standard deviations and ranges. Differences in initial and end of work MVC were tested with the Student's *t* test for pair observations and the correlation analysis. The results were considered statistically significant when $p < 0.05$.

3 RESULTS

3.1 *Subjective workload estimation*

The SWI scores, as presented in Table 1 showed a very high strain (4.4 ± 0.8) with a significant positive correlation between the SWI scores for fatigue, risks and concentration ($P < 0.05$). The correlation between the SWI score and work satisfaction (interest in the job and autonomy) were significantly negative ($P < 0.05$). The complaints of body pain regions in percentage as shown in Figure 3 were found as follows: 33.33% (low back / waist), 24.24% (legs /knees / feet), 18.18% (neck / shoulder / upper back), 15.15% (arms / hands / fingers) and 9.09% (other).

3.2 *Electromyography*

3.2.1 *Muscle workloads*

Muscle load is expressed in the percentage of the measured value in between the absolute rest value (zero) and the maximal capacity (MVC) of the selected muscles. The MVC was compared to the values before and after the effort. The MVC of each muscle was presented in Figure 4 and the occupational risk (percentage distribution of the selected muscles) during patient lifting activities was shown in Figure 5.

Table 1. Subjective workload index (Means ± SD).

Factors	Scores (0–10)
Fatigue[a]	8.2 ± 1.4
Risks[a]	8.2 ± 1.3
Concentration[a]	8.1 ± 1.5
Complexity	6.9 ± 1.7
Rhythm	6.7 ± 1.5
Responsibility	9.3 ± 1.0
Interest in the job[b]	6.1 ± 2.4
Autonomy[b]	6.0 ± 1.6
Subjective workload index	4.4 ± 0.8

[a] A significant positive correlation with the SWI ($P < 0.05$).
[b] A significant negative correlation with the SWI ($P < 0.05$).

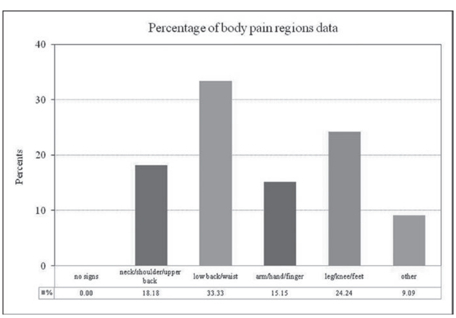


Figure 3. Percentage of complaints body pain regions.

3.2.2 *Changes of maximal voluntary contraction*

A comparison between before and immediately after lifting a patient in the right Erector spinae,

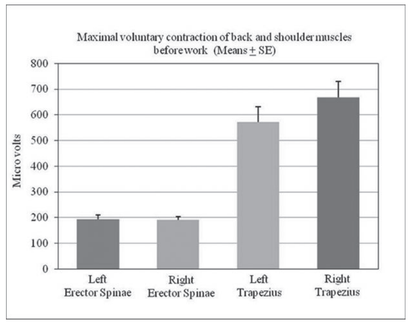


Figure 4. The maximal voluntary contraction in micro-volts of each muscle before work (Means ± SE).

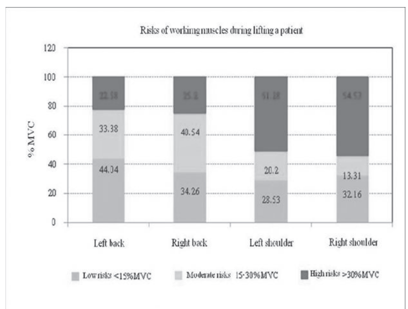


Figure 5. Risks of muscle workloads in percentage.

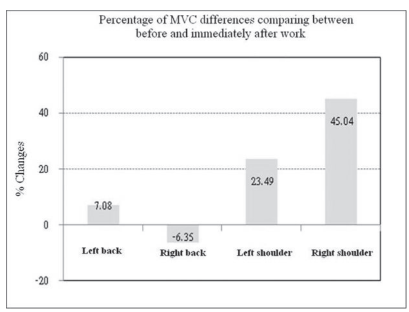


Figure 6. A comparison of MVC between before and at the end of work shown in percentage.

MVC showed a loss of 6.35%, whereas the left Erector spinae MVC slightly increased (7.08%). No loss was observed in the MVC of both the right and left Trapezius (+ 45.04% and + 23.49%, respectively), as represented in Figure 6.

4 DISCUSSIONS

4.1 *Subjective workload estimation*

The SWI is a useful tool to obtain qualitative data on the initiating feelings of discomfort. It is an onset indication of a LBP problem in exposed nurses. The SWI results showed high scores for fatigue, perceived risks and concentration were high. The compensatory factors as interest in the job and the degree of autonomy in the jobs were rather low. The lack of these motivating factors seemed to result in a mental strain.

The total SWI obtained very high (4.4 ± 0.8) and required improvement measures on short term.

4.2 *Muscle workload*

The EMG results obtained during patient lifting work of the participants showed that there was much higher loss of the main muscle power (M. Erector Spinae) in the right side than that of the left side. This findings were caused from awkward position of the participants when lifting a patient or turning position of a laying patient on bed.

In general, warning signs for biomechanical risks when the decline of MVC was more than 15%, the right Trapezius workload was closed to their limit point. As the results in Figure 5 shows, >50% in the Trapezius muscles is an almost non-acceptable muscle load, and requires improvement measures.

The loss in muscle power, as found in Figure 6 indicates a high degree of fatigue for the relative short period of observation. This confirms the finding of Hui et al. (2001) studied low back neuromuscular fatigue patterns following a daytime work shift of 21 female nurses working in geriatric wards and found fatigued muscles after work corresponding with a significant drop in initial Median Frequency (MF) of EMG and increased in negative slope of the MF/time plot of the back muscles after work suggesting muscles fatigue.

4.3 *Recommendations of prevention*

MusculoSkeletal Disorders (MSD) and low-back pain is multi-factorial in origin and reflect an unbalance between workload and capacity. Therefore, strategies to reduce musculoskeletal disorders have to include an ergonomics approach, which attempt to integrate equipment, tasks, personnel, and the work environment in order to create a well balanced job. Wongthanakit (2005) showed that 76 percent of the nurses had a moderate level of low-back pain preventive behavior.

The main possibilities can be considered:

a. Technical solutions: adjusting the workplace technically: height adjustable beds and/or environmental space so that nurses can move freely in the rooms. Technical aids such as sliding cloths, patient lifts, etc. are techniques which can be considered as medium to long term.
b. Human operator: increasing the individual capacity: physical, operational capacities, skills. A program of regular exercise for nurses, with an emphasis of ergonomics can reduce musculoskeletal symptoms (Alexandre et al. 2001). Successful intervention programs require a major commitment to providing help for individual nurses rather than providing "training" in groups (Harber 1990). This principle also takes time and could only solve problems temporarily as fatigue will occur in a later phase. Among predisposing factors, it was found that nurses with more practical skills (lifting techniques), are more accessibility to useful techniques which should be promoted.
c. Organization resources & techniques: job rotation, alternating muscle load following the position of the nurse around the bed, exposure time and recovery periods i.e. activities during which other muscles are solicited in a different way (dynamic contractions) will improve blood supply reducing acidosis.
d. Increasing knowledge and awareness about low back pain, health information with positive attitudes toward its symptoms and preventive behavior tended to have a higher level of low back pain preventive behavior.

5 CONCLUSIONS

It was concluded that the total muscle strain in nurses belongs to highly risky jobs even when lifting with a team of 4 nurses and the objective results the subjective experienced SWI scores. The imbalance in the M. Erector Spinae is a worrying issue.

Measures on short and medium term refer to the organization, individual awareness and actions as to technical improvements involving all personnel, management and authorities.

ACKNOWLEDGEMENTS

The authors are very grateful to the Social Security Office, the Ministry of Labour, Thailand for their financial support in carrying out the research project. The authors would like to thank all the participants, Rangsit University colleages and Dr. Suwannee Laoopugsin for their contribution in the project, and Miss Techika Tantimas for her advice on English of the manuscript.

REFERENCES

Alexandre, N.M.C., Moraes, M.A.A., Filho, H.R.C. & Jorge, S.A. 2001. Evaluation of a program to reduce back pain in nursing personnel. *Rev Saude Publica* 35(4): 356–361.

Chavalitnitikul, U., Roongreangthum, N. & Ibrahim, F. 1996. An investigation of back pain in the Thai nursing profession. *Journal of Safety and Environment Review* 6: 57–66.

Harber, P. 1990. Occupational back pain of nurses. *West J Med* 152(2): 175–176.

Hignett, S. 2008. Work-related back pain in nurses. *Journal of Advanced Nursing* 23(6): 1238–1246.

Hui, L., Ng, G.Y.F., Yeung, S.S.M. & Hui-Chan, C.W.Y. 2001. Evaluation of physiological work demands and low back neuromuscular fatigue on nurses working in geriatric wards. *Applied Ergonomics* 32(5): 479–483.

Maul, I., Laubli, T., Klipstein, A. & Krueger, H. 2003. Course of low back pain among nurses: a longitudinal study across eight years. *Occup Environ Med* 60: 497–503.

Taptagaporn, S. 1999. Occupational back pain. *Fact Sheet - Occupational Health* 2(9): 1–5.

Wongthanakit, S. 2005. Factors related to low back pain preventive behavior among nursing personnel in Governmental Hospitals, Nonthaburi province. *Journal of Public Health* (Thai) 35(2): 109–118.

Yoopat, P. 2004. Ergonomics in practice: Reducing physiological strain in construction work in Thailand - An ergonomic method. *Asian Pacific Newsletter* 11(1): 15–18.

Yoopat, P., Deerod, P., Rujipong, C. & Vanwonterghem, K. 2006. Assessment of muscular disorder in garment industry in Thailand. *Proceedings of International Ergonomics Association Conference: Meeting Diversity in Ergonomics.* Maastricht, the Netherlands, 10–14 July 2006.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Biomechanical risks in Thai construction workers

Pongjan Yoopat
Ergonomics Division, Faculty of Science, Rangsit University, Patumthani, Thailand

Kamiel Vanwonterghem
CERGO-MENSANA, Sijsele-Damme, Belgium

ABSTRACT: MusculoSkeletal Disorders (MSD) affect the quality of life of employees worldwide due to overexertion of the human movement system caused by physical working conditions. The objectives of this study was to assess the impact of workload on the biomechanical risks of the human body in construction jobs that can be partly realized by objective criteria—body postures and movements through muscle strain (electromyography) and partly by paying attention to the subjective experiences of the exposed workers. In total 50 experienced construction workers, equally spread over both genders participated voluntarily in the project. Through EMG measurements which showed a total muscle load exceeding in total about 20% above the risk threshold set at 15% MVC and the subjective ratings of SWI >3.0 specific risks were recognized for which preventive advice was formulated.

Keywords: musculoskeletal disorders, biomechanical risks, muscle strain, electromyography, experienced complaints

1 INTRODUCTION

MusculoSkeletal Disorders (MSDs) have a worldwide effect on the quality of life of employees. Neck and upper limb musculoskeletal disorder (WRULDs) are a significant problem within the EU member states. The cost of WRULDs problems in the European Nordic countries was estimated at 0.5–2.0% of gross national products (Toomingas, 1998).

In Great Britain, The Health and Safety Executive (HSE) estimated the cost related to WRULDs was 1.25 billion pounds per year (Davies & Teasdale 1999). In the United States, MSDs represent 40% of compensated injuries and cost between $45 to 54 billion per year (Institute of Medicine USA, 2001).

In Thailand, a report related to musculoskeletal disorders by the Ministry of Public Health stated 303 (cases per thousand) in the year 1999, 360 in 2000, and 723 in 2002 (Ministry of Public Health Thailand, 2004).

However, these asymptotic estimates have to be reconsidered in the future because the actual standard protocol to diagnose MSD cases is still in the ratification process.

In 2006, construction work in Bangkok employed about 4.7% of the total population and had an important impact on the incidence rates in the national statistics (National Statistic Thailand, 2010).

They also include accident claims classified per accident type during the period 2006–2008 (WorksafeBC, 2010). Of these, falls accounted for 25%, cases of overexertion were about 23%, impacta and collisions for 22% and others e.g. repetitive motions in the upper limbs, whole body postures, etc. counted for a total of about 30%.

Although risks were recognized, it has to be stated that—as statistics prove—the strategy of WRULD-prevention was not very successful because it was based on post-factum analysis whereas, an anticipatory approach could be more profitable.

The objective of this research project was to assess the problems related to the human movement system (musculoskeletal activity) by means of objective measurement of muscle strain in combination with subjective perceived WRULD problems in male and female construction workers.

2 MATERIALS AND METHODS

2.1 *Subjects*

Fifty subjects equally distributed over both genders, and all experienced construction workers were selected out of a total screening group of 175 subjects. The selection criterion was a SWI (Subjective Workload Index, see 2.2.1) strain level of between 2.5 and 5.0. All subjects participated voluntarily in the project and signed consent of agreement for the proposed methodology.

The average age was 34 ± 3 and 35 ± 3 years with a professional experience of 7.5 ± 5 and 6 ± 4 years for male and female workers respectively. The average Body Mass Index (BMI) was 21.3 ± 2 for male and 24.9 ± 6 for female workers.

The hand grip strength for male operators was 40.7 ± 5.9 kg and 28 ± 5.6 kg for female. Most of the employees were right handed and their heart rate and blood pressure at rest were within the range of normal healthy people.

2.2 *Methods*

2.2.1 *Subjective Workload Index (SWI)*

The SWI is a validated questionnaire developed to detect work related problems experienced by employees exposed to an occupational load (Vanwonterghem, 1986).

The general analysis covers the self-rated work factors. Six items represent load aspects (LF): problems about fatigue, concentration, risks, task complexity, work rhythm and problems with responsibility. Two factors (CF) have a compensating function, namely interest in the job and the degree of autonomy.

The problems experienced were then quoted on an 11-point scale (0 = no problems at all to 10 = extremely annoying, painful) and the compensating factors vary from 0 = no interest or no freedom at all to 10 = very high interest in the job and full autonomy.

The SWI is then calculated as follows:

$$SWI = [\,\Sigma(LF) - \Sigma\,(CF)]/8$$

SWI results of <2 refer to small or no problems and the higher the SWI, the more critical the workload. From the level of SWI >2.5 a detailed analysis is advised which should give indication for medium term (months) improvement. SWI >3 to 4 calls for interventions on short term i.e. within 1 month, >4 to 5 within one week and at SWI >5 work should be interrupted if no measures are taken immediately.

The detailed analysis includes information of tasks (types and duration) with specific indications of the body (posture and movements), the environment (climate, noise, vibrations, lighting, air quality, etc.), organization and references to the sensorial system.

2.2.2 *Muscle strain*

Muscle contractions (intensity in µV) characterizing the postural and movement strain of the body, are measured by means of surface electromyography (SEMG, equipment MEGA-8000P). The intensity (in micro-Volts) was recorded during work and stored in a solid state memory. A PC processed the absolute EMG values by means of the MEGAWIN software which transferred the raw signals into relative data (% distribution between the rest value and the MVC (Maximal Voluntary Contraction). The advantage of working with relative data is that the results attenuate the inter- and intra-individual differences. The relative results are then compared to strain thresholds.

The software proposes risk-threshold ranges from >0 – <15% MVC, >15% – <30% MVC and >30% MVC representing respectively the MSD risk of low, moderate and high. In this study the 15% MVC threshold was used as this level serves as a pre-indicator for problematic muscle strain.

The maximal capacity of each individual (MVC: Maximal Voluntary Contraction) was determined before the start of the work.

2.2.3 *Activities and operations*

An observer recorded the type and the duration of the different work phases. Those appearing in an irregular sequence were manually totalized per operation in the final processing.

The totaled duration of operations is expressed in a percentage to the total time (% time) for male and female operators as presented in Table 1.

Informal rest: included resting, waiting for materials to finishing previous tasks. Preparations: included, materials, products, tools, surface-cleaning as shown in Figure 1 concerning cleaning the work area after construction is finished. The physical work is characterized by adverse postures, upper limb movements and hand grips and lifting materials or debris of heavy weight when collected in baskets or containers.

Other operations were:

- Walking with or without supplying materials;
- Equipment: hand driven tools and/or mechanical equipment (drill machine, pneumatic hammer).
- Specific construction concerned bricklaying, carpentering, concrete mixing, metal work.
- Material handling (lifting, pushing, pulling, and piling materials). Figure 2 shows a typical handling of 50 kg cement bags, performed by two operators.

Table 1. Operations duration in % of total time.

Operations	Male	Female
Informal rest	24.63	17.25
Preparations	16.87	10.78
Walking/transport	8.18	9.97
Equipment	6.87	8.76
Specific construction	32.32	34.10
Material handling	11.13	19.14

Figure 1. Preparing work surface.

Figure 2. Material handling.

On average the periods of measurement took 75 ± 12 min and 82 ± 24 min for male and female operators respectively.

2.2.4 *Statistical analysis*

Data of SWI and EMG were analyzed by descriptive statistics: statistical mean, standard error and percentages. The statistical processing of the raw EMG signals and the threshold distributions were realized by means of the MEGAWIN software but tasks and duration were processed manually by the observers.

3 RESULTS

3.1 *Subjective Workload Index*

The SWI scores were respectively 3.1 ± 1.1 and 3.0 ± 1.0 for male and female workers which reach the level of the requirement of a detailed analysis and 'proposals for improvement' within one month.

For male workers 'concentration' (8.2 ± 1.7) and 'responsibility' (8.1 ± 1.9) have the highest score and also 'responsibility' (7.9 ± 2.3) was important for the female workers.

'Risks' were important for both genders (6.9 ± 2.1 and 6.7 ± 1.8) and fatigue scored somewhat higher for female (6.8 ± 1.6) than for male workers (6.6 ± 1.4). 'Complexity' and 'work rhythm' seemed not to be a problem for female operators whereas 'work rhythm' (6.6 ± 1.6) gave a moderate pressure to the male workers. Male workers scored higher for all factors (with exception of fatigue) and the small difference in the overall SWI (± 1.0) can be explained by the fact that male operators have higher job satisfaction and more freedom (autonomy) than the female workers.

The main complaints on medium term (months) for male subjects concerned problems/pain of lower back (26%), neck and shoulder (22%) and upper extremities with negligible problems in the other parts of the body. For females low back pain scored higher (35%) and an almost identical load for the upper extremities, neck and upper back

Concerning the environmental factors: two major complaints were mentioned, namely heat and the presence of dust. The main complaints for medium term.

3.2 *Muscle load*

The 8 muscle groups—which were selected from the detailed SWI analysis—concern the shoulder muscles (M. Trapezius, left and right, TL and TR), because they give the essential stability for the arms, the M. Deltoids (LD, RD), and the forearm muscles: the left and right Wrist Flexors and Extensors (LF, RF, LE, RE).

The MVC values are normal for male and female construction workers (Table 2) and are quite consistent for the individual averages.

The small differences between left and right, especially for the forearms are linked to the left/right handed preference of working.

As an objective value for the total muscle load per muscle group it has been decided to use the sum of all percentages which exceed the 15% MVC threshold of all participants, and the sum

Table 2. MVC male and female operators.

Muscles	Male (N = 25) MVC (μV)	Female (N = 23) MVC (μV)
M. Left Trapezius [LT]	610 ± 44	458 ± 41*
M. Right Trapezius [RT]	610 ± 55	510 ± 49
M. Left Deltoid [LD]	589 ± 51	500 ± 56
M. Right Deltoid [RD]	619 ± 54	511 ± 57
M. Left Wrist ext [LE]	491 ± 40	287 ± 23*
M. Right Wrist ext [RE]	607 ± 70	341 ± 27*
M. Left Wrist Flex [LF]	545 ± 61	300 ± 21*
M. Right Wrist Flex[RF]	553 ± 53	289 ± 19*

* significant different from male at $p < 0.05$.

is then expressed as a percentage to the total sum >15% MVC. Figure 3 represents the total load per muscle group. From this we can observe that the forearm muscles perform the highest contraction intensity, whereas the shoulders have lower values.

Obviously the shoulder region offers a more static supporting stability to the arms which have to handle weights and the specific manual handlings in all types of operation. Especially in the female forearms there is an important difference with the male operators, which suggest a higher WRULD risk.

The differences in the Standard Error (SE) refer also to dynamic muscle activities in which high and low contractions alternate and this can be explained by the fact that the nature of work is different.

Whereas females perform rather non skilled operations (preparations, in this case also demolishing with pneumatic hammer, concrete preparation, etc.) the male operators are more concerned with skilled activities (brick laying, iron, carpentering) and heavy physical work e.g. handling cement bags of 50 kg. These solicit more static contractions with only small fluctuations in dynamic work for hands (e.g. brick laying, iron binding).

The following figure 5, shows the muscle efforts in the different tasks for male and female workers.

As it can be observed from Figure 5, only walking (with transporting some materials) male operators score higher than female but in all other activities the female workers have the most important muscle strain.

The most intensive work was drilling with a pneumatic hammer when demolishing. Besides the weight of the pneumatic tool this also gives strong vibrations in the forearms.The low shoulder muscle load in male operator refers to static postural contractions.

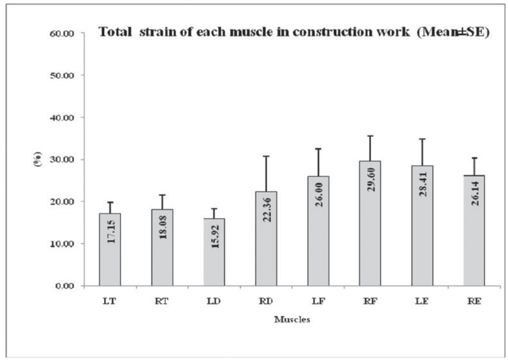


Figure 3. Total strain (Σ >15% MVC) for all subjects.

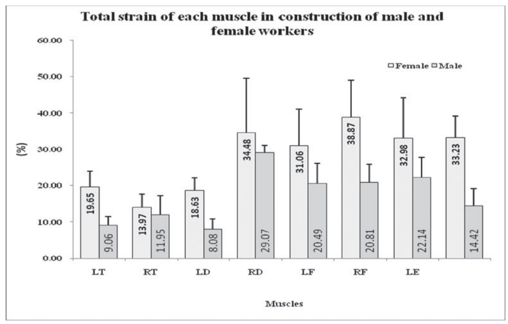


Figure 4. Total strain (Σ >15% MVC) per gender.

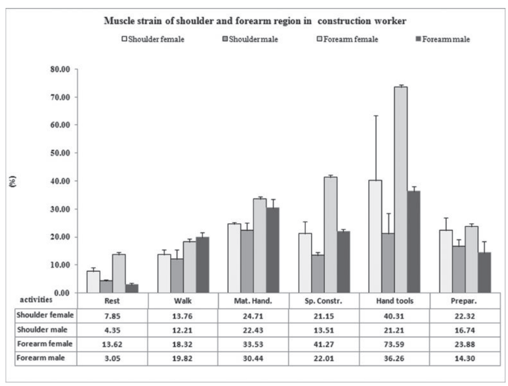


Figure 5. Muscle strain (Σ >15% MVC) per operation and per gender.

4 DISCUSSION-CONCLUSIONS

Prevention of WRULD in the construction industry should be oriented to the assessment of muscle activity as manual physical tasks are dominantly present in the operators' activities.

Postural load as static component for back and shoulder muscles must be managed in combination with problems of manual work with a dynamic character. This can be obtained by using the electromyography technique to assess biomechanical damage risks.

Handling of materials with different sizes, shapes, weights etc. has an energetic and a mechanical aspect whether the lifting operation is done by a machine or by a human. In both cases the operator needs the essential dynamic movements and, the 'capacity' (power) to realize the task. For humans this power is present in the biomechanical structure with its muscle potentials. For a certain lifting task, the external mechanical power for moving bricks (e.g. brick of 1.5 kg) is the same for each operator: 1.5 kg is 1.5 kg. Depending on the individual capacity the 'load' aspect changes as the higher the capacity the easier work can be performed, and subsequently represents a lower MSD risk.

In order to evaluate the 'load' or 'risks' the relative value (the load/capacity ratio) will indicate which problems should be handled with priority and, as the body will 'feel' the shortness in capacity, the subjective complaints have to be taken into account as well. In this project the detailed SWI indicated problems in the shoulder-arm region and the detailed analysis (EMG) confirmed the high WRULD-risks in the forearms of the female operators.

Besides technical proposals such as adjusting working heights, choice of lower weights for materials, mechanical aids etc. special attention was paid to organizational items such as work-rest schemes, task allocation (the 'demolishing' should not belong to females' job). A study of product/materials flow over the working site could avoid a series of material handling and transport tasks.

For prevention purposes, the principle of anticipating the risks should be intensified and implemented in a practical strategy.

This strategy could include a 'risk' screening based on subjective complaints, followed by a detailed analysis by measuring physiological reactions and threshold comparison.

ACKNOWLEDGEMENTS

This study is supported by Ministry of Labor, Department of Social Security Office. The researcher thanks to subjects, and all cooperative colleages and advisors, and my dearest Mom. The authors would like to thank Mr. Graham Harvey for his English suggestions of the manuscript.

REFERENCES

Davies, N. & Teasdale, P. 1999. The cost to the British economy of work accidents and work-related ill-health. Sheffield: Report, *Health and Safety Executive.*

MPH 2004. *Report of Musculoskeletal disorder*. Ministry of Public Health, Thailand.

NRCIM 2001. Musculoskeletal disorders and the workplaces: Low back and upper extremities. National Research Council and Institute of Medicine (NRCIM). Retrieved from *http://books.nap.edu/catalog/10032.html*

Toomingas, A. 1986. Methods for evaluating work: Work-related musculoskeletal neck and upper extremities disorders in epidemiological studies. National Institute for Working life. Göteborg University. *Arbete och hälsa vetenskaplig skriftserie* 6: 1–20.

Vanwonterghem, K. 1986. A postural load evaluation technique: Based on energy expenditure, heart rate and subjective appreciation. In N. Corlett, J. Wilson, I. Manenica (eds) *The ergonomics of working posture; Proceedings of the First International Occupaitional Ergonomics Symposium*, Zadar, 15–17 April 1985. London: Taylor & Francis.

Work Safe BC. 2010 Statistic for construct. Retrieved from: *http://www2.worksafebc.com/Portals?construction?statistics.asp*

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Correlation between MSD among milking parlour operatives and specific work place design

Martina Jakob
Leibniz-Institute for Agricultural Engineering Potsdam-Bornim e.V., Germany

ABSTRACT: The constructive work place dimensions in combination with anthropometrics and bovimetrics have a severe impact on comfort and efficiency. A significant rate of work absenteeism among milking parlor operatives is the background for an investigation on the impact of the work place design regarding the work load. As a result the optimal working height was defined. In practice the working height varies from cow to cow and cannot be adjusted. Working above shoulder level increases the incidence of MSD. Therefore on farm measurements were carried out to quantify the individual risk for the milking staff depending on the average distance between the udder and the floor within the herd. Furthermore the correlation between working above shoulder level and the incidence of MSD is checked.

Keywords: MSD, work place design, milking parlor, work absenteeism

1 BACKGROUND

1.1 *MSD among dairy farmers*

The incidence of symptoms in the neck, shoulder and upper extremities among European workers on dairy farms and especially among parlour operatives has increased with mechanisation. Although in general the overall workload is reduced in parlour milking in comparison to tie stall milking, the health status among the workers was not improved. Different studies from Sweden, Finland and Germany report that. For example every third Finnish milk producer working in a loose housing barn is affected by symptoms in the neck-shoulder region (Tuure & Kartunen 2007). A longitudinal study by Pinzke (2003) reported an increase of MSD among 1465 dairy farmers in Sweden asked in 1988 and again in 2002. There is an incidence that symptoms increase with the number of cows and parlour size (Tuure & Alasuutari 2009). Structural changes in the dairy industry go along with increased productivity allowing less muscular recovery (Stal et al. 2003). Pinzke et al. (2001) showed that the most physically demanding milking tasks were cleaning the udder, premilking and attaching the unit.

1.2 *Work place design*

The general construction of milking parlours is very similar all over the world. When cows are held in loose housing barns they are brought to parlours for milking. Parlour constructions slightly differ, but the worker normally stands in a pit between 80 and 95 cm below the level of the cows to avoid awkward working postures. The animals are positioned in different ways, depending on the parlour type. The most common parlour types are herringbone, side-by-side, tandem and the rotary (see Fig. 1). In the tandem system the cows stand in a row, and each "box" is entered on demand. The milker has full sight of the cow; the cow stands parallel to the milking pit. Tandem parlours are usually found on smaller farms because they are rather spacious. In the herringbone system the cows are overlapping and positioned in an angle of 33° or 45° to the edge of the pit. This is the case in a rotary parlours, too. In the side-by-side system the milker is facing the back of the cow attaching the milking

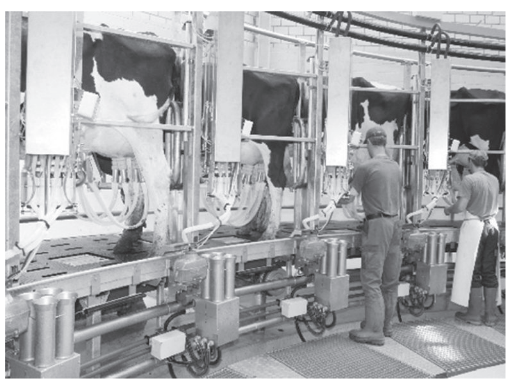


Figure 1. Rotary parlour equipped with MultiLactor.

unit through the back legs. Some rotary systems also have side-by-side positioning. Regarding the work place design the different parlour types do vary in the horizontal distance between udder and the edge of the pit, which influences the necessary reach to attach the milking unit. The vertical distance varies with udder height and body height of the worker. A total range of app. 30 cm for the udder and in addition to that another 30 cm for the worker (5th female percentile—95th male percentile) is possible. This should well picture the possible diversity within the work place design in milking parlours.

The height of the udder significantly influences the body posture. Working 15 cm or more above shoulder level was proven to increase the upper arm elevation and the muscular load of the corresponding muscle groups whereas when working below shoulder level upper body inclination increases. Smaller persons are more affected because they also have shorter arms.

1.3 *Ergonomic interventions*

Based on preliminary studies the workload can be reduced, if the teat ends are at the same level as the shoulder of the worker. The use of light milking units (1,5 kg or less) also significantly reduces the workload. A quarter individual milking device (MultiLactor® by Impulsa AG, Germany) improved the worker comfort in comparison to a conventional light milking unit (see fig. 1) (Jakob & Liebers 2010). Adjustable floors can be used to reduce the anthropometric influence and to adjust the working height. Service arms help to reduce the static load while holding the cluster.

2 METHODS

2.1 *On farm measurements*

Data acquisition is still going on. A number of three farms, distributed all over Germany, is selected for this presentation. Bovimetrics were assessed as well as the shoulder height of each worker. Out of this data a personal work load profile is calculated.

Based on the results of the preliminary study (Jakob & Liebers 2009) the optimal working height is given, when shoulder and teat ends are on the same level.

Farm data assessment was done by paper pencil method using a folding rule. There are three categories of data, the farm specifics, the worker specifics and the cow specifics. The last was measured during a milking period and for the complete herd.

Data was evaluated with MS Excel and JMP statistical software.

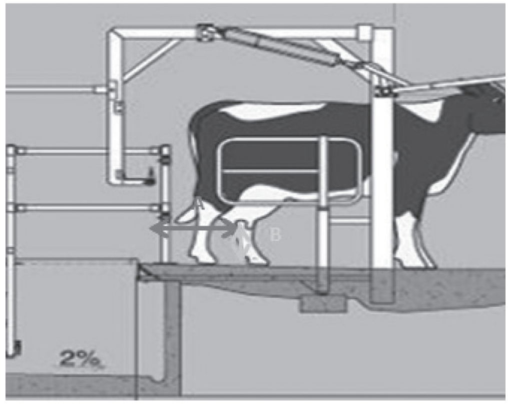


Figure 2. Measures taken on farm measurements.

The animal specific measures taken were the distance between udder and floor (B), the horizontal distance between the middle of the udder and the edge of the pit (A) and the diagonal distance between two teats (C). For a better understanding A and B are displayed in Figure 2 shows a side-by-side parlour design.

2.2 *Assessing the incidence of MSD using the standardized Nordic questionnaire*

The standardized Nordic questionnaire (Kuorinka et al. 1987) was handed out to every milker on each farm to analyze the incidence of musculoskeletal symptoms. The workers filled in the questionnaire themselves, because not all workers were present at the time of the visit.

If a worker reports symptoms, the locations of the symptoms are compared with the specific work load profile on the farm. The correlation of symptoms in the upper extremities with working above shoulder level is checked once a larger number of farms is included in the investigation.

3 RESULTS

3.1 *On farm data*

The results presented here are preliminary results including three farms and ten milking parlour operatives. To be able to statistically validate a correlation between symptoms in the upper extremities and the workload profile with excessive work above shoulder level more farms will be included in this survey.

The three presented farms had implemented a 33° herringbone parlour. The herd size ranged from 44 to 200 cows. Farm 2 has an adjustable floor.

Table 1. Average values for the cow specific on farm measures.

	A (in cm)	B (in cm)	C (in cm)
Farm 1	38,5	59	17,7
Farm 2	40,5	55,6	18,5
Farm 3	45,4	57,8	16,5

Table 2. Parlour and worker specific data.

	Pit depth (in cm)	Shoulder height (in cm)
Farm 1	89	134–149
Farm 2	94	148–156
Farm 3	97	143

The results for the defined values A and B, explained in figure 2 and C are shown in Table 1. C is the diagonal distance between two teats. It displays the size of the udder. Depending on the parlour type ½ C has to be added to A to know the furthest point of the cow's udder when attaching the teat cups. The average distance between udder and floor was 57,7 cm for the three farms and a group of 321 cows. The main breed was Holstein Frisian.

3.2 *Incidence of MSD symptoms*

Out of the group of the ten milkers two of them had to get medical treatments because of arm and shoulder problems. Both are female, one of them of average size (143 cm shoulder height) the other one rather small (134 cm shoulder height). The workload profile for her is presented in Figure 3.

3.3 *Personal work load profiles*

To get an impression of the work load the shoulder height of the worker is subtracted by the depth of the milking pit. The ideal height is given when the result equals the distance between udder and floor. For the calculation of the profile all values including ±5 cm of the distance between udder and floor that equal the described margin were defined to be within the optimal range. The rest is rated either above or below shoulder level.

Due to the range within the distances between udder and bottom of the floor, there are no work load profiles that are completely within the optimal range.

Figure 3 shows the percentage of the working height distribution of one person. The red area marks the percentage of work above shoulder level. If the depth of the pit was adjusted according to the body height of this worker presented in

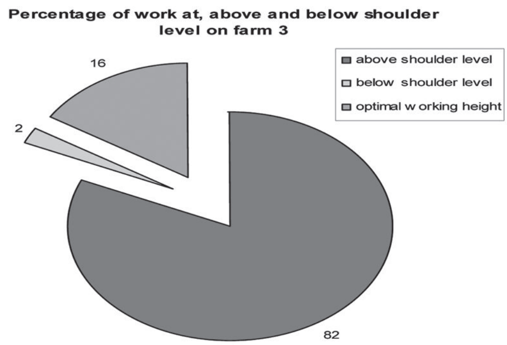


Figure 3. Example of workload profile.

Table 3. Examples of workload profiles for six workers on two farms.

	B (cm)	Pit depth	Set	% opt.	% bsl	% asl
A	58	97	47	16	2	82
B	59	89	58	61	13	26
C	59	89	60	68	19	13
D	59	89	49	16	2	82
E	59	89	52	20	7	73
F	59	89	45	12	0	88

bsl = below shoulder level, asl = above shoulder level.

Figure 3, an optimal working height during 55% of the milking time could be achieved.

Table 3 shows six examples of specific workload profiles of two farms. Only two workers do find the theoretically ideal working height, when the distance between udder and floor (B) equals the set (individual shoulder height—pit depth). Worker A and F were the ones that needed medical treatment. The profiles also clearly show that the variability of the animals cannot be compensated by an adjustable floor. There always remains a group of animals that induce the necessity of working above or below shoulder level. The percentage of inevitable work beyond the optimum ranges from 32–45%.

Another factor influencing the body posture is the arm length. It ranges from 60–79 cm, resulting in an effective length between 37 and 44 cm. The average horizontal distances ranged from 39 to 45 cm (see A in table 1). Therefore in many cases the full arm length is needed to place the cluster underneath the udder. When the worker has to work above shoulder level the distance between shoulder and udder becomes larger. This is the same when working below shoulder level. The horizontal distance between the middle of the udder and the edge of the pit also represents the length of the level arm to calculate the force the milking unit is putting on the operator.

4 DISCUSSION AND CONCLUSIONS

Much effort is spent on improving cow comfort, but the development for improving the work place milking parlour has come up with little changes once it was invented. Nevertheless task repetitiveness and the duration on one job as well as specialization are continuously increasing the work load. The milking operation is still reliant on human input, but it is more and more driven by greater efficiency and throughput due to the structural changes. The number of cows that is milked in one hour was drastically increased. The throughput also depends very much on the parlour type and size.

The precise analysis of the work place design presented here shows that depending on their body size some workers may barely find an optimal working height when they are milking the cows. Constructive limitations and the varying parameters such as body height and bovimetrics make it impossible to provide optimal working heights at all times for the common parlour constructions. These facts can result in physical overload.

For smaller persons the percentage of work above shoulder level seems to be a major risk factor. Regarding the presented results two out of ten workers required medical treatment due to symptoms in the upper extremities. The work load profile of one of them is shown in Figure 3 clearly indicating a need for improvement. A possible load reduction could be achieved, if the depth of the pit floor was lowered on this farm, but a remainder of 45% of the animals out of the optimal range clearly indicates the necessity of additional measures, for example lighter milking units.

The investigation is still going on to finally statistically validate a threshold (percentage of work above shoulder level) when workers are more likely to develop muscular skeletal disorders.

REFERENCES

Jakob, M. & Liebers, F. 2009. *Body posture variation during machine milking regarding weight of milking unit and working height—Experimental study*. BAB Heft 66; ISSN 0947-7314.

Jakob, M. & Liebers, F. 2010. Milking parlour ergonomics. Landtechnik, 1, 24.26. *https://www.landtechnik-online.eu/en/archive/2010/heft-12010/building-and-planning/milking-parlour-ergonomics/*

Kuorinka, I., Jonsson, B. & Kilborn, A. 1987. Standardized Nordic questionnaire for the analysis of musculoskeletal symptoms. *Applied Ergonomics* 18: 233–237.

Pinzke, S.M. & Hansson, G.-A. 2001. Physical workload on upper extremities in various operations during machine milking. *Ann Agric Environ Med* 8: 63–70.

Pinzke, S. 2003. Changes in working conditions and health among dairy farmers in southern Sweden—A 14-year follow-up. *Ann Agric Environ Med* 10: 185–195.

Stal, M., Pinzke, S., Hansson, G.-A. & Kolstrup, C. 2003. Highly repetitive work operations in a modern milking system: A case study of wrist positions and movements in a rotary system. *Ann Agric Environ Med* 10: 67–72.

Tuure, V.-M. & Karttunen, J. 2007. Työntekijän kannalta toimiva pihatto. In: T. Kivinen etc. (Eds). Lypsykarjapihaton toiminnalliset mitoitusvaihtoehdot, *MTT:n selvityksiä* 137: 1–149.

Tuure, V.-M. & Alasuutari, S. 2009. Reducing work load in neck-shoulder region in parlor milking. BAB Heft 66, ISSN 0947-7314.

Part V: Bio-signals & ergonomic assessment

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Identification of improper state of highway driving by using EEG

Sang-Young Yoon & Jung-Yong Kim
Hanyang University, Ansan, Republic of Korea

ABSTRACT: The psycho-physiological responses indicate the psychological state of human based upon mental workload that could be a factor increasing the risk of accident. However, the relationship between the signals and the risks of accident has not been clearly established. Therefore, the characteristics of the electroencephalogram investigated to identify the mental workload patterns. A field experiment was conducted with a test vehicle with data collecting equipment during actual highway driving. Participants drove the vehicle under various driving conditions with high, moderate and low mental workloads defined by road geometry and accident records. EEG was used and parameterized to describe the psycho-physiological state of individual drivers. Parameters were quantified into descriptive statistics and plotted to find the stable state of driver. By parameterizing the characteristics of the EEG response, this study has provided quantitative information on the EEG signals of drivers experiencing various mental workloads.

Keywords: mental workload, driving, electroencephalogram

1 INTRODUCTION

Previously, psycho-physiologists have attempted to find the pattern of a driver's physiological signals during driving. Moreover, they found that the factors such as sleepiness, mental workload, and prolonged driving time had an effect on the driver's physiological signals, especially electroencephalogram (EEG). Torsvall and Åkerstedt (1987) and Kecklund and Åkersted (1993) found that the sleepiness made the spectral density of α band increase. Brookius and Waard (1993) used the $(\theta + \alpha)/\beta$ parameter and demonstrated the possibility of using a driver's EEG response to understand their behavior when driving. Sterman and Mann (1995) found that cognitive processing (in-flight tracking test) caused the suppression of α band magnitude. Kim et al. (2002, 2003, 2004) found that the β/α parameter increased along with mental workload increased.

Previous studies aimed to identify the relationship between the descriptive statistics of EEG data and the driving behavior. However, the size of the individual variability of the signal made it difficult to have a reliable result. Thus, practical application to the drivers is yet to be further investigated.

Therefore, this study investigated how the EEG signals of driver respond during actual highway driving in terms of quantitative parameters. Then, the distribution of individual parameter was examined and proper range of parameter was defined. We defined the psycho-physiological state out of normative range as improper state in this study. Lastly, the possibility of practical application of the distribution and data was discussed.

2 METHOD

2.1 *Participants and apparatus*

For actual highway driving experiment, seventy volunteers participated in the experiment. Their average age was 25.35-year-old (SD = 3.17) and average driving experience was 4.58 year (SD = 2.47).

A 4 × 4 vehicle with an automatic transmission was equipped with measuring devices that included seven channels of electroencephalogram (EEG).

2.2 *Experimental road conditions*

In this study, the road condition was defined according to the mental workload expected for safe driving. The expected mental workload was termed as the mental demand in this study, and this workload was determined based on the accident record, horizontal alignment and vertical grade of the designated target area which was approximately 200 m in length. We used three levels of mental demand: low, moderate and high. Each demand level was defined in Table 1.

Table 1. Definition of three road condition: high, moderate, low demand.

Demand level	Accidents Per year	Characteristics
High	6	Near an interchange. Slight curve to the right with a downward grade (radius of curvature: 700 m, longitudinal slope: 5%).
High	7	End of a tunnel. Almost straight with a downward grade (radius of curvature: 1500 m, longitudinal slope: 2~5%).
Moderate	None	Almost straight with a slight downward grade (radius of cuvature: 1500 m, longitudinal slope: 5%).
Moderate	2	Curved to the left, with a downward grade (radius of curvature: 630 m, longitudinal slope: 5%).
Low	None	Straight, with almost no vertical alignment (radius of curvature: 3000 m, longitudinal slope: 1%).
Low	None	Straight, with almost no vertical alignment (radius of curvature: 3000 m, longitudinal slope is 1~2%).

2.3 *Quantification and selection of parameters*

The slope parameter was designed to represent the increasing and decreasing trends of the signals in order to consider the temporal nature of the signals. The collected data were quantified in terms of the slope parameter as well as the amplitude parameter.

Signals were collected during experiment. Then the signals were cut by 50-second period and averaged every 10 second. The target area was determined to be between 30 and 40 seconds. In this way, the driver's mental state both 30 seconds prior to entering the target area and 10 seconds after exiting the target area could be monitored and analyzed. Frequency and power spectral analysis were carried out using EEG data to compute the alpha, beta and theta values. Those EEG values collected for 50 seconds were normalized based on individual baseline data collected during resting and before driving to minimize inter-subject variability. The data analysis procedure was summarized in Figure 1.

2.4 *Procedure*

Each participant had driven for 30 minutes to familiarize with test vehicle before attaching electrodes. Then, the baseline data were collected for normalization before experiment. The experiment lasted about 50 minutes on actual highway without traffic control. The experiment was performed only on sunny or partly cloudy weekdays. Participants were instructed to drive safely and not to exceed 100 km/h. The participants were not allowed to talk, listen to music or smoke while driving. A participant took part in a single test on a given day without repetition. The same participant drove the test vehicle under each of the three different mental demands (high, moderate and low) to minimize inter-subject variability. The order of demand for individual participants was randomized to minimize the carry-over effect.

To collect the data, EEG sensors were placed on the following locations using a headcap: F3-F7 and F4-F8 of the frontal lobe, C3-Cz of the central lobe, P3-Pz of the parietal lobe, O1-O2 of the occipital lobe, and T3-T5 and T4-T6 of the temporal lobe. These locations were determined

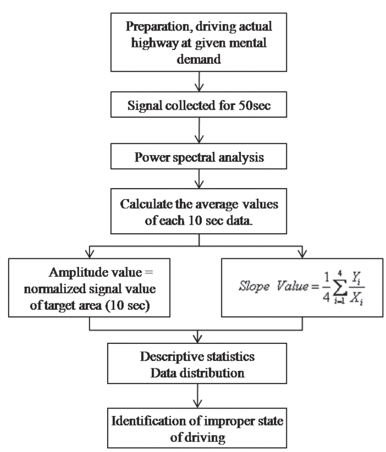


Figure 1. Procedure of data analysis.

based on the 10–20 method of the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (Cooper, 1980). The data sampling rate was 256 Hz, and high-and low-pass frequencies were 0.5 Hz and 30 Hz, respectively.

3 RESULTS

3.1 *Descriptive statistics*

Descriptive statistics of amplitude and slope are displayed in Table 2 and Table 3. All parameters which indicated mental effort did not show any significant difference ($\alpha = 0.1$) between different demand levels in a one-way analysis of variance (ANOVA) because of large variance.

The amplitude of α, θ and (θ + α)/β from the left frontal lobe and θ from the right frontal and parietal lobe increased as the demand level increased, though there were large variance of amplitude (Table 2). The slope of θ and β/α from the left frontal lobe also increased as the demand level increased (Table 3). Other parameters did not show a consistent pattern as demand level varied.

3.2 *The distribution of parameters and improper state of psycho-physiology of drivers*

To identify the unique characteristics and improper state of individual psycho-physiology, the frequency distribution function of the amplitude and slope parameters was investigated. Then, a normality test was used to identify the characteristics of distribution for each parameter.

As a result of normality test, the amplitude parameter of β/α and (θ+α)/β did not have normal distribution except the central lobe. But, other amplitude parameters had normal distribution. For slope parameters, α, β, θ, (θ + α)/β and β/α from five lobes had normal distribution except β/α from parietal lobe. Though 80% of parameters showed a normal distribution, many parameters did not show a clear difference among mental demand levels.

For improper state of driving, the extreme data were identified as improper state of psycho-physiology of drivers. Extreme values were observed at both tails of the normal distribution. The data that was out of 1.64 σ from average was classified into improper state in this study. Extremely high or low EEG data are usually the

Table 2. Average and standard deviation of amplitude parameter.

Lobe and parameter		Demand level Low	Moderate	High
Left frontal	α	22.7 ± 6.9	23.9 ± 6.8	24.2 ± 5.8
	β	19.3 ± 7.4	18.1 ± 7.6	17.8 ± 8.1
	θ	38.8 ± 11.7	39.2 ± 11.6	39.6 ± 12.4
	(θ+α)/β	4.1 ± 3.3	4.3 ± 2.9	4.6 ± 2.7
	β/α	9.5 ± 5.2	8.6 ± 6.6	7.5 ± 4.4
Right frontal	θ	39.3 ± 10.0	40.7 ± 10.3	41.1 ± 9.8
Central	θ	37.4 ± 8.3	37.1 ± 8.5	36.6 ± 8.8
Parietal	α	25.8 ± 7.3	23.8 ± 5.4	23.1 ± 7.1
	β	18.3 ± 7.0	17.7 ± 5.0	17.0 ± 5.8
	θ	33.5 ± 9.4	36.2 ± 12.5	39.2 ± 12.0
Occipital	β/α	10.2 ± 10.7	8.9 ± 3.3	7.9 ± 3.4

Table 3. Average and standard deviation of slope parameter.

Lobe and parameter		Demand level Low	Moderate	High
Left frontal	α	2.4 ± 8.0	–1.0 ± 7.6	–1.8 ± 7.2
	θ	–2.9 ± 15.7	0.3 ± 12.7	1.0 ± 11.5
	(θ+α)/β	–0.1 ± 3.5	–0.3 ± 2.9	–0.4 ± 2.6
	β/α	–0.1 ± 5.9	0.8 ± 7.7	1.3 ± 5.6
Parietal	θ	1.5 ± 9.8	–0.3 ± 11.0	–0.3 ± 8.9
Occipital	β/α	0.1 ± 5.3	0.0 ± 4.6	0.0 ± 4.5

result of excessive tension or boredom which is an improper state of mental effort associated with risky driving.

Likewise, the psycho-physiological response within the $\mu \pm 1.64\ \sigma$ was categorized as a proper state of driver's response. The proper state was calculated for forty parameters, and eight typical examples were summarized in Table 4 and Table 5. The percentage of data out of the proper state for each parameter was displayed in Table 4 and Table 5. Based on the current analysis, approximately 10% of participants experienced the improper level of mental or physiological workload while driving. The four examples showing proper range were displayed in Figures 2–5.

Table 4. The relative percentage of proper/improper state of psycho-physiological response using amplitude parameters.

		Proper state of amplitude parameter	Improper
Lobe and parameter		Range	%
Left frontal	β	5.7~31.0	5.3
Right frontal	θ	24.1~56.7	8.9
Central	θ	23.0~50.96	10.2
Parietal	β	7.8~27.5	7.2%

Table 5. The relative percentage of proper/improper state of psycho-physiological response using slope parameters.

		Proper state of slope Parameter	Improper
Lobe and parameter		Range	%
Left frontal	α	−12.9~12.5	13.1
Left frontal	$(\theta+\alpha)/\beta$	−5.1~4.6	10.8
Parietal	θ	−18.2~18.2	8.3
Occipital	β/α	−7.7~7.8	12.0

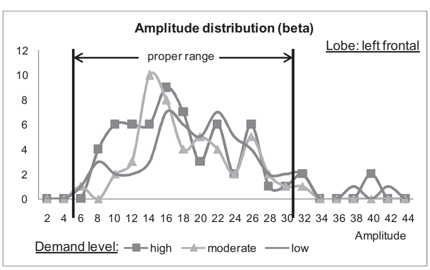


Figure 2. Distribution and the proper range of beta amplitude from left frontal lobe.

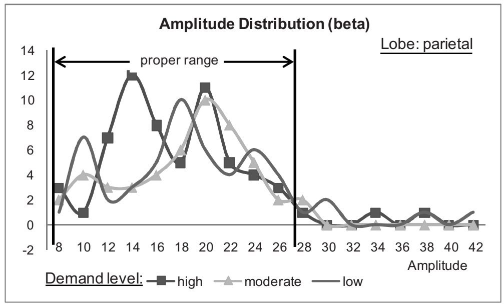


Figure 3. Distribution and proper range of beta amplitude from parietal lobe.

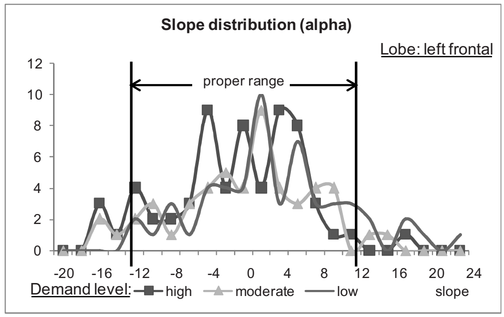


Figure 4. Distribution and proper range of alpha slope from left frontal lobe.

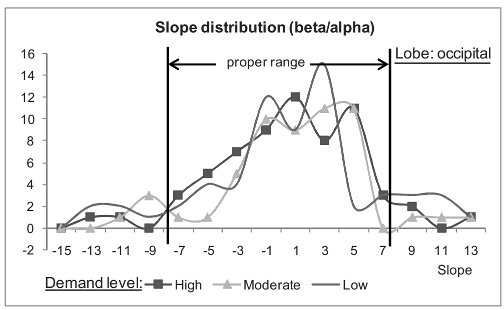


Figure 5. Distribution and proper range of beta/alpha slope from occipital lobe.

4 DISCUSSION

The main purpose of using the slope parameter was to observe the change in drivers' mental efforts over a certain time period, unlike previous researchers who observed only the instant state of the mental effort. This slope parameter included the rate of increase or decrease in physiological response during driving for 50 seconds. Using this method, we were able to observe continuing physiological changes due to different workloads or mental demands. Such slope parameters were used to identify extreme physiological values out of the proper state. For example, the extreme EEG signals were observed in both tails or just one side of the distribution. So we could determine that the signals positioned on either side of the tails represented either a hyper-or hypo-psycho-physiological response that could be considered an improper level of mental effort related to uncomfortable or risky driving. Moreover, the slope values used in this study have shown their utility in identifying the state of drivers with a particular pattern of psycho-physiology.

In this study, the proper state was used to identify the driver's mental state of normal and presumably comfortable driving. This approach was useful in determining the extreme physiological responses of EEG parameters. However, this approach is only effective when the parameter has a normal distribution.

Some parameters in this study were useful in evaluating a driver's mental workload, although the sensitivity of the parameters could be improved. The parameterization technique suggested in this study can be further used to examine the temporal characteristics of the signals. Additionally, the determination of proper state can be used as an alternative approach to defining normal and comfortable driving. Field practitioners can use experimental and quantitative approaches to identify the states of drivers that may lead to risky driving on the highway. Further research can be conducted in various fields using a different parameterization method or experimental validation.

5 CONCLUSIONS

The purpose of this study was to identify the characteristics of the EEG signals of drivers during highway driving. To do so, various driving workloads were defined and tested using different geometrical road conditions and accident records. The psycho-physiological response corresponding to various driving workloads were also defined in terms of numerical parameters, such as the slope and the amplitude of the parameters. The results of this study are expected to clarify the risks of driving associated with the driver's mental or physical strain.

REFERENCES

Brookhuis, K.A. & Waard, D.D. 1993. The use of psychophysiology to assess driver status. *Ergonomics* 33: 1099–1110.

Kecklund, G. & Åkerstedt, T. 1993. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics* 36: 1007–1017.

Kim, J.Y., Yoon, S.Y. & Whang, M.C. 2002. The development of a "demand-effort model" to evaluate the psychophysiological response of drivers on the highway. *4th International Conference on Psychophysiology in Ergonomics,* The University of Glasgow: 28–33.

Kim, J.Y., Yoon, S.Y. & Whang, M.C. 2003. The validation of the "demand-effort model". *Proceedings of the15th Triennial Congress of International Ergonomics Association*, Seoul: Korea, 4: 122–125.

Kim, J.Y., Yoon, S.Y. & Park, J.S. 2004. The study to diagnose the road-driver compatibility II: data collection, variables selection and parameter quantification. *Journal of the Korean Institute of Industrial Engineers* 31(1): 50–57.

Sterman, M.B. & Mann, C.A. 1995. Concepts and applications of EEG analysis in aviation performance evaluation. *Biological Psychology* 40: 115–130.

Torsvall, L. & Åkerstedt, T. 1987. Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroencephalography and Clinical Neurophysiology* 66: 502–511.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Skin temperature increase as a fitness indicator for hand-mouse interface evaluation—A preliminary infrared camera study

Hui-ju Lin, Yung-hui Lee & Wei-yu Chen
Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

ABSTRACT: Computer mouse is the most prevalent input device among VDT users. The "fitness" of mouse-hand interface is one of the most important factors to reflect the comfortable status. In this study we compared different size of mouse which used by one subject. Infrared thermographic camera was used to observe the temperature distribution of skin. The originality of this research is primarily in the use of thermography to measure thermal effects in hand-mouse interface. One male right-handed user subject took part in this study. He performed two tasks in this study to record the temperature of hand and mouse after the activities. The first part of test was static holding of mouse for 3 minutes. The second part of test is dynamic using condition included point and dragging tests and freely browse website for 5 minutes. Infrared (IR) camera was used to record hand and computer mouse surface temperatures. Contact area can show the hand-mouse interface clearly. The skin temperature change is relatively higher at posterior palm side (3.7°C higher), and the temperature change of mouse is higher in the posterior-right side (increase 7.8°C) after 15 min website browsing. Infrared thermographic camera is a useful method which can quickly investigate the fitness of hand-mouse interface.

Keywords: computer mouse, infrared, skin temperature, human computer interface

1 INTRODUCTION

1.1 *Computer mouse interface*

Of all the input devices of computer (including mice, trackpads, trackballs, joysticks, light pens, etc.), computer mouse is the most prevalent among VDT users, both in terms of number and daily operation time (Jensen et al., 1998). In the past few years, impressive developments have taken place regarding computer mouse interface. Using mouse some people complained about heating sensation on the palm and fingers. When the mouse is hold against the palm, it functions as a heat insulator. Thereby it reduces the heat that is normally lost to the air. For mouse designed with a "form fit" concept, to achieve a best fit interface between the shape of the hand and the surface contour of the mouse, may cause additional increase skin temperature. Holding a mouse while maintaining finger tip pressure on the button and dragging also resulted in skin heating sensations, however, users may exhibit sustained, static finger lifting behaviours to prevent inadvertent activations by avoid finger pressure on the buttons (Lee, et al., 2008). Our research would like to use a new method of infrared (IR) camera to evaluate the the change in skin temperature between computer mouse and hand interface. This finger lifting behaviours could, on the other hand, release the cumulative heat. It is the size, the form, and the key location of the mouse that determined the operating postures of wrist, hand, and fingers in mouse use and the amount of heat that was insulated on the surface skin of the hand.

1.2 *Infrared (IR) camera as a skin temperature detector*

Thermographic camera was used to observe thermal distributions over a surface and obtain thermometric data (Gulyaev, et al., 1995). Recent development in this technology incorporated infrared imaging for material evaluations, medical diagnosis, and thermal comfort assessments (De Oliveira, et al., 2007). Taurisano and Vorst (2000) used thermographic camera to observe thermal distributions over the skin surface of the ear regions and found the average temperature increase on the ear lobe range from 1.0 to 2.4 ± 0.2°C. Examining factors that causing complaints about burning sensations of the ear region for mobile phone users, Straume, et al. (2005) found that the insulation and the electrical power dissipation lead to statistically significant rises in skin temperature, while the radio frequency exposure did not. Theoretical methods to calculate the skin temperature elevation due to various factors can be found in Bernardi et al.'s study (2001).

1.3 *Purpose of this study*

From a design point of view, we are interest to know would skin temperature rise be an effective indicator for subjective evaluation of human sensation while using mouse with different design features. The originality of this research is primarily in the use of thermography to measure thermal effects in hand-mouse interface. It was hypothesized that skin temperature dynamics could be used to investigate the interaction of different factors involved in comfort rating of a hand-mouse system. The distribution and dynamics of surface temperature of the hand-mouse interfacing system in this study were examined different duration of mouse exposures. Other disturbing factors such as the environmental temperatures, subject's personality, mood, and psychological state were considered as controlled variables in the study.

2 MATERIALS AND METHODS

2.1 *Participant*

The same subject (male, 23 years old) was used for the whole experiment with hand length (wrist joint to 3rd finger tip) 17.6 cm, hand breadth 8.0 cm. The subject was restricted to any activities which will increase the heart rate and metabolic rate to avoid increase the body temperature. And also not allowed to eat or drink anything cold or hot during and 30 minutes before the sessions. Smoking and taking snuff as well as heavy meals was not allowed later than 1 h before a session. This subject was right handed user and without any past history of hand and wrist dysfunctions.

2.2 *Experimental design*

There were two parts of the experiment. The first part of test was static holding of mouse for 3 minutes. Measurement of skin temperature was taken before test and for each following 1 minute. The second part of test is dynamic using condition (freely browsing website use mouse only) for 15 minutes. This study use one mouse to test (height: 3.9 cm × width: 5.9 cm × depth 11.6 cm; weight: 98.3 g; symmetrical design with a scroll in the middle side). For all tasks, participants were seated at a workstation with an adjustable chair. Before the test start, all participant can adjust the seat to a comfortable position with both leg can touch the ground and arm can rest at the table.

2.3 *Infrared (IR) camera settings*

Infrared (IR) camera (Therma CAM P25, FLIR System Inc.) was used to record skin and computer mouse surface temperatures. The camera has a 24° × 18° field of view and an uncooled microbolometer 320 × 240 pixels. The minimum focus distance was 0.3 m and the thermal sensitivity was 0.08°C at 30°C. Measurement accuracy was ±2°C and ±2%.

The IR camera was place on the tripod besides the mouse working area and the lens toward to table with a height of 60 cm (Fig. 1). The area of working area and IR picture taking area were separated to avoid interfered of background temperature. Participant was sitting in front the monitor performing the tasks. In the static task, pictures were taken immediately before the task and after 3 minutes. In the dynamic using mouse test, pictures were taken immediately before the task and after first test, second test, and browsing website.

2.4 *Data analysis*

The temperatures of hand were recorded into five areas as following: thumb, index and middle finger, ring finger and little finger, anterior palm side, posterior side (Fig. 2). And the computer mouse was divided into four areas: anterior-right side, anterior-left side, posterior-right side, posterior-left side (Fig. 3).

The surface temperatures of skin and computer mouse were recorded for descriptive analysis. And histogram was analyzed for the centralization and distribution shift of the temperature.

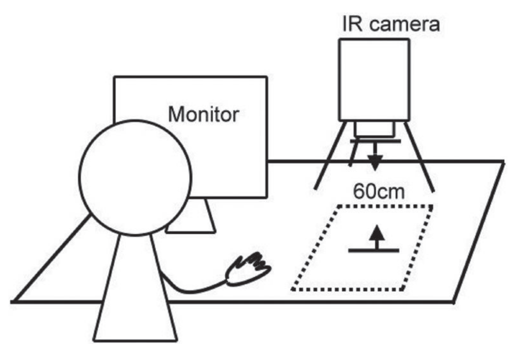


Figure 1. The experiment apparatus of IR camera and monitor. The distance from the camera lens to the table was 60 cm.

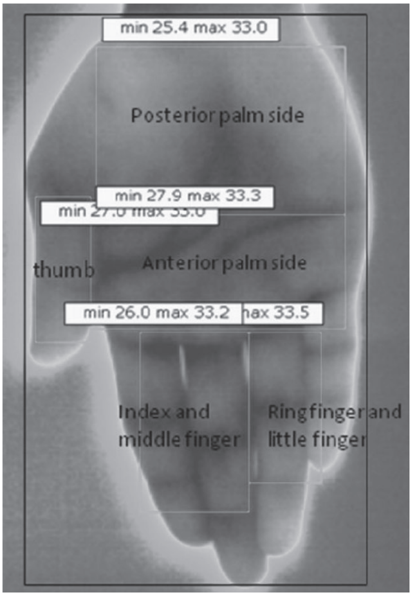


Figure 2. The areas divided into five parts to analysis the temperature change of the hand.

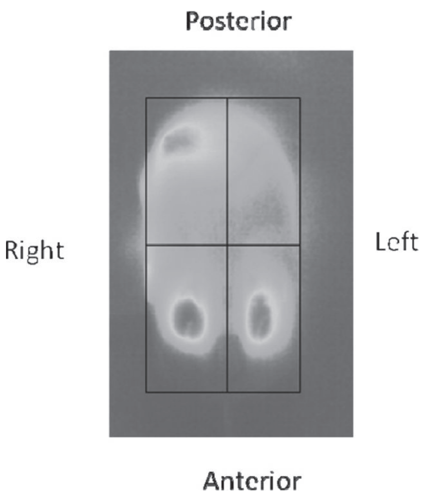


Figure 3. Computer mouse was divided into four parts to analysis the temperature change.

3 RESULT

3.1 *Temperature during static holding*

The temperature changes during 3 minutes static holding were shown in Table 1.

Thumb and the anterior palm area were found the highest change in the temperature. Both increase 1.9°C of this two area. And the temperature changes in computer mouse were shown in Table 2. The IR image of 0 minute and 15 minute were shown on Fig. 4 and Fig. 5.

Table 1. Skin temperature (°C) and temperature change during 3 minutes static holding of computer mouse.

Hand area	T0	T3	Temperature change
Thumb	31.9	33.8	1.9
Index-middle finger	32.2	33.3	1.1
Ring and little ginger	32.2	33.4	1.2
Anterior palm	31.7	33.6	1.9
Posterior palm	32.3	34	1.7

Table 2. Computer mouse temperature (°C) and temperature change during 3 minutes static holding of computer mouse.

	T0	T3	Temperature change
Ant-right	24.7	28.2	3.5
Ant-left	25	27.8	2.8
Post-right	24.8	28.3	3.5
Post-left	25	27.2	2.2

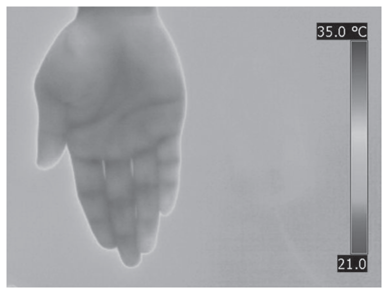


Figure 4. The IR camera image of hand and computer mouse before holding the mouse. The color spectrum showed the temperature between 21°C to 35°C.

Figure 5. The IR camera image of hand and computer mouse after static holding the mouse for three minutes. The contact areas of hand and mouse interface were mainly at index and third finger area and posterior palm area.

3.2 *Temperature during dynamic using mouse*

The temperatures in the beginning and after 15 minutes dynamic using the mouse were shown in Table 3 and Table 4 respectively.

Surface skin temperature get higher of 3.7°C in the posterior palm side was the most remarkable area. The IR image of after 15 minute using mouse were shown on Fig. 6. The temperature rises around the area of posterior part of computer mouse corresponding to the area of hand.

3.3 *Histogram of temperature distribution*

The histogram of temperature distribution in posterior palm side and the posterior-right side of computer mouse during the dynamic mouse using task were shown in Figure 7 and Figure 8 respectively. The frequency means the quantity of temperatures pixels detected by the IR camera in each temperature. The temperature distribution of hand was shifted and centralized to higher degree after using mouse for 15 minutes. And the temperature change/of computer mouse was more spread after 15 minutes use.

Table 3. Surface skin temperature of hand (°C) and temperature change during 15 minutes using of computer mouse.

	T0	T15	Temperature change
Thumb	31.2	33.6	2.4
Index-middle finger	30.6	33.7	3.1
Ring and little ginger	30.5	33.9	3.4
Anterior palm	30.5	33.8	3.3
Posterior palm	30.8	34.5	3.7

Table 4. Computer mouse temperature (°C) and temperature change after 15 minutes using of computer mouse.

	T0	T15	Temperature change
Ant-right	24	29.3	5.3
Ant-left	24	30	6
Post-right	24.2	32	7.8
Post-left	24.2	31.5	7.3

Figure 6. The IR camera image of hand and computer mouse after dynamic web browsing for 15 minutes. The contact areas of hand and mouse interface were mainly at index and third finger area and posterior palm area.

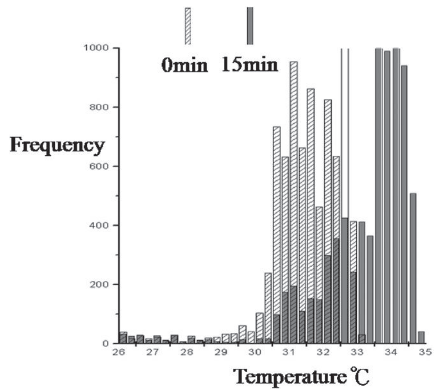


Figure 7. The histogram of temperature distribution of posterior palm side at the beginning and 15 minutes us of computer mouse.

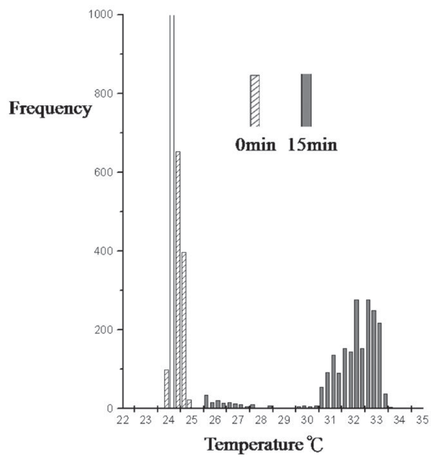


Figure 8. The histogram of temperature distribution of posterior-right side of computer mouse.

4 DISCUSSION

4.1 *IR camera as a equipment to evaluate hand-mouse interface*

The first purpose of our study is to use IR thermography as a tool to evaluate the hand-mouse interface. Hand-mouse interface variances depending

Figure 9. The static holding position taking with a regular camera. In this photograph we can only see the outside feature of gripping a mouse but not the real contact area.

on the form of the mouse, the size of the user's hand, and the way they used to hold the mouse (Lee et al., 2008). As Figure 9 we can see our subject hold the computer mouse with a "holly grip" position which only the finger tips and posterior palm contact to the mouse. This regular photograph cannot represent the real contact situation of the interface. In this case the IR camera took picture after 3 minutes static holding the mouse (Figure 5) can give a very clear contact picture. The "form fit" problem of mouse can show on the mouse in which part it is contacted. This is a very efficient application for designing the computer mouse.

4.2 *Temperature change in different area of hand and computer mouse*

In the study we analyzed the temperature of hand into five areas corresponding to the functional use of click (index and third finger), thumb (holding), ring and little finger (holding), anterior and posterior palm (supporting). When using the computer mouse, the increase of temperature is mainly cause by the insulation of heat (Straume, et al., 2005). During static holding of mouse for 3 minutes, the temperature slightly increase from 1.1 to 1.9°C in hand, and 2.2 to 3.5°C (Table 1, Table 2) in the mouse. After 15 minutes using the mouse, the the temperature increase in 2.4 to 3.7°C in hand, and 5.4 to 7.8°C (Table 3, Table 4) in the mouse. The short term use of the mouse can show the initial contact of the hand-mouse interface. It is very valuable to reveal the form-fit problem which useful for designer. The prolong use of mouse can show the area by heat insulation. It is correlated to the subjective feeling of comfortable. The ISO thermal sensation scale collected the subjective feeling of thermal change. People had "hot" sensation after the temperature increase 3°C in average (Olesen 2002). So in our case, most area of hand temperature increase over 3°C, especially in the posterior palm side and posterior-right mouse area. We analysis those areas in temperature histogram furthermore (Figure 7, Figure 8). The histogram showed hand temperature was increase and centralizes after using 15 minutes. And the mouse temperature was in crease and spread after use. This means in certain areas the insulation effect more and the other were not. This message can be use further to test different design, size, user how they fit and comfort using the mouse.

4.3 *Methodological control of IR camera*

Measurement of surface skin temperature may have many factors to control. Many factors will cause temperature change including which like blood flow, heat transfer between neighboring regions, and water evaporation from the surface. Psychological factors may also affect the temperature change like emotional change (Gulyaev et al., 1995). To avoid the confounding factors may affect the temperature, the subjects should control activity, and the environment convention should be controlled in the same situation (Straume, et al., 2005). In our study, we found that after holding the mouse, the temperature of mouse and hand have to cool down after at least 5 minutes. And this may be the recommendation time interval of each test for further test.

5 CONCLUSION

IR image is a new method to exam the hand—mouse interface efficiently. It can reveal the problem of mouse design in the early stage and save cost from inappropriate design. Also, it can distinct the specific area of the interface. We need further examination to see the difference between size, and shape of mouse, and also different hand size of user. The further completed experiment is need for variety of test conditions.

REFERENCES

Bernardi, P., Cavagnaro, M., Pisa, S. & Piuzzi, E. 2001. Power absorption and temperature elevations induces the human hard by a dual-band monopole-helix antenna phone. *IEEE Transactions on Microwave Theory and Techniques* 49(12): 2539–2546.

De Oliveira, F., Moreau, S., Gehin, C. & Dittmar, A. 2007. Infrared imaging analysis for thermal comfort assessment. *Proceeding of the 29th annual international conference of the IEEE EMBS.* Cité Internationale, Lyon, France.

Gulyaev, Y.U., Markov, A.G., Koreneva, L.G. & Zakharov, P.V. 1995. Dynamic infrared thermography in humans. *IEEE engineering in medicine and biology* 14(6): 766–771.
Jensen, C., Borg, V., Finsen, L., Hansen, K., Juul-Kristensen, B. & Christensen, H. 1998. Job demands, muscle activity and musculoskeletal symptoms in relation to work with the computer mouse. *Scand J Work Environ Health.* 24(5): 418–24.
Lee, D.L., McLoone, H. & Dennerlein, J.T. 2008. observed finger behaviour during computer mouse use. *Applied Ergonomics* 39: 107–113.
Straume, A., Offedal, G. & Johnesson, A. 2005. Skin temperature increase caused by a mobile phone: A methodological infrared camera study. *Bioelectromagnetics* 26: 510–519.
Olesen, B.W. & Parsons, K.C. 2002. Introduction to thermal compfort standards and to the proposed new version of EN ISO 7730. *Energy and Building* 34: 537–548.
Taurisano, M.D. & Vorst, A.V. 2000. Experimental thermographic analysis of thermal effects induced on a human head exposed to 900-MHz fields of mobile phones. *IEEE transactions on Microwave Theory and Techniques* 48(11): 2022–2032.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluation of cyclic one-handed handling by electromyography

Seung-Min Mo, Jongseon Kwag & Myung-Chul Jung
Department of Industrial and Information Systems Engineering, Ajou University, Suwon, South Korea

Yong-Ku Kong
Department of Systems Management Engineering, Sungkyunkwan University, Suwon, South Korea

ABSTRACT: The purpose of this study is to evaluate the trends in the frequency and energy for repetitive lifting and lowering tasks by means of surface Electromyography (sEMG). A total of fifteen male subjects participated in the experiment, and sEMG signals were recorded from three right-sided muscles of the anterior-deltoid, posterior-deltoid, and upper-trapezius. Four independent variables were considered, namely, the weight (1, 4, and 7 kg), frequency (2, 6, and 10 cycles/minute), height (30, 60, and 90 cm), and motion (lifting and lowering), which affected one-handed handling; a full factorial experimental design was employed. This study considered two kinds of dependent variable: (1) the slope of the instantaneous mean power frequency (IMNF) and (2) the slope of the Instantaneous Energy (IE). The results were analyzed in terms of the main and interaction effects. The significance level was set at 0.05 for all the tests. Repetitive dynamic sEMG signals were segmented into three sections (first, middle, and last) and time-frequency analysis conducted for each segment by Continuous Wavelet Transforms (CWT). Then, using linear regression analysis, the trends in the slope for the two parameters were estimated. As a result, the average slope of IMNF with regard to the weight, frequency, and height was significantly lower as the levels of these factors increased. However, the average slope of IE was significantly higher as the level of the frequency factor increased. The trend with regard to the frequency shifts was similar to the results of previous studies that analyzed the mean or median frequency values by means of the time-frequency transform. Furthermore, the increase in the slope of IE can be described in terms of muscular fatigue that is related to the phenomenon of decreasing frequency. This study can suggest a proper method for sEMG signal processing for movement that occurs cyclically and under varying force.

Keywords: one-handed handling, risk factors, electromyography, continuous wavelet transform, fatigue

1 INTRODUCTION

Recently, most of the tasks to be directly performed by workers have been replaced by the mechanical power of machinery in industrial fields. Accordingly, the ongoing automation of production facilities has significantly decreased the need for muscular strength on the part of workers; working procedures have become simplified and easy to accomplish. However, there are still some areas where a machine cannot take over from workers who can do elaborate tasks. Various fields of industry have been adopting automated and mechanized systems; however, tasks still remain that require workers' strength (e.g., simply repetitive tasks, micro-level operation of machines, awkward lifting in a confined space, etc.). Furthermore, in the future, Manual Material Handling (MMH) by industrial workers will continue to be required.

Most of the improvements for industrial workers are associated with MMH that causes high compensation costs and labor losses (Ciriello 2005, Murphy & Courtney 1996). MMH consists of lifting, lowering, pushing, pulling, carrying, and keeping, but especially the lifting task has been considered as the main concern in terms of work-related safety and health issues in industrial fields. Typically, the lifting task occurs repeatedly over a long time; as a result, the shoulders of workers will be accompanied by chronic pain that is caused by musculoskeletal disorders. This is confirmed by the fact that most workers suffer from a high degree of shoulder pain (Brox 2003, Sommerich & Huges 2006).

The previous studies on MMH have concentrated on vertical lifting and lowering and horizontal pushing and pulling motions based on the symmetric use of both hands. The prevention of musculoskeletal disorders, as related to

MMH, was proposed by the National Institute for Occupational Safety and Health (NIOSH) in the US in 1981 for the maximum permissible limits under symmetrical, two-handed handling. Zhu & Zhang (1990) also reported that the previous studies conducted in the US and Europe had a limitation to apply to relatively small Chinese people. As a result, the maximum acceptable weights and workload suggested in the previous studies should be lower than the suggestions of the previous studies. Lin et al. (2006) evaluated the physical stress of manual lifting jobs by using the ratio of joint moment to joint capacity based on a nine-link joint model. As a result, the proposed method found having a high correlation with the NIOSH lifting index. Parakkat et al. (2007) analyzed differences in workload under asymmetric lifting between novice and experienced subjects. The biomechanical lower back loads were found to be similar but novice subjects had larger discomfort than experienced subjects. It may be reported that novice subjects seemed to have a lower threshold of discomfort.

The previously reported studies were undertaken by using both hands related to MMH. Based on surface electromyography (sEMG) and subjective discomfort rating, some studies were conducted to analyze the fatigue of the shoulder part in relation to combined, overhead, one-handed work (Sood et al. 2007; Garg et al. 2006). Also, another study was designed to investigate the effects of the temperature on the kinematic data and characteristics of sEMG signals (EMG gaps) during the repetitive handling of light-weight objects in the horizontal plane. Besides, only a limited number of studies have been undertaken to simply separate static and dynamic components of sEMG signals and consider relative Maximum Voluntary Contractions (MVC) of some muscles (Strasser & Müller 1999; Kothiyal & Kayis 2001).

Most workers perform tasks for a long time with regard to repetitive movements in manufacturing and production lines in actual industrial fields. In this regard, the movements require dynamic muscle contraction instead of steady-state processes (Gerdle et al. 2000). During dynamic states, several factors, e.g., changes in the fiber length and angle of joint through the range of motion, may increase the nonstationarity of the sEMG signal (Larsson et al. 1999). This is a major reason why other suitable signal processing techniques must be used to analyze the signals of dynamic muscle contraction. Knaflitz and Bonato (1999) analyzed the differences of frequency shifts related to muscle fatigue according to the time-frequency transform technique. They reported the need for proper sEMG signal processing techniques in accordance with the characteristics of contraction for obtaining reliable data during dynamic muscle contraction.

This study evaluates the workload of the shoulder for repetitive lifting and lowering tasks by means of sEMG. Localized muscle fatigue and energy generation is analyzed by using a dynamic signal processing method in order to consider the main risk factors of weight, frequency, and height related to MMH.

2 METHODS

2.1 *Subjects*

A total of fifteen male subjects volunteered to participate in the experiment. All fifteen subjects were right-handed. None of the subjects reported any history of musculoskeletal disorders in the previous year. Also, all subjects had previously completed a training session.

2.2 *Experimental design*

The experimental design consisted of four within-subjects factors of weight (1, 4, and 7 kg), frequency (2, 6, and 10 cycles/minute), height (30, 60, and 90 cm), and motion (lifting and lowering) for the independent variables. Subjects participated in a total of 27 (3^3) experimental conditions with a full-factorial design (except for the motion). All the experimental conditions were randomized. This study considered two kinds of dependent variable in three muscles, namely, the slopes of (1) the Instantaneous Mean Power Frequency (IMNF) and (2) the Instantaneous Energy (IE) from continuous wavelet transforms and linear regressions.

2.3 *Apparatus*

For recording physiological signals of repetitive muscle contraction, sEMG equipment was used (DataLINK DLK 900, Biometric Ltd., The UK). The standing workstation height was 90 cm based on the floor of our laboratory. Dumbbells were used for the load with the grip diameter of 3.04 cm. Furthermore, a metronome was used for maintaining the frequencies.

Table 1. Summary of anthropometric data for the fifteen male subjects.

Anthropometry	Mean	SD
Age (year)	25.47	2.1
Height (cm)	175.49	4.0
Weight (kg)	69.41	5.3
Elbow height (cm)	109.81	3.5

2.4 *Procedure*

To reduce the electrical impedance, the skin was shaved and cleaned with alcohol. Subsequently, silver-silver Chloride (Ag/AgCl) bipolar electrodes (Biometrics SX-230, Biometric Ltd., The UK) with a diameter of 10 mm and an inter-electrode distance of 20 mm were attached to each subject. All the data were converted from analog to digital (13 bits) at a sampling rate of 1 kHz. The CMRR of the unit was 96 dB, the gain was set at 1000, and the noise was below 5 μV.

The sEMG signals were recorded from three right shoulder muscles of the Anterior-Deltoid (AD), Posterior-Deltoid (PD), and Upper Trapezius (UT). These muscles were selected because they are the major muscles related to vertical lifting and lowering movements (Hislop and Montgomery, 2007). The position of electrodes also was in accordance with the SENIAM recommendations (Hermens et al. 2000). A reference electrode was attached over the dorsal surface of the right wrist.

The subjects performed a repetitive one-handed handling task during a ten-minute period. The experiment progressed until the subject was no longer able to conduct the task. Figure 1 shows a cycle of a one-handed handling movement.

During the experiment, the subjects were not allowed to bend their trunks. A break for about 10 minutes was provided after each experimental condition. During the break, all the electrodes were not allowed to touch the subjects.

2.5 *Signal processing and statistics*

This study used the Continuous Wavelet Transform (CWT) that is widely known as one of the appropriate signal processing techniques for dynamic movement. The CWT technique is able to analyze time and frequency domains at the same time and is suitable for nonstationary biological signal analysis (Ismail & Asfour 1998, Karlsson et al. 2001). Barcier et al. (2009) and Kellenberg et al. (2009) used the CWT technique according to dynamic muscle contraction.

The signals of sEMG collected from each experimental condition were segmented into initial, middle, and final sections. The signal for each of the three segments was repartitioned into lifting and lowering sections to consider the independent variable of motion. The signal segmentation followed the method that Neptune et al. (1997) used. Each segmented signal was filtered by means of a fourth-order Butterworth bandpass filter (10–500 Hz). Then, instantaneous mean power frequency (IMNF) and Instantaneous Energy (IE) were estimated from each segmented signal using a CWT with the Morlet basis wavelet function. The slopes of IMNF and IE were computed by using linear regression analysis applied to the values of initial, middle, and final sections in order to obtain the slope for the two parameters of MNF and energy.

All the data processing and statistical analyses were performed using the MATLAB (Ver. 7.0.4, Mathworks, The US) and SAS (Ver. 9.1, SAS Institute, Inc.) programs. The results were analyzed in terms of the main and interaction effects with respect to the average and standard deviation by means of an analysis of variance and post-hoc test (Tukey's HSD). The significance level was set at 0.05 for all tests.

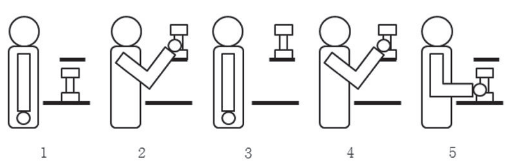


Figure 1. One cycle of a one-handed handling task: 1) resting, 2) lifting after grasping the dumbbell, 3) resting again, 4) grasping the dumbbell, and 5) lowering.

3 RESULTS

3.1 *Slope of IMNF*

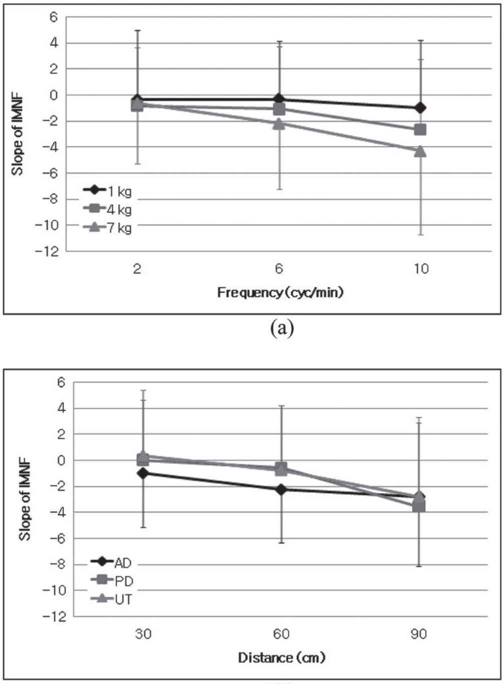
The analysis of variance for the slope of IMNF indicated that there were the significant main effects of weight ($p < 0.0001$), frequency ($p < 0.0001$), and height ($p = 0.0204$). Generally, the average slope of IMNF for a significant factor decreased as the levels of the factors increased.

According to the result of the *post-hoc test* (Tukey's), no significant difference for the weight factor was found between the 4 and 7 kg loads. No significant difference for the frequency factor was found between 2 and 6 cycle/min. In the case of the height factor, all the levels were analyzed as having significant differences.

The weight * frequency ($p < 0.0001$), weight * distance ($p = 0.0177$), frequency * distance ($p = 0.0020$), distance * muscle ($p = 0.0117$), and weight * frequency * distance ($p = 0.0088$) interaction effects were also significantly different. Figure 2 illustrates some interactions.

3.2 *Slope of IE*

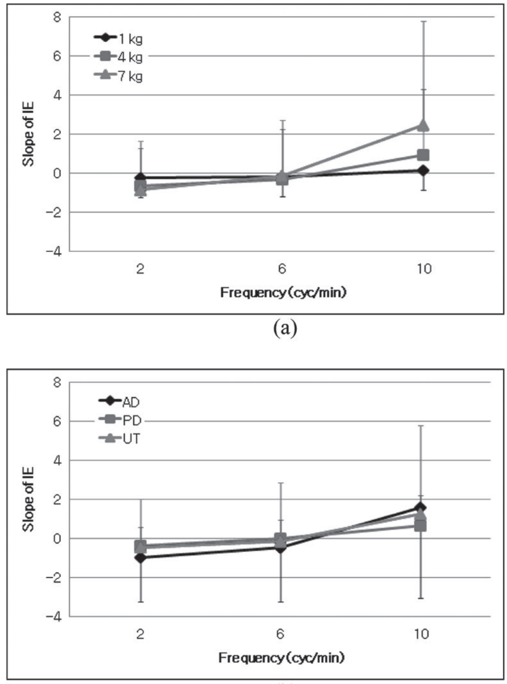
The analysis of variance according to the slope of IE indicated that the main effects of frequency ($p < 0.0001$) and motion ($p = 0.0204$) were significantly affected. Generally, the average slope of IE with the frequency was significantly higher as the levels of the factors increased.



(a)

(b)

Figure 2. Average slope of IMNF (average ± SD) in relation to interaction effects between the muscles and factors. (a) Weight * Frequency. (b) Distance * Muscle.



(a)

(b)

Figure 3. Average slope of IE (average ± SD) in relation to interaction effects between the muscles and factors. (a) Weight * Frequency. (b) Frequency * Muscle.

According to the result of the *post-hoc test* (Tukey's), no significant difference was found for the frequency factor between 2 and 6 cycle/min. Also, the slope of IE recorded was significantly higher during the lifting motion than the lowering motion.

Additionally, the weight * frequency ($p < 0.0001$), frequency *muscle ($p = 0.0016$), frequency * motion ($p = 0.0039$), motion * muscle ($p = 0.0048$), weight * frequency * muscle ($p = 0.0118$), and weight * frequency * motion ($p = 0.0092$) interaction effects were also revealed significant differences. Figure 3 illustrates some interactions.

4 DISCUSSION

The IMNF significantly decreased as the levels of the main risk factors of weight, frequency, and height increased. The results of the analysis were similar to those of the previous studies that analyzed the mean or median frequency in terms of time-frequency transform technique (Bonato et al. 2001; Englehart et al. 1999; Pope et al. 2000). In the case of continuous muscle contraction, a decreasing frequency represented muscle fatigue resulted from slower rate of depolarization of muscle membranes because of the reduction of the Motor Unit (MU) conduction velocity along the myofibrils and the depletion of metabolites within the cells (Potvin and Bent, 1997). Also, Doud & Walsh (1965) reported that a decreasing frequency reflected localized muscle fatigue due to the increased accumulation of lactic acid owing to higher Motor Unit Action Potentials (MUAP). Therefore, a negative slope of IMNF means that as the levels of the factors increased, the rate muscle fatigue increased.

This study analyzed the IE value that was calculated in terms of scale bands by summing up all the absolute values of coefficients of CWT, not the amplitude and magnitude expressed by the Root Mean Square (RMS) or Mean Absolute Value (MAV). Figure 3(a) shows that the slopes of IE significantly changed to negative to positive; as the level of weight increased, the rate of change in the slope of IE increased. The increase in the slope of IE can be explained with muscle fatigue related to the repetitive movements (Merletti et al. 1990). Task frequency refers to the workload per minute in actual industrial fields. Accordingly, the firing rate of MUs has gradually decreased and the early recruitment of MUs during continuous repetitive muscle contraction has reduced (Kellenberg & Hermens 2008). Due to these effects, it may be considered even though the same strength was generated, a relatively larger amount of MUs were recruited compared with the early phase of the lifting and lowering task (Sbriccoli et al. 2003). In addition, according to the *post-hoc test* results regarding the slope of IE, the frequency levels of 2 and 6 cycles/minute had the same slopes. This could explain low fatigue resulted from the resting period between trials.

5 CONCLUSION

This study has demonstrated the dynamic sEMG signal processing method with continuous wavelet transforms for prolonged and cyclic dynamic muscle contraction. This study will help to quantitatively identify the relationship between energy and fatigue in terms of weight, frequency, and height. In future, empirical studies is need to recommend guidelines to workers who suffer from occupational muscular skeletal disorders.

REFERENCES

Barcier, S., Halin, R., Ravier, P., Kahn, J.F., Jouanin, J.C., Lecoq, A.M. & Buttelli, O. 2009. The vastus lateralis neuromuscular activity during all-out cycling exercise. *Journal of Electromyography and Kinesiology* 19(5): 922–930.

Bonato, P., Row, S.H., Knaflitz, M. & DeLuca, C.J. 2001. Time-frequency parameters of the surface myoelectric signals for assessing muscle fatigue during cyclic dynamic contraction. *IEEE Transactions on Biomedical Engineering* 48(7): 745–753.

Brox, J.I. 2003. Shoulder pain. *Best Practice & Research Clinical Rheumatology* 17(1): 33–56.

Ciriello, V.M. 2005. The effects of box size, vertical distance, and height on lowering tasks for female industrial workers. *International Journal of Industrial Ergonomics* 35(9): 857–863.

Doud, J.R. & Walsh, J.M. 1995. Muscle fatigue and muscle length interaction: Effect on the EMG frequency components. *Electromyography and Clinical Neurophysiology* 35(6): 331–339.

Englehart, K., Hudgins, B., Parker, P.A. & Stevenson, M. 1999. Classification of the myoelectric signal using Time-Frequency based representations. *Medical Engineering & Physics* 21(6/7): 431–438.

Garg, A., Hegmann, K. & Kapellusch, J. 2006. Short-cycle overhead work and shoulder girdle muscle fatigue. *International Journal of Industrial Ergonomics* 36(6): 581–597.

Gerdle, B., Larsson, B. & Karlsson, S. 2000. Criterion validation of surface EMG variables as fatigue indicators using peak torque: a study of repetitive maximum isokinetic knee extensions. *Journal of Electromyography and Kinesiology* 10(4): 225–232.

Hermens, H.J., Freriks, B., Disselhorst-Klug, C. & Rau, G. 2000. Development of recommendations for SEMG sensors and sensor placement procedures. *Journal of Electromyography and Kinesiology* 10(5): 361–374.

Hislop, H.J. & Montgomery, J. 2007. *Daniels and Worthingham's muscle testing: techniques of manual examination (8th ed.)*. Saunders/Elsevier.

Ismail, A.R. & Asfour, S.S. 1998. Continuous wavelet transform application to EMG signals during human gait. *Asilomor Conference on Signals, Systems & Computers* 1: 325–332.

Karlsson, S. & Gerdle, B. 2001. Mean frequency and signal amplitude of the surface EMG of the quadriceps muscles increase with increasing torque—A study using the continuous wavelet transform. *Journal of Electromyography and Kinesiology* 11(2): 131–140.

Kellenberg, L.A.C. & Hermens, H.J. 2008. Behaviour of a surface EMG based measure for motor control: Motor unit action potential rate in relation to force and muscle fatigue. *Journal of Electromyography and Kinesiology* 18(5): 780–788.

Kellenberg, L.A.C., Preece, S., Nester, C. & Hermens, H.J. 2009. Reproducibility of MUAP properties in array surface EMG recordings of the upper trapezius and sternocleidomastoid muscle. *Journal of Electromyography and Kinesiology* 19(6): e536–e542.

Knaflitz, M. & Bonato, P. 1999. Time-Frequency methods applied to muscle fatigue assessment during dynamic contractions. *Journal of Electromyography and Kinesiology* 9(5): 337–350.

Kothyial, K. & Kayis, B. 2001. Workpalce layout seated manual handling tasks: an electromyography study. *International Journal of Industrial Ergonomics* 27(1): 19–32.

Larsson, B., Månsson, B., Karlberg, C., Syvertsson, P., Elert, J. & Gerdle, B. 1999. Reproducibility of surface EMG variables and peak torque during three sets of ten dynamic contractions. *Journal of Electromyography and Kinesiology* 9(5): 351–357.

Lin, C.J., Wang, S.J. & Chen, H.J. 2006. A field evaluation method for assessing whole body biomechanical joint stress in manual lifting tasks. *Industrial Health* 44(4): 604–612.

Merletti, R., Knaflitz, M. & DeLuca, C.J. 1990. Myoelectric manifestations of fatigue in voluntary and electrically elicited contractions. *Journal of Applied Physiology* 69(5): 1810–1820.

Murphy, P.L. & Courtney, T.K. 1996. Low back pain disability: relative costs by antecedent and industry group. *American Journal of Industrial Medicine* 37(5): 558–571.

Neptune, R.R., Kautz, S.A. & Hull, M.L. 1997. The effect of pedaling rate on coordination in cycling. *Journal of Biomechanics* 30(10): 1051–1058.

Parakkat, J., Yang, G., Chany, A.M., Burr, D. & Marass, W.S. 2007. The influence of lift frequency, lift duration and work experience on discomfort reporting. *Ergonomics* 50(3): 396–409.

Piedrahita, H., Oksa, J., Malm, C., Sormunen, E. & Rintämaki, H. 2008. Effect of cooling and clothing on vertical trajectories of the upper arm and muscle functions during repetitive light work. *European Journal of Applied Physiology* 104(2): 183–191.

Pope, M.H., Aleksiev, A., Panagiotacopulos, N.D., Lee, J.S., Wilder, D.G., Friesen, K., Stielau, W. & Goel, V.K. 2000. Evaluation of low back muscle surface EMG signals using wavelets. *Clinical Biomechanics* 15(8): 567–573.

Potvin, J.R. & Bent, L.R. 1997. A validation of techniques using surface EMG signals from dynamic contractions to quantify muscle fatigue during repetitive tasks. *Journal of Electromyography and Kinesiology* 7(2): 131–139.

Sanders, M.S. & McCormick, E.J. 1992. *Human factors in engineering and design (7th ed.)*. McGrow-Hill, Inc.

Sbriccoli, P., Bazzucchi, I., Rosponi, A., Bernardi, M., DeVito, G. & Felici, F. 2003. Amplitude and spectral characteristics of biceps brachii sEMG depend upon speed of isometric force generation. *Journal of Electromyography and Kinesiology* 13(2): 139–147.

Sommerich, C.M. & Hughes, R.E. 2006. Aetiology of work-related disorders of the rotator cuff tendons: research and theory. *Theoretical Issues in Ergonomics Science* 7(1): 19–38.

Sood, D., Nussbaum, M.A. & Hager, K. 2007. Fatigue during prolonged intermittent overhead work: reliability of measures and effects of working height. *Ergonomics* 50(4): 497–513.

Strasser, H. & Müller, K.W. 1999. Favorable movements of the hand-arm system in the horizontal plane assessed by electromyographic investigations and subjective rating. *International Journal of Industrial Ergonomics* 23(4): 339–347.

Zhu, Z. & Zhang, Z. 1990. Maximum acceptable repetitive lifting workload by Chinese subjects. *Ergonomics* 33(7): 875–884.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluation of motor preparatory activity in manual control using the measurement of ERP

Masato Sakurai & Takuya Ichikawa
Tokyo University of Science, Tokyo, Japan

Miwa Nakanishi
Keio University, Kanagawa, Japan

Sakae Yamamoto
Tokyo University of Science, Tokyo, Japan

ABSTRACT: To examine the relationship between the Event-Related Potential (ERP) and the changes of EMG in the right hand of observers in manual control, it was measured the physiological responses of EEG and EMG when steering a wheel to handle the object through the gate as the task. In the results, there are the negative slopes of ERP before steering the wheel. At first, the potentials rise gradually negative (N1) and change rapidly (N2) to the peak at the time of steering the wheel. It is suggested that N1 is the preparatory activity to steer the wheel and N2 is the activity to exercise the muscle in motor preparation from the relationship between the start points of N1, N2, and the responses of EMG, and the amount of steering the wheel with the time lag.

Keywords: ERP, motor preparatory activity, manual control, time lag, EMG

1 INTRODUCTION

There is a system of handling an object with a time lag such as steering large-size ships. Although such system recently advances to control an object with automation instead of manual control, it is not possible to apply the automatic control to all the scenes, for example, an unusual case such as the system does not work properly. It is important to investigate the property of the operator when manually handling the object with a time lag in the fields of human factors and ergonomics. Most of the previous studies have been reported to examine the property of the operator in the performance of a certain task for the system with a time lag (Hayashi (1967), Yukimachi (1973), Hayashi (1975)). On the other hand, there are few studies to evaluate manual control with the measurement of the physiological responses while the operator performs the task in the system with a time lag. In this paper, we shed a light to study the activities from the brain to motor control system of the operator when manually handling the object with a time lag. In the measurement of physiological responses, an electroencephalogram (EEG) and electromyogram (EMG) of right arm of the operator were recorded during the task.

2 PURPOSE

The purpose is to evaluate the effect to the activities between brain and motor control of the operator with the time lag for the manual control system using the measurements of physiological responses (EEG and EMG) during the task. In particularly, it is examined the relationship between the Event-Related Potential (ERP) and the changes of EMG for the right hand of observers when steering a wheel to handle an object through the gate in the task.

3 METHODOLOGY

3.1 *Measurement of EEG*

An electroencephalogram (EEG, a Nihon Kohden EEG1100) was recorded with a disk electrode of silver/silver chloride arrayed at the five points, Fz, Cz, Pz, C3, and C4 (see Fig. 1) in accordance with the 10/20 electrode system and referenced to linked earlobes. The EEG was measured using a high-frequency filter with cut-off at 30 Hz, a low-frequency filter with cut-off at 0.03 Hz, and a sample rate of 1000 Hz. In addition, the electrooculogram (EOG) for the right

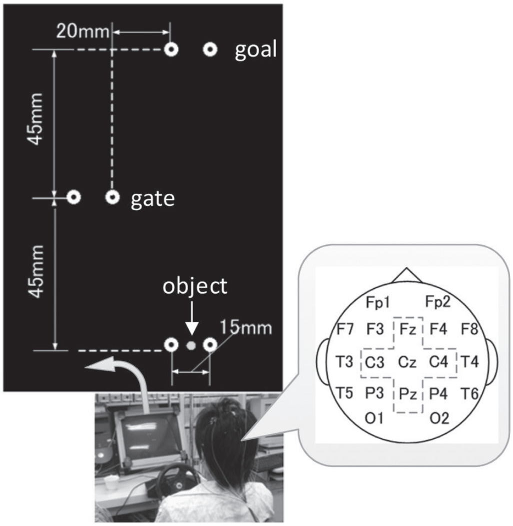


Figure 1. Experimental scene. Top: course of the task. Bottom: experimental scene. Right: measurement points.

eye was recorded to monitor any artifacts such as blinking into the EEG.

3.2 *Measurement of EMG*

An electromyogram (EMG) was recorded with a disk electrode of silver/silver chloride arrayed at the three points of the muscles in forearm, arm, and shoulder for right hand of the observer. The EEG was measured using a high-frequency filter with cut-off at 120 Hz, a low-frequency filter with cut-off at 10 Hz, and a sample rate of 1000 Hz.

3.3 *Task*

In this experiment, it was used the simulator of the system with the time lag in the first order for steering the wheel. The task was to steer an object to a goal through the gate. Observers were instructed to guide the object on a CRT screen with a steering wheel to the goal through the gate. There is one gate to the goal as shown in Fig. 1. The dynamic characteristics were set a first order lag. Each observer handled the object with the three conditions of a time constant (τ) of zero, three, or six seconds lag in the response to the steering wheel input. The object went forward to the top of the screen with the velocity of 4 mm/s in constant. The angle of steering the wheel was recorded by the computer in transferring into the amount of voltage.

They were required to repeat this task for 200 times of each condition. The total was 600 times repeated randomly for two days taking rests.

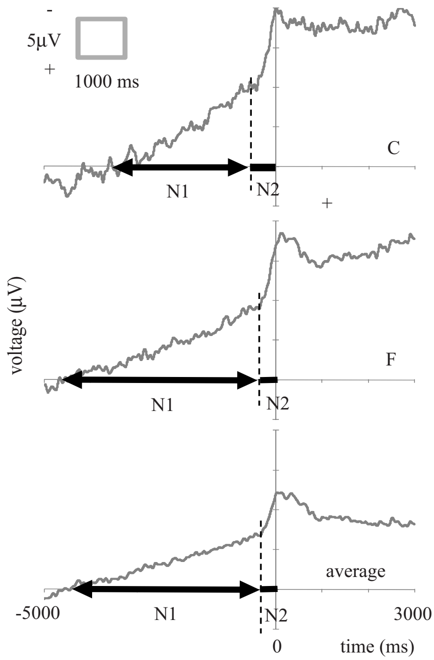


Figure 2. ERP of Cz, $\tau = 3$ s for observer C, F, and the average result in ten observes. Top: observer C, Middle: observer F, Bottom: average.

3.4 *Observers*

Written and informed consent was obtained from ten male observers in age with an average of 23.5 years and an SD of 1.84. Prior to the experiment, observers had an enough time to practice for this task.

4 RESULTS AND DISCUSSIONS

4.1 *ERP in steering the wheel*

Figure 2 shows the results of ERP at Cz, $\tau = 3$ s when steering the wheel to the right to through the gate. Each graph corresponds to the results of observer C, F, and average based on the waveforms of all the observers, respectively. The waveforms of observer C and F were obtained by averaging the waveform of all the measurements for each observer except in the cases of such as blinking the eye. The horizontal axis indicates the time (ms) for eight seconds and the zero means the point when steering the wheel to the right. The vertical axis indicates the amplitude (μV) and the mean potential of 5000 ms to 4000 ms before steering the wheel was taken as the baseline.

As shown in this figure, there are the negative slopes of the potential before steering the wheel for all the graphs. At first, the potentials rise gradually negative (N1) and change rapidly (N2) to the peak at zero in time. It is considered to be able to divide into two stages of N1 and N2 in terms of the motor preparatory activity. It is the same tendency at the others.

4.2 *Relationship between ERP and EMG in motor preparatory activity*

Figure 3 shows the results of ERP at C4, $\tau = 3$ s for observer C, EMG of the muscle in forearm for the right hand, and the amount of voltage with steering the wheel. The top, middle, and bottom graphs correspond to the waveforms of ERP, EMG, and amount of voltage with steering input, respectively. The horizontal axis is the same as Fig. 2. The vertical axis indicate the amplitude (μV or mV) depended on each graph.

As shown in this figure, before the potential of the steering input rises up in bottom graph, the potential of the muscle in forearm changes rapidly (M) in middle graph. The start point of N2 for ERP in top graph is roughly same as M in the potential change for the forearm in right hand. It is

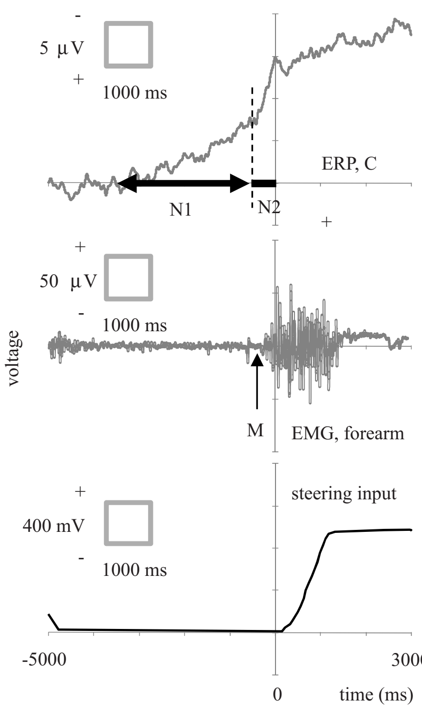


Figure 3. ERP of C4, $\tau = 3$ s for observer C, EMG of the muscle in forearm of right hand, the amount of the voltage with the steering input. Top: ERP, Middle: EMG, Bottom: steering input.

suggested that N2 means the activity to exercise the muscle in motor preparation of human body.

4.3 *Effect of the time lag for ERP*

Figure 4 shows the results of ERP at Cz based on the waveforms of all the observers. Each line represents the results of each time lag condition, zero, three, and six seconds, respectively. The horizontal and vertical axes are the same as Fig. 2.

As the results of Fig. 4, the start point of N1 and the peak of the potentials are different from each condition. It seems to rise earlier and increase in peak with increasing of the time lag in conditions.

4.4 *Underlying mechanism for N1, N2, and peak of ERP in motor preparatory activity*

To obtain the start points of N1 and N2 and the peak in ERP, each stage of N1 and N2 in the waveforms was fitted by two straight lines using least squares method according to the previous study (Kornhuber and Deecke (1965)). Fig. 5 shows the results of start points of N1 and N2 (N1 start and N2 start) and the peak at Cz based on the results of all the observers. Each graph corresponds to the results of N1 start, N2 start, and the peak of ERP, respectively. Each bar represents each condition of the time constant in time lag. The vertical axis indicates the time (ms) and amplitude (μV), respectively.

In Figure 5, the times of N1 and N2 start are faster with increasing the time constant. Also, the peak of the potential in negative slope increases with the time constant as the mention above. It is similar tendency in all the measurement points.

Figure 6 shows the comparison of ERP at Cz based on the waveforms of all the observers and the amount of voltage with the steering input for three conditions of the time lag. The top and bottom graphs correspond to the ERP of average results and the change of voltage with the steering

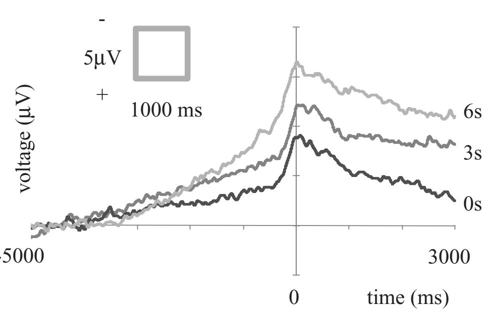


Figure 4. ERP at Cz based on the waveforms of all the observers.

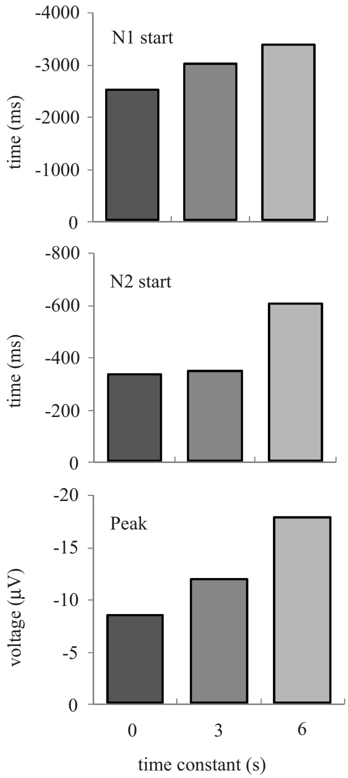


Figure 5. N1 start, N2 start, and peak at Cz. Top: N1 start, Middle: N2 start, Bottom: peak.

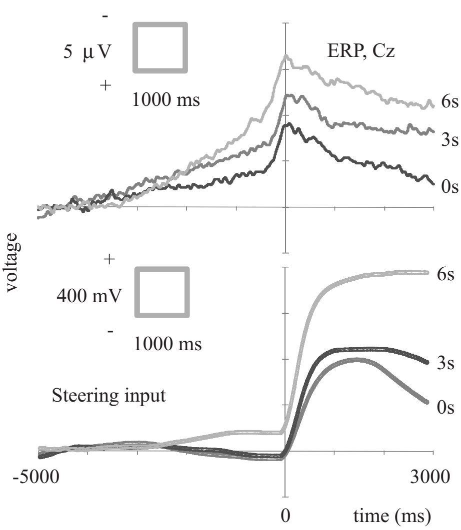


Figure 6. Comparison of ERP at Cz and the steering input in each time constant. Top: ERP at Cz in the average results, Bottom: voltage with steering input in each time constant.

input. The horizontal and vertical axes are similar to Fig. 3.

As shown in this figure, the amount of the voltage with the input increases with the time constant in time lag. According to the increasing the amount of the voltage with the input in this figure and the results in Fig. 5, it seems that the time of N1 start is faster and the peak of the negative slope in ERP increases. It is suggested that N1 and the peak means the preparatory activity to steer the wheel and the activity to steering the wheel. In the previous study of Shibasaki et al. (1980), they measured the ERP when moving the fingers. The results were the similar tendency that there are two stages before the movement. In their consideration, one is the preparation of motor control such as N1, the other is the activity to exercise the muscle in motor preparation such as N2.

5 CONCLUSION

To examine the relationship between the Event-Related Potential (ERP) and the changes of EMG in the right hand of observers, it was measured the physiological responses when steering a wheel to handle the object through the gate as the task. It was obtained the phenomenon that the activity in brain occurs earlier to exercise the muscle in motor control with increasing the time lag of the system in this experiment. It is considered to need the strong activity in early stage from the relationship between N1 start, the peak of negative slope in ERP and the value of time lag in the system. From the results of N2 start with the time lag, it seems that brain orders to act earlier for exercising the muscle in motor control to correspond to the time lag of the system.

REFERENCES

Hayashi, Y. 1967. Human factors on manual control. *The Japanese Journal of Ergonomics* 3(4): 266–274. (in Japanese).

Hayashi, Y. 1975. On the human tracking characteristics of previews control. *The Japanese Journal of Ergonomics* 11(2.3): 69–70. (in Japanese).

Kornhuber, H.H. and Deecke, L. 1965. Changes in the brain potential in voluntary movements and passive movements in man: Readiness potentials and reafferent potentials. *Pflugers. Arch. Gesamte.Physiol. Menscchen Tiere* 10: 284: 1–17.

Shibasaki, H., Barrett, G., Halliday, E. and Halliday, A.M. 1980. Components of the movement-related cortical potential and their scalp topography. *Electro-encephalography and Clinical Neurophysiology* 49: 213–226.

Yukimachi, T. 1973. Manual preview control and multivariate information. *The Japanese Journal of Ergonomics* 9(4): 163–169. (in Japanses).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Assessment of customers' perceptions toward attributes of bedroom space

Yi-Lin Chen & Szu-Yu Tzeng
Department of Architecture and Interior Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

Chi-Ming Hsieh
Department of Community, Agriculture, Recreation & Resource Studies Michigan State University, Michigan, USA

ABSTRACT: Home is a place where people can pursue a carefree way of living. Bedroom aesthetics is about allowing customers an enjoyable experience when using the bedroom. Aesthetics is not the only thing that interior design is involved with; instead, aesthetics must also build upon functionalism and utilitarianism. A design is successful when it gains an insight into customers' needs and thus can be realized and completed. The most important stage in this process is to build a communication platform. Through this platform, designers can adequately communicate and reach a consensus with their customers, so that they can build a home space satisfying customers' needs. This study adopts questionnaire survey and develops a "bedroom spatial attribute evaluation scale". This scale selects adjectives used by general customers when describing attributes of a bedroom, so that it projects bedroom images in customers' perceptions. After being tested, this scale is proved to meet validity and reliability. Therefore, it can be used as a tool to construct bedroom spatial attributes and allow designers to effectively grasp customers' demands for bedroom design. Finally, this study suggests on future research and practical applications.

Keywords: bedroom space, attributes, evaluation, scale development, confirmatory factor analysis

1 RESEARCH BACKGROUND AND OBJECTIVES

People spend a third of their life time in sleep. The quality of sleep affects people physically, mentally, and emotionally. After a tiring day at work, people expect to return home to an environment where they can really relax and rest. Bedroom is exactly a place for resting and sleeping. Hence, privacy, independence, and quality sleep are priorities for a good bedroom. Regarding this, bedroom is a space with infinite possibilities. With owners' personalities, preferences, and demands as premises, the making of a bedroom should strike a balance between aesthetics and functions and create an atmosphere to fit its functionality. People of different personality traits, occupations, genders, and ages have their own ideas about bedroom space. With the help of different types of bedroom and even cognitive psychology, designers can understand customers' expectations and preferences. Therefore, it is expected that interior design can become more academic and professional. Based on the above-mentioned motives, the purposes of this study include the following:

1. Explore related theories of bedroom spatial attributes
2. Collect attributes of bedroom design
3. Construct an evaluation model for bedroom spatial attributes
4. Provide this model as a reference to interior designers.

2 LITERATURE REVIEW

Home life should be an expression of aesthetics. The creation of atmosphere and setting of a bedroom as the soul of a home is above all the most important. Bedroom is the most relaxing and private place for people to sleep. The style of a bedroom varies in accordance with different users' requirements. Since the 1990s, numerous domestic and international scholars (Chin, 2001; Chang, 2004; Deng, 2002; Scott, 1993; Yang, 2003; Zhuang, 2008) have started research on the evaluation of spatial attributes to assess the connection between spatial aesthetics and people's preferences. Chin (2001) categorized consumers' images, descriptions, and perceptions of space into adjectives of young, bright, brand-reflective, fashionable, and clean. The categorization allowed consumers to describe settings and styles of space with various adjectives. Deng (2002) used adjectives on spatial design of restaurant space. He designed restaurants with installations of world's famous landmarks, bridges, streets, or pictures of natural landscapes, playgrounds, snow scenes, windmills, glacial erosion lakes, etc., These designs inspired consumers to describe a space with adjectives such as splendid, classical, gracious, interesting, exciting, passionate, stony, natural, and restricted. These verbal descriptions immerse consumers into the space and allow them a different way to perceive the space.

Yang (2003) used the public's images about the forms of sculptures to categorize styles of outdoor sculptures in public space. She also collected adjectives suitable for describing sculptures and conducted a survey on sculpture images. Chang (2004: 19–20) applied the theory of Mehrabian and Russel (1974: 233–252) and used the three elements of pleasure, arouse, and dominance to reflect human emotions. His semantic differential scale used opposite terms like "pleasant—unpleasant," "excited—bored," and "relaxed—nervous" to evaluate people's emotional reactions to environment and explore their preferences in terms of environmental aesthetics. Zhuang (2008) explored terms used to describe cultural signs and patterns of Taiwanese temples. With patterns which were described as "happy," "safe and well," and "auspicious." Zhuang (2008) provided inspirations to product design. In addition, he concretely presented traditional signs and creative associations to improve product design and consumers' recognition of the design. The atmosphere of a bedroom arises from its functions. If the design of a bedroom is integrated with emotional factors, a designer can use colors and lights to create a variety of styles. Zhang (2007) believed with the considerations of users' perceptions and demands for quality, bedroom design should include the following features: (1) the integration of outline and framework, (2) the preservation of natural and original style, and (3) the psychological influence of shapes like circles, curves, and rectangular on users. Regarding this, the dimensions of bedroom space are vast, its atmosphere is homey, its conversation is about humanities, and its spirit is being practical. Interior design is not just about decorating. Instead, it should interpret every space perfectly to allow users to sleep in peace. This is the original intention of design, and by doing so, a designer can truly present the atmosphere of a space.

With regard to the descriptions above, it is obvious that bedroom design is like putting together stylish outfits, where every person has his or her own unique style and taste. A bedroom too speaks for its owner's taste and style. Exploring spatial images and preferences with adjectives is a suitable method for the studies of interior design. Bedroom design involves not only the designer but also people living in it, who are the main roles to the bedroom. In addition to the presentation of a space, discovering consumers' preferences is equally significant. Some related researches adopt factor analysis approach to grasp consumers' psychological perceptions. For further considerations of the public's preference and consumer-oriented design, however, there is a lack of a valid and reliable scale due to the highly subjectivity and short of an evaluation approach (Aaker, 1981). Therefore, this study applies confirmatory factor analysis approach to establish a valid and reliable evaluation. Additionally, it uses a measurable scale as the measuring tool, so that designers can interpret consumers' subjective concepts of values about styles and aesthetics. By doing so, the designers can reach a consensus with the consumers and reduce the risk of a cognitive gap and poor communication.

3 METHODS

The construction of the bedroom spatial attribute evaluation scale was divided into four stages. In the first stage, experts selected representative bedroom designs whose types and styles were suitable to the experiment. These selected designs served as stimulants for the projection experiment. In the second stage, adjectives suitable for describing bedroom space were selected during the interview with consumers. This study interviewed non-professional consumers in an attempt to choose general and comprehensible attributive adjectives. In the third stage, the bedroom spatial attribute evaluation was constructed and its aspects were completed. In the last stage, this study carried out confirmatory factor analysis to build a reliability and validity index for the evaluation scale and to examine the theoretical and practical values of the studies of bedroom design.

In the first stage, this study collected existing literatures, magazines of interior design, and published or unpublished works of existing interior designers in the past decade as representative samples for the experiment. By excluding the consideration of levels and costs of these designs, this study sought to find a collection which covered preferences of as many consumers as possible. Initially, 128 designs were collected. After being screened by three experts, the sample size of the stimulants was downsized to 14 types of bedroom design (Table 1). These descriptions were then represented with projected images, which were put into the questionnaire survey for the respondents.

In the second stage, through the projection experiment, this study selected attributive adjectives used by the respondents in response to the stimulants. During the experiment, 10 respondents with different educational backgrounds and life experiences were asked to contribute 2 to 5 adjectives to each of the 14 design cases in an open-end questionnaire. Considering the uneven verbal abilities among the respondents, the questionnaire provided terms as a reference to the respondents but only when they required for it. With the help of the stimulants, this study could understand consumers' subjective judgments and descriptions of the attributes of bedroom space. When analyzing the results of the questionnaire survey, this study selected attributive adjectives which had been used for 10 times or more to be the evaluative items for this research. These adjectives were (01) romantic, (02) steady, (03) simple, (04) refreshing, (05) elegant, (06) passionate, (07) mysterious, (08) bright, (09) homey, (10) plain, (11) popular, (12) comfortable, (13) retro, (14) benevolent, (15) classical, (16) vivid, (17) gentle, (18) natural, (19) dreamy, and (20) gorgeous.

In the third stage, a scale ranging from 0 to 10 marks was assigned to each of the 20 attributive adjectives selected in the second stage. Therefore, there were a total of 11 marks for each adjective. The more the marks a respondent gave to an adjective, the more agreeable she was to it. Conversely, the less agreeable the respondent was to the adjective. The scale was put to the test upon the completion of its construction. After the test results were collected, they were analyzed with SPSS for windows 12.0.

In the fourth stage, this study applied AMOS 17.0 for confirmatory factor analysis. By examining the significance of the path coefficients of the variables, it could verify its hypothesis. Structural equation modeling was used to explore the causal relations among variables and verify the goodness-of-fit of the observations. By observing the connections among variables, this study corrected the model until the best fitness was obtained.

3.1 *Research samples*

This study adopted sampling approach and conducted a questionnaire survey about bedroom spatial attributes on general public. A total of 155 questionnaires were distributed during the survey. After incomplete questionnaires were excluded, 140 valid ones were collected, marking a return rate of 90%. Among the respondents, there were 30.4% (63 persons) males and 37.2% (77 persons) females.

Table 1. 14 types of bedroom design.

Figure 01

Figure 02

Figure 03

Figure 04

Figure 05

Figure 06

Figure 07

Figure 08

Figure 09

Figure 10

Figure 11

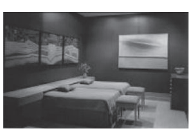


Figure 12

Figure 13

Figure 14

Most of the respondents were 21 to 30 years old, which accounted for 32.9% of the valid samples. The remaining age groups were 31 to 40 years old (23.7%), 41 to 50 years old (6.8%), and 50 years old and over (4.3%). 33.3% of the respondents had an educational background of college (or junior college). 16.9% had an educational background of senior high school (vocational school) or below, 14.5% had a master's degree, and 2.9% had a doctoral degree or were still studying for it. As many as 20.3% of the respondents worked in agriculture, industry, business, electronic industry, and service industry. 17.4% were students majoring in arts or designs. 12.1% served in the military, public sectors, or education. 6.3% were housewives. 1.9% were students majoring in subjects other than arts or designs. 4.3% were unemployed, retired, freelancers, or otherwise.

3.2 *Item analysis*

According to Qiu (2006), item analysis is the most critical to scales or questionnaire preset. The purpose of item analysis is to figure out item difficulty and discriminability of a proposed scale. This study used SPSS12.0 for item analysis. The testing methods were missing value estimation method, descriptive statistics (means, standard deviation, and coefficient of skewness), and homogeneity test (correlation coefficients and factor loadings). Items failed to achieve significance did not have discriminability, and thus could not reflect the responses of different respondents. Regarding this, insignificant items were removed. The mean of this scale was 4.33, and the standard deviation was 1.37. When the figure exceeded two positive or negative standard deviation, an item was deleted (standard deviation larger than 7.07 or smaller than 3.59 showed low discriminability). The means of the scale ranged from 4.04 to 5.97, meaning that none of the means was too large or too small. The top 27% was the group with high marks, and the bottom 27% was the group with low marks. The two groups were used for t test to obtain the critical value. According to Qiu (2006), CR value larger than 3.0 is considered acceptable. Based on these criteria, all items of this scale achieved the standard. Additionally, the results of testing for mean differences all achieved significance. Correlation coefficients of all items exceeded 0.4, meaning that all items met the standard of discriminability. A more precise analysis revealed the results of respondents' factor loadings and internal homogeneity. All items in the scale were with factor loadings larger than 0.3, and the attributions were highly relevant. The Cronbach's α of the scale as a whole was 0.97, which fell within the range of high reliability.

4 DATA ANALYSIS

4.1 *Exploratory factor analysis*

This scale used exploratory factor analysis to establish factors for bedroom space. The communality demonstrated the parts of observable variances which could be explained. The higher the communality was, the lower the uniqueness. KMO sampling adequacy was 0.95, and Bartlett's Test of Sphericity showed that this scale had good sampling adequacy. Principal component analysis and promax method in oblique rotation kept items with factor loadings greater than 0.3. With the help of pattern matrix and structure matrix, four factors were established. The following shows the naming of the factors, which was based on the values of factor loading matrix.

Variances of the exploratory factor analysis showed that the first variance explained was 22.80%, the second was 20.06%, the third was 18.30%, and the fourth was 15.59%. The cumulative variances of the four factors were 76.76%. In other words, the percentage of these four exploratory factors to "bedroom spatial attribute evaluation scale" was 76.76%. According to the reliability analysis of the 20 items under the four factors, the Cronbach α of the first factor was 0.921, the Cronbach α of the second was 0.921, the Cronbach α of the third was 0.883, and the Cronbach α of the fourth was 0.876. The Cronbach α of this scale as a whole was 0.966, which fell within the range of high reliability. The 20 items of the bedroom spatial attribute evaluation scale were categorized into the four factors respectively.

1. The first factor covered items (03) simple, (10) plain, (04) refreshing, (02) steady, (17) gentle, and (08) bright. Homes with these descriptions are simple, plain, refreshing, steady, gentle, and bright, and such homes present a feeling of purity and simplicity. Therefore, this factor is named "pure and plain style."
2. The second factor covered items (01) romantic, (05) elegant, (09) homey, (19) dreamy, (20) gorgeous, and (12) comfortable. Homes with these descriptions are romantic, elegant, homey, dreamy, gorgeous, and comfortable, and such homes present a romantic and warm atmosphere. Therefore, this factor is named "luxurious and elegant style."
3. The third factor covered items (16) vivid, (11) popular, (18) natural, (14) benevolent, and (06) passionate. Homes with these descriptions are vivid, popular, natural, benevolent, and passionate. Such homes give a mixed feeling of modern, natural, passionate and friendly. Therefore, this factor is named "fashionable and natural style."

4. The fourth factor covered (13) retro, (15) classical, and (07) mysterious. Homes with these descriptions present a classical and mysterious atmosphere. Therefore, this structure is named "retro and classical style."

4.2 *Confirmatory factor analysis*

The major function of Confirmatory Factor Analysis (CFA) is to verify factor validity and factor correctness of a proposed model. Such an analysis approach is suitable for theories waited to be verified or empirical research (Byrne, 2001; Qiu, 2003). In his research of education and psychology, Li (2006) proposed that reliability and validity of a measuring tool could be verified with the help of CFA. When using CFA, a researcher can define some constraints based on the theories or hypotheses of his/her model. The researcher can decide correlations among some common factors, how some observable variances will be affected by other common factors, which observable variances are unique factors, and which unique factors correlate to one another (Cheng, 1993; Doll, Raghunathan, Lim & Gupta, 1995).

AMOS 17.0 can test the significance of path coefficients in a structural equation model and verify the causal relation of the model. Generally, it is suggested that linear structural relation analysis be used for larger sample sizes to obtain better results (Qiu, 2003). There were 140 valid samples in this study. Therefore, it could use maximum likelihood estimation and hypotheses of multivariate normality to estimate the coefficients and fit of its model (Ding, Velicer, & Harlow, 1995).

The 20 items of the bedroom spatial attribute evaluation scale were designed into a path model for first order confirmatory factor analysis. The retrieved index served as a standard to test model fit. The result showed that χ^2 was 2.7305976, RMSEA was 0.112, RMR was 0.123, SRMR was 0.069, GFI was 0.74, AGFI was 0.67, NFI was 0.832, RFI was 0.81, IFI was 0.89, NNFI was 0.87, CFI was 0.885, PGFI was 0.580, PNFI was 0.718, and PCFI was 0.76. In the hypothetic model, the values of (χ^2/df), RMSEA, GFI, SRMR, AGFI, NFI, RFI, IFI, NNFI or TLI, and CFI were not ideal.

According to Bagozzi and Yi (1988), when MI value exceeds 3.84, there is a need for correction (Cheng, 2003). Through MI, covariance of error, and factor loadings, this study corrected parameters and removed items with residual values larger than 2 or smaller than negative 2. In the end, 6 items were removed. (Figure 1) F1 eliminated v3—simple, v4—refreshing; F2 eliminated v5—elegant, v20—gorgeous; F3 eliminated v16—vivid; F4 eliminated v7—mysterious. Table 2

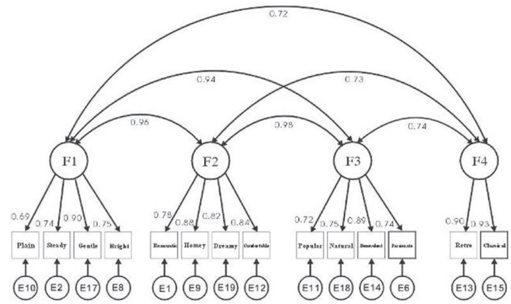


Figure 1. The best fitting model.

Table 2. Measures of the best structural model fit.

Item	Indices fit	Criteria
Chi-Square/Chi-Square DF	2.731	<3
(GFI)	0.89	>0.9
(AGFI)	0.83	>0.8
(CFI)	0.963	>0.9
(NFI)	0.92	>0.9
(RMSEA)	0.08	<0.08

Source: This study.

demonstrates the model fit after the correction. Based on absolute fit index, except for the chi square value, the rest of the values of GFI (0.89), RMSEA (0.08), RMR (0.069), and SRMR (0.038) were all acceptable. GFI and AGFI fell within the range between 0.80 and 0.89, which were also reasonable. Models with values equal to or exceed 0.90 are regarded to have better fit (Joreskog and Sorbom, 1989). Based on absolute fit measures, the model of this study as a whole was acceptable. Based on incremental fit measures, NFI (0.921), RFI (0.90), IFI (0.96), NNFI (0.95), and CFI (0.963) either equaled to or exceed 0.90, meaning that the model as a whole was acceptable. Based on parsimonious fit measure, PNFI (0.72), PCFI (0.75), and PGFI (0.599) all exceed 0.50, showing that they were also acceptable. Generally, the model of this study was acceptable.

Composite reliability and AVE were applied to explain the four factors. The CR of the first factor of "pure and plain style" was 0.86, and AVE was 0.6; the CR of the second factor of "luxurious and elegant style" was 0.9, and the AVE was 0.7; the CR of the third factor of "fashionable and natural style" was 0.86, and AVE was 0.6; and the CR of the fourth factor of "retro and classical style" was 0.91, and AVE was 0.8.

This study used other model fit indices of Amos 17.0 to correct the model. The factor loadings of the model all achieved significance and fell within the range of an ideal model fit. In terms of the internal structure fit of the "bedroom spatial attribute evaluation scale," the reliability of each item (similar to R^2 of regression coefficients) was 0.69, 0.74, 0.90, 0.76, 0.78, 0.88, 0.83, 0.83, 0.72, 0.75, 0.89, 0.74, 0.90, and 0.93 respectively. The result of internal quality showed the model and information fit of the bedroom spatial attribute evaluation scale was acceptable.

5 CONCLUSIONS AND FUTURE SUGGESTIONS

By referring to domestic and international literatures of interior design, this study built the "bedroom spatial attribute evaluation scale" with 14 items, including romantic, steady, passionate, bright, homey, plain, popular, comfortable, retro, benevolent, classical, gentle, natural, and dreamy. Through factor analysis approach, the four factors of "pure and plain style," "luxurious and elegant style," "fashionable and natural style," and "retro and classical style" explained the items in the bedroom spatial attribute evaluation scale. The Cronbach's α of the scale was between 0.88 and 0.92. CFA examined the convergence of each of the factor. The result showed that the values of (χ^2/df), RMSEA, GFI, SRMR, AGFI, NFI, RFI, IFI, NNFI or TLI, and CFI did not achieve good fit. However, when examining MI, covariance of error, and factor loading for parameter correction, absolute fit measure showed that RMSEA (0.08) was not less than or equal to 0.05, GFI (0.89) was not larger than or equal to 0.90, and RMR (0.069) was not less than or equal to 0.05. It is suggested that Chi square (χ^2/df) should be equal to or less than 3 (Kline, 2005). Additionally, in absolute fit measure, it is suggested that RMSEA should be equal to or less than 0.08 (Browne & Cudeck, 1993), SRMR should be equal to or less than 0.08 (Hu & Bentler, 1999), and CFI and NNFI should be 0.90 and 0.95 respectively (Hoyle & Panter, 1995; Hu & Bentler, 1999). Based on the criteria of model fit proposed by the scholars, the model fit of this study was acceptable ($\chi^2 = 126.596$, df = 71, $p < 0.001$, RMSEA = 0.08, SRMR = 0.0387, CFI = 0.963, NNFI = 0.95).

After modifications, the "bedroom spatial attribute evaluation scale" covered four factors. Factor one "pure and plain style" included items (10) plain, (02) steady, (17) gentle, and (08) bright; factor two "luxurious and elegant style" included items (01) romantic, (09) homey, (19) dreamy, and (12) comfortable; factor three "fashionable and natural style" contained items (11) popular, (18) natural, (14) benevolent, and (06) passionate; and factor four "retro and classical style" included items (13) retro and (15) classical. For practical applications, the "bedroom spatial attribute evaluation scale" was transferred into an evaluation index to measure people's satisfaction about a bedroom space. It is expected that this evaluation scale can serve as a solution for interior design and a reference for future research. It should be notified that, however, whether this scale can be used in a wide variety of applications depends on different situations and requirements. It is suggested that future research adjust this evaluation scale to cater to different bedroom settings. Given that factors of a model are determinants in terms of model fit, removal of any of the items in this evaluation scale requires thorough consideration. It is a challenge for a good scale to cover all essential items. Regarding this, it is suggested that future research carry out in-depth studies.

REFERENCES

Browne, M.W. & Cudek, R. 1993. Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models*: 136–162. Newbury Park, CA: Sage.

Chang, H.P. 2004. *The relationship between aesthetic factors and aesthetic responses in indoor environments: Using living room as an example*. Chung Yuan Christian University, Interior Design master's thesis.

Cheng, B.L. 1993. Second-order confirmatory factor analysis of the theory and applications. *Journal of Education and Psychology* 16: 23–82.

Chin, H.C. 2001. *A study of the relationship between spatial images and boutique design features*. Chung Yuan Christian University, Interior Design master's thesis.

Deng, J.H. 2002. Symbol, image, spectacle—Taiwan Food Culture pedigree (above), Garden City culture, Ltd: 64.

Doll, W.J., Raghunathan, T.S., Lim, J.S. & Gupta, Y.P. 1995. A confirmatory factor analysis of the user information satisfaction instrument. *Information Systems Research 6*(2): 177–189.

Jöreskog, K.G. & Sörbom, D. 1989. *LISREL 7: A guide to program and applications* (2nd ed.). Chicago: SPSS.

Kline, R.B. 2005. *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Qiu, H.Z. 2006. *Quantitative research and statistical in social & behavioral sciences*. Taipei: Wunan book.

Suzanne, C.S. 1993. *Visual attributes related to preference in interior environments*. University of Wisconsin Madison, Wisconsin.

Yang, Y.J. 2003. *A study into images of form of outdoor sculptures in public space*. Nanhua University, Department of Aesthetics and Visual Arts, master's thesis.

Zhang, Z.Z. 2007. *Furniture Viewpoint*. DECO Chi, Oct. No. 69: 182–183.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Visibility evaluation for vehicle interior ergonomics based on augmented reality technology

Lifeng Zhang
CIM Research Institute, School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai, P.R. China

Xiumin Fan
CIM Research Institute, School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai, P.R. China
Shanghai Key Lab of Advanced Manufacturing Environment, Shanghai, P.R. China

Xu Jing & Shiguang Qiu
CIM Research Institute, School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai, P.R. China

Huanchong Cheng
CIM Research Institute, School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai, P.R. China
Shanghai Key Lab of Advanced Manufacturing Environment, Shanghai, P.R. China

ABSTRACT: In the process of product development for automotive vehicle, human-machine ergonomics assessment is an important factor which has influence on the product development cycle time and cost, and it is really an important factor for the user. In order to provide a quick and human-in-loop visibility evaluation method for style designer when designing vehicle interior, an ergonomics assessments system based on augmented reality technology is proposed for vehicle interior design evaluation in conceptual design phase. Some key technologies are discussed, such as, camera & optical tracking mixed registration for real and virtual objects. A set of preliminary hardware and software prototype system for ergonomics evaluation is developed, and data is collected to calculate blind zone for A-pillar based on Eye Ellipse Method to verify the function of this prototype system. The result shows the prototype system can act as an aided tool for vehicle interior ergonomics evaluation.

Keywords: vehicle interior design, ergonomics, visibility evaluation, augmented reality, vision & optical hybrid registration

1 INTRODUCTION

For nowadays product development, human-machine efficacy design has become a key factor for vehicle customers' acceptance and a critical part of automobile design. It not only helps to reduce accidence caused by drivers' fatigue, but also improves vehicles' safety, comfort, and convenience of communication system and meets the user's individual needs (Milosevlc 1997).

Traditional analysis of vehicle ergonomics design is through repeated trial of prototype, until it meets with designers' criteria. However, establishment and modification of physical prototype is time-consuming. As a result, the cost is high and it cannot catch up with the pace of vehicle updating. With the appearance of virtual reality technology, human-machine ergonomic evaluation system, based on 3D virtual modeling of human body, began to emerge in industry. Through the interactions between virtual models of human body and vehicle, ergonomics evaluation can be conducted. Even though a lot of time is used to accomplish the realistic expression of this system, the virtual human models still fails in fully expression of real human and the evaluation system cannot cover all aspects of a real driving procedure (Geuss 1998; Ekern et al. 1997).

Augmented Reality (AR), which is a variation of Virtual Reality (VR), is a computer application and human-machine interaction technology (Azuma 1997). It was developed on the basis of computer vision and graphics, sensor, network and GPS technologies. It is accomplished through superimposing virtual objects upon the real world, using photoelectric display, interactive and visualization technology. By sensing and vision registration method, the virtual object is precisely 'put' in real environment and the resulting augmented scene combined both virtual and real objects is presented to users on display (Azuma et al. 2001; Feiner et al.

2004). Many AR application systems were developed with ARToolkit (Kato et al. 1999) library. The ARToolkit used vision techniques to obtain the position and orientation of a camera relative to marker for real-virtual world registration.

Augmented reality technology can achieve precise and real-time superimposing of virtual vehicle model to the real vehicle prototype. Designers can see the real scenario after this synthesis and carry on the evaluation through interaction with the virtual models. In this way, the development cycle and cost can be reduced. Augmented reality technology to support the automobile development was also mentioned in (Fründ et al. 2004), and the author has also developed mobile augmented reality testing platform cooperated with Volkswagen.

This paper is focused on visibility analysis based on Eye Ellipse Method in a static augmented reality environment. The rest section of this paper is organized as follows. Section 2 is about principle for vehicle interior visibility evaluation based on Eye Ellipse Method. Section 3 is about augmented reality system structure for ergonomic evaluation and key technology. Section 4 is case study about head movement data collection and blind zone calculation for a testing operator. Section 5 is conclusion and further research work.

2 METHOD FOR VEHICLE INTERIOR VISIBILITY EVALUATION

There are three kinds of field of view visibility evaluation; the first one is front view, such as view evaluation for front road signs and billboards; the second one is indoor view of the vehicle interior parts, the third one is rear view through rearview mirrors.

This paper focuses mainly on the visibility evaluation of A-pillar's influences on the front view. The influence evaluation of A-pillar is carried on first by drawing the eye ellipse, as shown in Figure 1, and then calculating its position according to the location of car seat, thereby assessing the A-pillar by graphing (Yiming et al. 1999).

As shown in Figure 2, the evaluation process for vehicle A-pillar through the eyes ellipse drawing is as follows:

1. On the front-view of eye ellipse, draw the projection profile of A-pillar. On the top-view, draw the sectional view of A-pillar.
2. On the top-view, find out point B on eye ellipse which has the shortest distance to A-pillar. From point B, draw a tangent line B' A1 to A-pillar, point A1 is the tangential point. Make point B projection to the major-axis of eye ellipse on front-view, and obtain point B1.

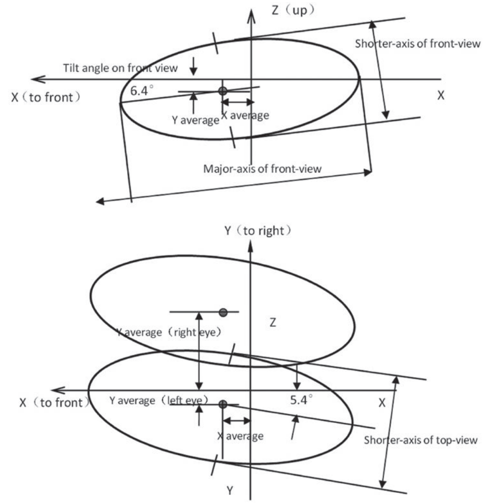


Figure 1. Drawing of eye ellipse.

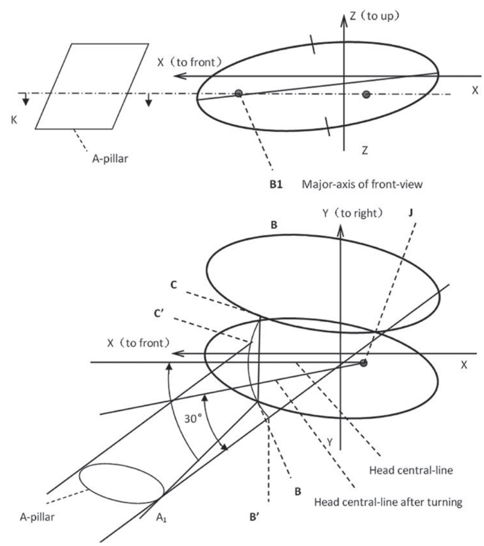


Figure 2. A-pillar analysis based on eye eclipse.

3. On top view, find out the right eye point C, BC represents the IPD, usually its value is 65 mm. Through the BC, make the perpendicular bisector, away from the center cross point to 98.6 mm, get the head rotation center, point J.
4. Since the A pillar is located 30° outside of right and left eye' front sight line, it means beyond the eye's turning angle 30°, the driver has to turn his head to make his sight line towards A-pillar. At the same time, point B and C will turn alone with, until the angle between head central-line and line B'A1 turns to 30°.

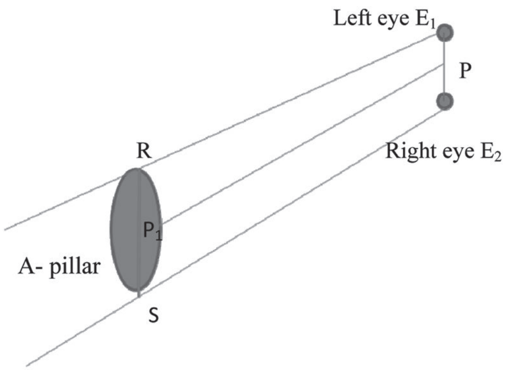


Figure 3. Calculation of the two eyes' blind angle.

5. On top view, from the point C', draw sight line CE, point E is the tangential point. The angle between the sight line C'E and sight line B'A1 is the A pillar blind zone of the two eyes.

The principle mentioned above is from statistic point of view, while for a specific driver, a simplified method can be used, as shown in Figure 3. In the evaluation process of measuring the relative position of the mid-point of human eyes P to the center point of the cross-section of A-pillar, P1, one can establish the equivalent line of the blocked area by A-pillar, RS, then measured its length in the vehicle virtual model, so as to calculate the space vector of line E1R and E2S, the angle between these two lines is the blind angle of the driver's two eyes.

3 AUGMENTED REALITY SYSTEM FOR ERGONOMICS EVALUTAION

With Augmented Reality technology, a preliminary augmented reality system named ARPEAS has been developed for the specific application of automotive interior evaluation. The system includes two parts; one is a static driving hardware experimental environment, and the other is the corresponding software system.

3.1 *Hardware components*

The hardware of ARPEAS system has several components including two 1394 cameras for vision-based registration, the NDI position tracking device for optical tracking, the frame for fixing cameras and NDI device, the semi-physical mock-up of car for simulating the driving environment, as well as one PC for image processing and synthesis. All of hardware are organized together, is shown in Figure 4.

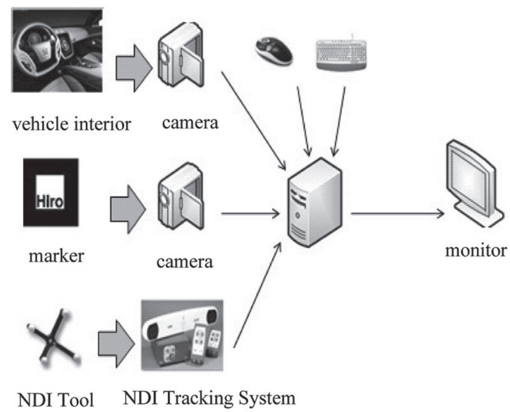


Figure 4. Hardware components for the ergonomic assessment system.

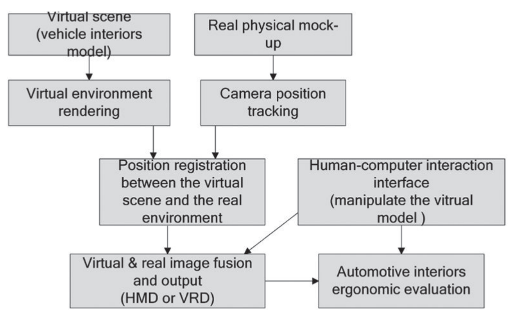


Figure 5. General scheme of AR-based automotive interiors ergonomic evaluation.

3.2 *Overall scheme and structure*

The general scheme of AR-based automotive interiors ergonomic evaluation is shown in Figure 5. Virtual scene rendering module is responsible for rendering the virtual vehicle interiors. Using the digital camera, the image of the marker and the real environment is captured for the marker-based tracking and registration. Through the accurate registration between real physical driving mock-up and virtual automotive interiors, the evaluator can interact with the AR environment to assess the ergonomic of the automotive interiors, and the motion data of the human body can be collected by the motion tracking devices for the further analysis.

3.3 *Hybrid tracking and registration*

The Marker-based AR visual registration is easy to realize, but has poor real-time performance and its registration accuracy is not high (Richard 1995).

Optical position tracking registration has faster sampling rate and high tracking accuracy, but the optical position tracking hardware is more expensive and the effective tracking range is not large (Thad et al. 1997; Wayne et al. 2003). Synthesize both advantages and disadvantages, we choose hybrid tracking technology for camera pose and human body motion tracking.

The hybrid tracking method used in this paper combined marker-based visual tracking and optical tracking by means of a dedicated infrared camera system-NDI Polaris Spectra which includes two infrared cameras and several NDI tools whose poses can be calculated in real time. The implementation of hybrid tracking system is shown in Figure 6. WCS is the global coordinate system in the real world and its origin is fixed at the datum mark in the real physical model. NT_1CS is the NDI tool1 local frame which is regarded as the conference coordinate system for optical tracking. So the transform matrix of the NT_1CS with respect to the WCS is the MNT1-W which can be measured in the pre-calibration process. The optical tracking system collects the position and orientation of every tracking tool named $M_{NDI1,}$ $M_{NDI2,}$ … M_{NDIn} in the NDI original coordinate system and calculates the relative pose matrix of the other tools with respect to the NDI $Tool_1$ named $M_{NDI2\text{-}NDI1,}$ $M_{NDI3\text{-}NDI1,}$ … $M_{NDIn\text{-}NDI1}$.

$$M_{NDIn\text{-}NDI1} = M_{NDIn} \times M_{NDI1}^{-1}. \quad (1)$$

So the pose matrix of the $NDITool_n$ in the WCS is

$$M_{NDIn\text{-}W} = M_{NDIn\cdot NDI1} \times M_{NT1\text{-}W}. \quad (2)$$

The camera coordinate system CCS1 represent the local frame of camera1 whose pose respect to the marker can be calculated by the vision-based AR registration method. The transform matrix of the camera1 with respect to the marker is MC1, also the MC2. So the transform matrix from CCS2 to CCS1 is $M_{C2}{*}M_{C1}^{-1}$. For the fusing display of the virtual scene and the real world, we assign the camera1 pose matrix MC1 as the virtual viewpoint position and orientation matrix in the virtual environment, that is, the virtual world coordinate system origin of the coordinate origin placed on the Marker. In order to the alignment of the virtual model and real physical muck-up, we measured the coordinate transformation between the datum mark and the AR marker $M_{M\text{-}W.}$ The virtual vehicle coordinate point which is relative to the datum mark moves the transform matrix $M_{M\text{-}W}$ in the virtual environment. The camera1 pose in the WCS can be calculated by

$$M_{C1\text{-}W} = M_{C1} \times M_{M\text{-}W}. \quad (3)$$

Then, the virtual vehicle model and the real physical mock-up have been integrated in the global coordinate system.

Through the above transformation, the virtual vehicle model, the real physical mock-up, the NDI tracking tools, and the cameras have been unified in the global coordinate system WCS; also, the vision-based tracking method and the infrared optical tracking technology have been integrated in the AR system.

3.4 *Visual cone drawing*

In the scene which combines the virtual objects and real world picture, there are two translucent visual cones attached to the position of the two human eyes, so that the user can see the scope of their visual field and the driving blind area caused by the car's A-pillar. The size of the virtual cones can predefine by the user. The pose of the virtual cones describing the user's eyes position and direction in the global coordinate system can be calculated in real time by the coordinate transformation of the data of one NDI motion tracking tool which binding in the

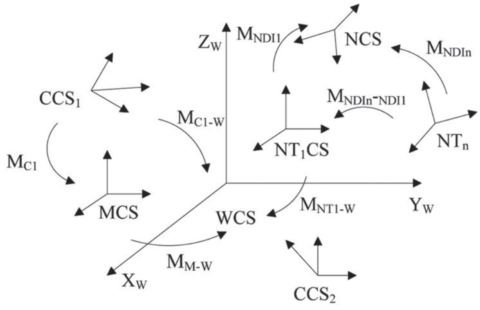


Figure 6. Coordinate systems for the hybrid tracking and registration.

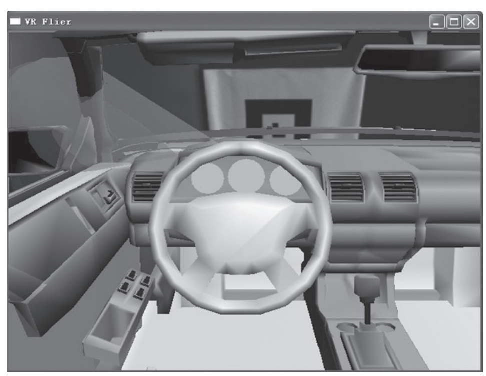


Figure 7. Virtual visual cones for visibility evaluation.

user's head. In the interactive process, the virtual cone's pose changed in accordance with the actual user's head movement in real time for interactive evaluation of driving vision field. The virtual cones in one certain time can be shown in Figure 7.

4 HARDWARE PRECALIBRALITION AND DATA ANALYSIS

Based on above technology of AR and registration, the hardware system is established and also the software system is realized, as shown in Figure 8.

In the system, taking the center point of the real car's steering wheel as the coordinate origin, the front direction as Z axis, the right direction as X axis and the head direction as Y axis, also making the VR coordinate coincides with this coordinate. Through calibration with the measurement device, $M_{NT1\text{-}W} = [480, 240, -10, 0, 0, 0]$, $M_{M\text{-}W} = [645, 342, 484, 0, 0, 0]$. Then, configuring the system with these parameters, the virtual vehicle model can be registered in the real vehicle platform and the hybrid tracking of infrared optical and vision-based AR register technology is realized.

During the interactive operation, the driver's head movement data $M_{NDI2\text{-}NDI1}$ is captured by the NDI tools, as shown in Table 1, the NDI_1 tool is fixed on the head of the driver and the NDI_2 tool on the platform statically, the capturing frequency is 5 times per second. In the NDI_1's coordinate, the position of the driver's eyes midpoint is $M_{NDI2\text{-}NDI1.}$ The section of the A-pillar is shown in Figure 9, the red profile is A-pillar's border and the 118.3 mm length blue line represents the equivalent blind line, which is measure in the virtual vehicle. Through measurement of the virtual vehicle model, in the global

Figure 8. Hardware system for vehicle interior ergonomic evaluation.

Table 1. Measured data for head movement.

$M_{NDI2\text{-}NDI1}$			$M_{E\text{-}S}$			Angle
–477.36	–676.82	405.52	321.34	488.32	648.82	3.232
–483.40	–680.90	405.30	315.30	488.10	652.90	3.239
–487.94	–677.86	406.17	310.76	488.97	649.86	3.257
–491.23	–680.32	405.58	307.47	488.38	652.32	3.261
–488.93	–681.33	405.11	309.77	487.91	653.33	3.253
–485.45	–686.28	403.81	313.25	486.61	658.28	3.235
–482.25	–692.08	402.48	316.45	485.28	664.08	3.215
–481.87	–691.75	402.37	316.83	485.17	663.75	3.215
–484.8	–689.38	403.54	313.9	486.34	661.38	3.226
–484.2	–686.65	404.02	314.5	486.82	658.65	3.230

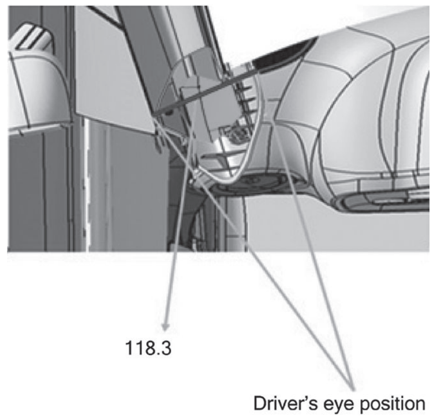


Figure 9. Section view of A-pillar.

coordinate, the center of the A-pillar's position $M_{S\text{-}W} = [-320.2, -301.0, 195.8]$. Then the relative position between the center of the driver's eyes and the A-pillar can be calculated by the equation: $M_{E\text{-}S} = M_{E\text{-}W} \times M_{N\text{-}W\text{-}1} \times M_{NDI2\text{-}NDI1}$, the calculated $M_{E\text{-}S}$ result is shown in table 1. Based on the principle of section 2, and take 68.5 mm as the distance of the driver's two eyes, the blind region angle can be computed, as shown in table 1. In the captured data domain, the minimum blind region angle is 3.21°.

Through the analysis result above, the virtual vehicle can be registered into the real physical environment accurately, based on the hardware technology, it provides designer with a real AR vehicle interior evaluation environment. The captured data can be taken as the basis for the car designer to make decision.

5 CONCLUSIONS AND FURTHER WORK

In the vehicle style design stage, the AR technology can really provide the designer an effective tool to evaluate the vehicle interior design from designer's subjective view. In this paper, through the analysis of the vehicle visibility assessment based on eye ellipse method, a static driving ergonomic assessment prototype system is built based on AR technology. With this system, the virtual vehicle interior can be registered to the real car environment, based on the mixed AR registration technology by NDI optical tracking system and marker-based vision registration technology. Through the analysis of captured data for A-pillar blind angle, it shows that the AR system can provide vehicle designer with visual tool for different vehicle interior design assessment.

Taking the AR technology into the vehicle interior assessment, it involves different kind of knowledge from different disciplines. Works described in this paper is still at the first stage, further research work is needed and improvement is necessary, such as:

1. Improvement of the efficiency of AR image processing algorithms, so as to prove the real time interaction.
2. Further research on the light and block between the real and virtual scene.
3. Development of better auxiliary tool for designer's subjective assessment data analysis.

ACKNOWLEDGMENT

The work is supported by Program "NCET" from MOE of P.R. China.

REFERENCES

Azuma, R. 1997. Survey of Augmented reality. *Presence: Teleoperators and Virtual Environments* 6(4): 355–385.

Azuma, R., Baillot, Y. & Behringer, R. 2001. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications* 11(15): 340–346.

Ekern, D.F. & Gedraitis, C.J. 1997. Kinetic computer modeling of human posture in automotive seats. *SAETeeh. PaP.*

Feiner, S., Maelntyre, B. & Hollerer, T. 2004. A touring machine prototyping 3D mobile augmented reality systems for exploring the urban environment. In: *International Symposium on Wearable Computers,* Cambridge*,* MA.

Fründ, J., Gausemeier, J., Matysczok, C. & Radkowski, R. 2004. Using Augmented Reality Technology to Support the Automobile Development. In: *CSCWD* 2004, Xiamen, China, May 26–28, LNCS 3168: 289–298.

Geuss, H. 1998. Optimizing the product design process by computer aided ergonomics. *SAETeeh. PaP.*

Kato, H. & Billinghurst, M. 1999. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: *Proceedings of IWAR*, San Francisco, USA, October 20–21, 85–94.

Milosevlc, S. 1997. Drivers' fatigue studies. *Ergonomoics*. 140(3): 381–389.

Richard, L.H. 1995. Registration errors in augmented reality systems. (Ph.D. dissertation) Department of Computer Science, University of North Carolina at Chapel HILL.

Thad, S., Steve, M. & Bradley, R. 1997. Augmented reality through wearable computing. *Presence, Special Issueon Augmented Reality* 6(4): 25–34.

Wayne, P. & Bruce, H.T. 2003. Augmented reality user interfaces and techniques for outdoor modeling. In: *ACM SIGGRAPH Symposium on interactive 3D Graphics*, California, USA, 228–230.

Yiming, Z. & Enrong, M. 1999. Ergonomics in vehicle engineering. *Beijing Institute of Technology Press.* (in Chinese).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A systematic ergonomic approach for the analysis, design and evaluation of hand-held tools

Helmut Strasser
Ergonomics Division, Siegen University, Siegen, Germany

ABSTRACT: Ergonomically designed hand-held tools must satisfy the equation "suitable for the human body = suitable for the hand," i.e., characteristics of the human hand-arm system such as motion ranges as well as limits of the various joints, have to be taken into consideration. The paper presents a systematic procedure for the ergonomic layout of the hand side of tools with respect to shape, dimensions, materials, and surface. The design must be preceded by a thorough analysis that examines what needs to be performed with the tool under what conditions and how and where and with which type of grip and coupling it needs to be performed. The ergonomic approach must always strive for holistic—rather than sectoral—goals and must consider interdependencies between the various design criteria. The systematic procedure is also helpful for the appropriate selection of truly ergonomically designed hand tools from an assortment of several variants.

Keywords: hand-held tools, ergonomic design, evaluation, systematic approach

1 ON THE IMPORTANCE OF HAND-HELD TOOLS OVER TIME

Hand-held tools have a long tradition. For centuries, their use provided the solid foundation for craftsmen. During the prime time of craftsmanship, they offered an amazingly high level of user-friendliness since they were manufactured based on the human body's dimensions. Contrary to today's practice in an industrial setting, such tools were not intended for multiple uses and the individual design measures were dominated by "popular standards" and "rules of thumb." Extremities were truly determining the dimensions as well as the shape of hand-held tools.

Unfortunately, today's mass production does not allow for such individualization, and the shelves of hardware and home improvement stores are typically dominated by standard sizes according to the metric system. Poorly shaped tool handles that are not customized lead to losses in the transmission of power, to an unbalanced pressure distribution in the palm and fingers with resulting blister formation, and thus to adverse effects on the tactile feedback. If hand-held tools are not manufactured according to the hand and the hand-arm system with their diverse anatomic and physiological characteristics, substantial detrimental effects on work efficiency and the human body can be expected. Repetitive use of poorly designed tools can even lead to work-related illnesses and occupational diseases.

Advertisers today hastily and frivolously sometimes use the term "ergonomic" for a product even if a critical examination would not reach such a conclusion since often only selected design approaches are realized. "*Ergonomically designed*" means more than a designer-created new shape of a product. The guideline for design measures must be the principle of compatibility (Strasser 1995). Technical design specifications must always be based on anatomical-physiological characteristics of the human body, e.g., the shape of joints and the range of motion of the hand with its dimensions. That is, in meeting the equation "*suitable for the human body = suitable for the hand*," it must be ensured that (response-response) compatible opportunities are created for the hand.

2 SYSTEMATICS FOR THE DESIGN OF THE HAND SIDE OF TOOLS

In the following, a systematic approach for the analysis and ergonomic design of hand-held tools will be described which has its roots already in the late 1970th (Bullinger & Solf 1979). It has been updated in the 1990th (see Bullinger 1994) and after having been applied successfully to lots of hand-tools, it has been presented again in an advanced version together with real-life examples and combined subjective rating and electromyographic assessment methods by Strasser & Bullinger (2007). As shown in Fig. 1, the actual design of the hand side of tools with respect to the correct *shape* and suitable *dimensions*, *material*, and *surface* should always be preceded by both a general and detailed analysis with various sub-items in order to scrutinize all constraints that can potentially impact on the design.

2.1 *General analysis*

In the analysis of the "*work task*," a crucial question concerns what must be accomplished with the tool how, where, and under what kind of work and environmental conditions (e.g., dry or "oily" handle surfaces). A number of criteria such as work resistance to be overcome, precision and time requirements, visual, auditory or tactile feedback, environmental influences and work safety, matter in this context that have an important impact on the "blocks" in the flow chart of Fig. 1.

The block "*Position and posture of the human body*" stands for body positions and postures that are required during work. Unfavorable body postures such as a twisted upper torso or a bent back lead to premature fatigue and cause long-term problems which should be avoided via appropriate designs. The same applies to constrained postures of the hand-arm system or the legs, which are necessitated by ergonomically poorly designed tools or difficult accessibility of the work location and tool and lead to unfavorable static muscle work. Especially during two-handed tasks in a closed kinematic chain, the design should allow for preferred directions of manual movements during work relative to the frontal body plane, e.g., 60° during filing with a vise. For repetitive movements of the hand-arm systems (as an open kinematic chain) in the horizontal plane, an angle of 30° between frontal plane and the direction of movement was found to be optimal while arm movements in any other direction require substantially more effort

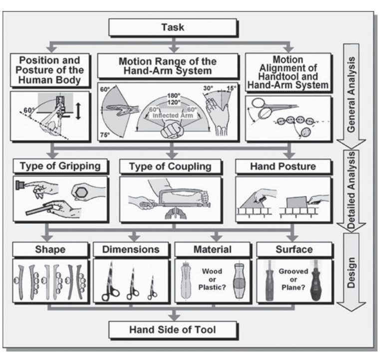


Figure 1. Flow chart for the design of the hand side of tools.

(cp. Strasser & Müller 1999). For the unaided holding of heavy hand tools, solutions must be found that, e.g., permit the optimal force potential of the biceps as a strong flexor of the forearm at an elbow angle of approximately 100°.

For the analysis of the "*Motion range of the hand-arm system*" knowledge of the various joints of the hand-arm-shoulder system with their ranges of motion is essential. The individual phalanges and the bones of the forearm and upper arm are connected through many different types of joints with differing degrees of freedom. Hinge joints and pivot joints each have 1 degree of freedom, saddle joints have 2, and ball and socket joints have 3 (rotatory) degrees of freedom. If the limits of motion of peripheral joints are exceeded, the joints that are one step closer to the torso must necessarily become active. This can ultimately lead to cumbersome movements of large body parts or even the torso itself.

As can be seen from the top middle part of Fig. 1, the wrist as the possibly most important joint allows substantial translatory dorsal and volar movements. Nonetheless, it has been shown in the case study on hairdresser scissors (see Strasser & Bullinger 2007) that even a range of 60° can be limiting for some tasks and results in constrained postures. The hand's range of motion that is limited to approximately 30° in ulnar direction is especially problematic, however. If the hand must be moved finger-dynamically repetitively in such a position, ailments and ultimately tendovaginitis can result.

As shown in Fig. 2, consideration should also be given in a tool's design to the advantages and disadvantages of a horizontal versus a vertical grip axis. For example, the hand's limiting horizontal motion range is not a factor with a socket wrench (cp. right part of Fig. 2) or a check lever on a lathe with vertical grip axis. Slippage of the grips is not an issue when the breakaway torque of a stuck screw must be overcome because the socket wrench is held in a positive coupling. Since joints allow straight or rotatory movements (which, e.g., are reduced from 180° with an outstretched arm to 120° with a bent arm) based on their type, attention must be paid to the compatibility of functional and anatomic joints. That is, rotatory movements should be carried out by a joint with a rotatory degree of freedom. The same applies to translatory movements with, e.g., hinge joints and ellipsoid joints.

Such facts will be discussed in the block "*Motion alignment of hand tool and hand-arm system.*" Representative for the various cases, in the upper right part of Fig. 1, the misalignment of the saddle joint of the thumb relative to the interphalangeal (hinge) joints of the fingers in the design of scissors cannot be without consequences.

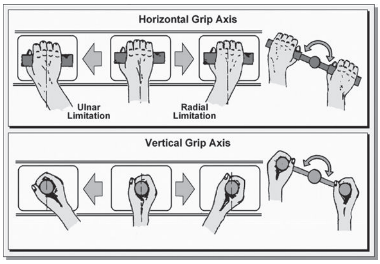


Figure 2. Horizontal grip axes of a handle on a hand-knitting machine and on a socket wrench that result in restricted range of motion due to the hand's limited ulnar and radial range (top) and vertical grip axes (bottom) that permit wide-ranging movements in the elbow and the shoulder joint using a power grip with positive coupling in a normal hand posture.

2.2 *Detailed analysis*

A detailed analysis of the "*Type of gripping*" must take the specific advantages and disadvantages of a contact, pinch, or power grip into consideration along with the different types of grip that the hand, in principle, can realize. While a contact grip allows for quick reactions, the fingers are not very resilient. A tool can be led with high precision using a pinch grip, but such precision is lost with a power grip in exchange for substantially higher forces and torques that can be transmitted.

With a power grip in which the distal, middle, and proximal phalanges of at least one finger as well as the thumb's distal and proximal phalanges—possibly together with the entire palm—are in contact with the handle, the coupling area is always largest and, assuming an ergonomic shape of the handle (see below), the pressure distribution in the hand is fairly even. Both positive and friction coupling, however, require a suitable dimensioning for a favorable bio-mechanical transmission of force (also see below).

With pinch grips on plier-like two-handle tools, e.g., all fingers are in contact with the two shanks. It must be possible for the legs to slide from the distal phalanx via the intermediate phalanx to the proximal phalanx when the tool is operated. The exerted force is, of course, minimal, when due to a rather large distance between the two legs only the weak distal phalanges couple. An absolutely minimum force occurs if even only the little finger's distal phalanx is involved. The maximum force can be achieved when the proximal phalanges are also able to reach the handle legs. Conversely, high precision and sensitivity can be expected when 2 or 3 distal phalanges couple in the two-finger grip when tweezers are used or with a dynamic three-finger pinch grip to hold a pen. When only the fingertips or the finger pads couple with the surface of, e.g., keyboards or a pushbutton switch in a contact grip, quick reactions are possible that are time-phased by the mobility of the various fingers or the hand. The thumb as the strongest finger can be used for even higher resistance, but it has deficits with regard to speed.

The analysis also has to consider whether the expected force when making a fist in normal "*Hand posture*" (when hand and forearm are aligned) can actually be mobilized or whether the hand's forced dorsal extension or volar flexion results in drastic force losses due to "tendon insufficiency." The contraction of the flexor digitorum superficialis during volar flexion of the hand and thus the exerted muscle force, e.g., has little effect on the phalanges since the tendons in their sheaths are too long in this position—similar to a bowden cable between actuator and actuating element. Dorsal extension of the hand is unfavorable as well since it leads to input tension of the system of muscles, tendons, and phalanges.

An unfavorable posture (ulnar deviation) can sometimes be the result of the geometry of a tool's work side, e.g., the triangular shape of mason trowel handles, which may have its advantages on other criteria. Repetitive finger movements in ulnar deviated hands (limited to a mean value of approximately 30°) pose the risk of tendovaginitis since the friction-reducing fluid at the deviating points in the tendon sheaths is displaced, thus resulting in friction when the tendons are moved.

Finally, how the hand's force is transferred to the tool is of crucial importance. The displayed hacksaw in Fig. 1's block on the "*Type of coupling*" exhibits the less favorable friction coupling for the right hand. As shown in Fig. 3, the only indirect transmission of force of the friction coupling is unfavorable for the finger musculature since higher forces are required than with positive coupling. Furthermore, positive coupling should provide for a forward-slanted grip surface similar to a pistol

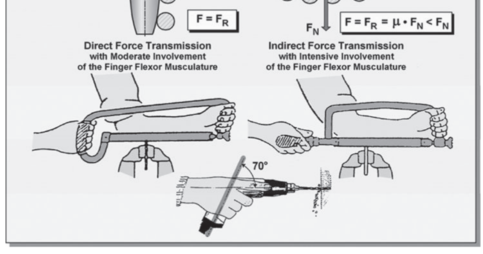


Figure 3. Friction coupling (top right) and positive coupling (top left) with examples of handles (bottom).

grip. It is also advantageous when the directions of force and function are aligned since tilting effects can be avoided.

2.3 *Design of the hand side of tools*

As far as the design of the hand side of tools is concerned, all details and characteristics that result from the general and detailed analyses must be coordinated and considered, which oftentimes results in compromises. For each type of grip and coupling, a highly distinct "*Shape*" of the tool's handle can result, while a number of common criteria must be realized. For example, it is advantageous to have shapes for finger or thumb contact grips on controls and on keyboards as well as on freely movable tools that exhibit concave cavities that are suitable for the dimensions of the coupling phalanges (cp. the section on "Dimensions"). This eliminates slippage on pushbutton, toggle and rocker switches as well as on pivot levers. Also, a change in direction of an actuation, e.g., with a throttle, does not require a time-consuming re-grasping (cp. top part of Fig. 4).

With 2-finger pinch grips (e.g., on industrial tweezers and keys), the shapes that are illustrated in the left and middle bottom part of Fig. 4 are found to be advantageous, and fan-like broadenings in the longitudinal contour of tweezers facilitate the handling. For picking up plane parts off a work surface, a tapering of the tweezers' hand side from the "work side" is helpful and allows for a normal hand posture. Writing utensils that are held with a 3-finger pinch grip should not have the round cross section that is tapered towards the tip that is commonly found on pens. Round or hexagonal pencils are not appropriate, either, and can lead to tension in the finger musculature. The only cross section that is compatible with a 3-finger pinch grip has the rounded triangular shape that is shown in the bottom right part of Fig. 4. It provides suitable support for the thumb's distal phalanx, the index finger pad, and the inside of the slightly lateral positioned middle finger. Rubber sheathing or slightly pressure-anthropomorphic (elastic) surfaces further help to avoid punctiform pressure. Engraving and dental tools oftentimes benefit from a triangular design as well.

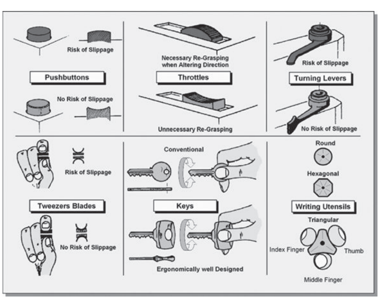
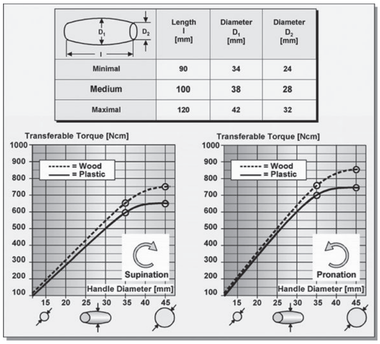


Figure 4. Basic shapes of controls and hand-held tools that are used with contact grip (top) and pinch grip (bottom), with and without compatible cross-sectional area and surface contour.

In the determination of the longitudinal contour of pliers' handles (cp. bottom left of Fig. 1), a spherical shape should always be considered since it follows the hand shape and ensures that all fingers are in contact with the handle in a pinch grip. This also applies to a power grip for single-handle tools. High finger forces can lead to very painful outcomes with the cambered handles that conform to a standard because it is not possible for all fingers to touch the handle surface without deformation of the hand. Additionally, the fingers get pressed against each other when force is applied. "Standardization does not prevent folly," and strictly following "Technical Rules" (cp. N.N. 1973) can sometimes lead to versions of tools that are questionable from an ergonomics point of view. For the cross section of handles on pliers, knives, and paring tools that must be held with a 4-finger pinch grip with the distal, intermediate, and proximal phalanges, the length of the middle phalanx must determine the thickness. Handles that are too thin can lead to contractions while shanks and handles that are too thick do not provide a good "fit." The optimal design for pliers' handles and diagonal cutters has a slightly rounded trapezoid cross section (Kluth et al. 2007).

As far as the "*Dimensions*" are concerned, they must be based on the hand with fingers and phalanges. Even if it will not be possible for practical and economic reasons to have sizes similar to clothing, at least three size groupings for large, medium-sized, and small hands would be desirable. For scissors (see bottom part of Fig. 1), that is actually absolutely essential since large eyes would lead to unacceptably poor coupling conditions for users with small thumbs and fingers. If handles, e.g., in pliers, do not have a spherical longitudinal contour for the pinch grip or for all tools that must be used with a power grip, large areas of the palm cannot couple with the tool. As a consequence, losses in the transmission of force or unnatural deformations of the hand occur.

Decades ago, Bullinger & Solf (1979) already recommended practical size gradations for the length and the diameter of handles. They referred to power grips and were suitable for rotatory as well as translatory motions in both frictional and direct, positive force transmission (cp. Fig. 5). For positive transmission for translatory motions, the diameters D_1 and D_2 could be approximately 4–6 mm smaller.



	Length l [mm]	Diameter D_1 [mm]	Diameter D_2 [mm]
Minimal	90	34	24
Medium	100	38	28
Maximal	120	42	32

Figure 5. Dimensions for power grips that are suitable for large, medium-sized, and small hands during rotatory activities with frictional coupling (top) as well as transferable torque with frictional coupling for supination (left) and pronation (right) with wooden and plastic handles of different diameters.

Pronation (inward rotation of the hand) with frictional force transmission generally, but especially with larger handle diameters permit substantially higher torques than supination (outward rotation). The reason is that the thenar is part of the coupling area during pronation. Studies with various models of screwdrivers (Strasser & Wang 1998) confirmed that the results clearly depend on the direction of the rotation, as visualized in the bottom part of Fig. 5. Additionally, it was shown that the differences in maximum torque between inward rotation of the dominant (typically the right) hand and outward rotation of the sub-dominant (typically the left) hand are in excess of 50%.

The higher operational output during pronation occurs even at lower levels of muscle strain, especially in the grasping musculature. This is of practical importance for the sparing use of physical force during manual work since even the sub-dominant hand during pronation is "stronger" than the dominant hand during the typical supination when screws are fastened.

The "*Material*" of choice for handles is nowadays plastic rather than wood. Plastic handles are typically acid-resistant and shatterproof and injection molding permits the simple and inexpensive production in a large variety of shapes that are suitable for the hand. Care must be taken to differentiate between greatly differing materials and "*Surfaces*" in the context of the friction coefficient. Since the surface and the material determine the friction coefficients they must be carefully chosen for all handles that are used with friction coupling. Cellulose acetate is slightly more expensive than polypropylene, but has a substantially better friction coefficient than the latter due to its slightly wax-like surface. If friction coefficients are too low, e.g., on brick trowels with lacquered wooden handles, slippage may occur during rotatory pivot movements. At the same time, friction coefficients can be undesirably high. Chiseled handles and surfaces with a rough texture carry the risk of high surface pressure for the hand and contractions of the skin with resulting blistering. Smooth, micro-textured surfaces result in the biggest coupling areas, which is optimal for force transmission. Comparative studies (cp. among others, Kluth et al. 2004a) have found multi-component surfaces to be particularly advantageous. Deviations from the ideal alignment of handle and hand, in particular during fast work, can be compensated via elastic, pressure-anthropomorphic surface parts. At the same time, the quasi-positive coupling leads to an increase in exertable force due to the elastic deformation. For areas on a handle with sliding coupling, hard micro-textured plastic was found to be more suitable. Pronounced malleable nubs, e.g., on so-called "power screwdrivers," are less suitable for the hand's surface, however.

It has been shown in several real-life examples (see Kluth et al. 2004b; Strasser et al. 1996; Strasser 2007) that the various design criteria must be coordinated. Selection of the wrong material can lead to worse outcomes even with the ergonomically appropriate shape. Especially with a handle that is used with friction coupling, pressure-anthropomorphic materials with soft, possibly slightly hollowed-out surfaces that conform to the hand and the convex phalanges are favorable.

3 EVALUATION OF ERGONOMIC QUALITY

The above shortly described complex system of work tool design can also successfully be used for the rating of ergonomic quality and the selection of tools. Additionally, tests that measure, e.g., deliverable torques or forces under real-life conditions are always recommended. Measuring "*physiological costs*," i.e., what the muscles involved in handling the tool must "*pay*" for the use of the tool is much more demandable. For details on the analysis of electromyographic time series with filtering, standardization, and extraction of ergonomically relevant parameters such as static and dynamic components see Strasser (2006). Such work physiological investigations and subjective assessments of various design criteria after standardized work tests by experienced users should always be invested in the assessment of the ergonomic quality of hand-held tools.

Screwdrivers will be used as an example to show that handles that conform to the hand's characteristics allow higher forces and torques. Reversely, a certain level of work can be accomplished with less physical effort. Amongst a set of various screwdrivers the standard handle (cp. top right part of Fig. 6) was the least suitable variant. The maximum torque that could be achieved was only half as high as with the handle that satisfied ergonomic requirements particularly well (cp. top left part of Fig. 6).

It is noteworthy that such operational output was achieved with almost identical physiological effort, i.e., with physiological costs that were measured in standardized myoelectric activity (sEA) values of the flexor digitorum as grasping muscle and the biceps as supinator. With a set point of equal sub-maximal demand of 20% or 40% of the individual maximum torque for 20 or 10 seconds, the determined electromyographic data showed that handles without ergonomic design require substantially more physiological resources (cp. bottom part of Fig. 6), i.e., that substantially more muscle force is required if the tool is not designed to match human dimensions.

As can be seen in the left part of Fig. 7, the number of test subjects (in %) who, on a specially designed questionnaire, indicated any kind of complaint with the cupped hand, thumb, and fingers after the completion of work with the standard model is substantially higher than with an ergonomic model.

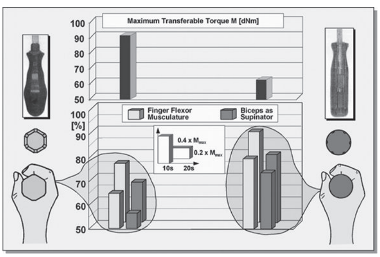


Figure 6. Maximum exertable torque and sEA-values of the biceps associated with an ergonomically designed screwdriver handle (left) and a standard model (right) at constant work (cp. Strasser et al. 1998).

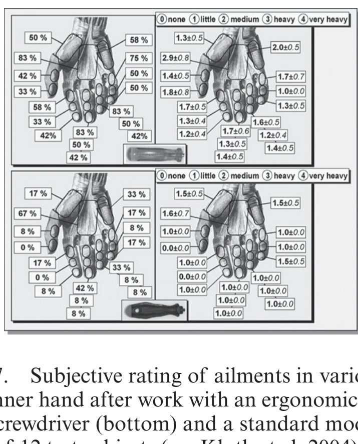


Figure 7. Subjective rating of ailments in various parts of the inner hand after work with an ergonomically optimized screwdriver (bottom) and a standard model (top). Means of 12 test subjects (see Kluth et al. 2004).

The results were similar for the perceived strength of an unfavorable distribution of pressure, as shown in the right part of Fig. 7. The preventive aspect of the ergonomic solution in the sense of reducing human effort as well as avoiding pressure marks and blisters should be apparent.

REFERENCES

Bullinger, H.-J. & Solf, J.J. 1979. *Ergonomische Arbeitsmittelgestaltung. I—Systematik. II—Handgeführte Werkzeuge—Fallstudien*. Forschungsberichte Nr. 196 und 197 der Bundesanstalt für Arbeitsschutz und Unfallforschung. Bremerhaven: Wirtschaftsverlag NW.

Bullinger, H.-J. 1994. *Ergonomie—Produkt- und Arbeitsplatzgestaltung*. Stuttgart: B.G. Teubner Verlag.

Kluth, K., Zühlke, D. & Strasser, H. 2007. Product-ergonomic evaluation of diagonal cutter handles. Chapter 1. In Helmut Strasser (ed.), *Assessment of the ergonomic quality of hand-held tools and computer input devices*. Amsterdam/Berlin/Tokyo/Washington DC: IOS Press: 197–2006.

Kluth, K., Chung, H.-C. & Strasser, H. 2004a. Verfahren und Methoden zur Prüfung der ergonomischen Qualität von handgeführten Arbeitsmitteln—Professionelle Schraubendreher im Test. Schriftenreihe Ergo-Med., Band 5. Heidelberg: Dr. Curt Haefner-Verlag.

Kluth, K., Kellermann, H.G. & Strasser, H. 2004b. Assessment of the ergonomic quality of file handles using electromyographic and subjective methods. *Occupational Ergonomics* 4(2): 133–142.

N.N. 1973. *Werkzeugnormen. Hand-Werkzeuge*. DIN-Taschen-buch Nr. 42. Berlin: Beuth Verlag.

Strasser, H. 1995. Ergonomic efforts aiming at compatibility in work design for realizing preventive occupational health and safety. *International Journal of Industrial Ergonomics* 16(3): 211–235.

Strasser, H., Wang, B. & Hoffmann, A. 1996. Electromyographic and subjective evaluation of hand tools: The example of mason's trowels. *International Journal of Industrial Ergonomics* 18(1): 91–106.

Strasser, H. & Wang, B. 1998. Screwdriver torque strength and physiological cost of muscles dependent on hand preference and direction of rotation. *Occupational Ergonomics* 1(1): 13–22.

Strasser, H. & Müller, K.-W. 1999. Favorable movements of the hand-arm system in the horizontal plane assessed by electromyographic investigations and subjective rating. *International Journal of Industrial Ergonomics* 23: 339–347.

Strasser, H. 2006. Electromyography: Methods and Techniques. In Waldemar Karwowski (ed.), *International Encyclopedia of Ergonomics and Human Factors. Vol. III. Methods and Techniques.* 2nd Edition. London/New York: Taylor & Francis: 3115–3118.

Strasser, H. & Bullinger, H.-J. 2007. A systematic approach for the analysis and ergonomic design of hand-held tools and control actuators—Visualized by some real-life examples. Chapter 1. In Helmut Strasser (ed.), *Assessment of the ergonomic quality of hand-held tools and computer input devices*. Amsterdam/Berlin/Tokyo/Washington DC: IOS Press: 1–22.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Photographic ergonomics: Holding a small digital camera

Egbert J. Hendrikse Jr. & James L. Smith
Texas Tech University, Lubbock, Texas

ABSTRACT: With the increasing reliance on an LCD screen with cameras, some of the strategies of holding a camera lose their effectiveness. This study investigated possible measures based on movement of the camera while the exposure is taken place. Independent variables included using one or two hands, hold the camera near or far, and holding the camera for 3 or 10 seconds before pressing the shutter button. The measures were based on movement data in the left to right (X) and up and down (Z) direction using a Motion Analysis system. The results were that the main effects of handedness and duration were significant for males, and handedness for females. The small sample size could be a limiting factor in the application of the results.

Keywords: Digital cameras, image stabilization, holding camera

1 INTRODUCTION

As the photographic industry has converted from a film-based media to an electronic (digital) media, cameras have undergone significant changes in physical shape as well as in their nature of image capture. For film-based cameras, which are typically larger than their digital counterparts, there were several strategies for handholding the camera to minimize image blur and obtain sharp images. Typical suggestions included keeping the elbows tucked in near the body, holding the camera with the right hand while supporting the camera/lens with the left hand, not breathing during shutter release, using a tripod for longer exposures, and releasing the shutter with a smooth depression force on the shutter release button (Koselj, 2008). The obvious question is whether or not these same guidelines apply to a smaller, lighter digital camera, especially cameras that have abandoned the eye-level viewfinder in favor of lcd screens that must be viewed as the camera is held away from the body.

Basic knowledge in ergonomics and how the body works can help explain the reasons behind the recommendations. Holding an item with two hands is far more secure than with only a single hand. Biomechanically, holding an item further away from the body requires far more muscle force than holding it close (Chaffin, Andersson, & Martin, 2006). If the muscles are then required to hold a static posture for an extended period of time, the muscles will become fatigued, causing discomfort for the individual and also result in some shaking (Chaffin et al. 2006). In order to see if these basic human responses hold true while holding a small camera, it would first require a measure that could adequately represent the shake while taking a photo.

Therefore, the main objective was to quantify the movement of a small "point and shoot" camera. Then second objective was determine how the number of hands you use to hold a camera, the posture while you hold a camera and the duration you hold a camera can influence the likelihood of blur. Therefore, our hypotheses were as follows: that two hands would result in less movement than holding the camera with one hand; that holding the camera further away from the body would result in more movement than holding the camera close; and that holding the camera for 10 seconds would result in more movement when compared to holding the camera for 3 seconds.

2 METHOD

2.1 *Subjects*

The subjects consisted of 18 students (12 male, 6 female) from an Industrial Engineering class. Ages ranged from 20–26 years of age. The mean weight was 89.0 ± 19.5 kg for males and 60.6 ± 11.1 kg for females. The mean height was 178.8 ± 5.7 cm for males and 163.6 ± 8.6 cm for females. A comparison of normative data (Tayyari & Smith, 1997) can be seen in Table 1.

2.2 *Apparatus*

The Motion Analysis motion capture camera system was used. This consisted of eight Eagle 4 cameras, and a computer to run the Cortex software (version 1.1.4.368). The camera used for analysis was a Nikon Coolpix 5900. An image of the camera and the markers that were attached to it can be seen in Figure 1.

2.3 *Experimental design*

The independent variables consisted of two hand conditions (one versus two hands), two duration conditions (holding the camera for three or ten seconds), and two postures (near and far). The near posture consisted of the subject holding the camera where the angle between the upper and lower arm was less than 90° and the far posture consisted of an angle greater than 90°. An illustration of the two postures can be seen in Figure 2.

Table 1. Comparison between measured weight (kg) and height (cm).

	Male	Female
Measured weight	89.0 (19.5)*	60.5 (11.1)
Comparison weight	74.9 (12.6)	63.7 (13.8)
Measured height	178.8 (5.7)	163.6 (8.6)
Comparison height	173.2 (6.9)	160.3 (6.6)

*84.3 (11.2) kg excluding one male subject who was over 99.9 percentile in weight (140.6 kg).

Figure 1. Photo of the camera used and the markers attached to the camera.

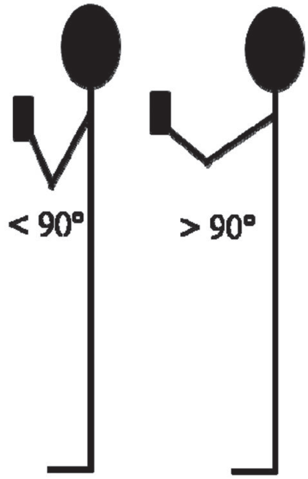


Figure 2. Illustration of camera postures near and far.

There were also two trials per condition. The dependent variables consisted of the movement as measured by the motion capture system. The exact measures will be explained in the data analysis section.

2.4 *Procedure*

Subjects were divided into groups of 4 and 5. At any given moment, one group was being tested, and one group helped with the data collection. The data collection group had a variety of roles. One member was in charge of training the member on their next test. Each subject was given two copies of a sheet with a list of 16 randomized trial numbers (1–16). One sheet was presented to the data collector, and one was presented to the trainer. Here the trainer would reiterate the procedure that needed to be followed, and gave the subject a chance to practice the posture. After the subject finished at the training station they would wait until their turn came. The group members cycled through so that each member only underwent one test in succession, giving each member enough rest (at least 2 minutes). When the subjects approached the designated area, they would be told their posture again by the data collector. The subject would show the data collector the posture and then lower the camera. Two fellow students that were part of the group helping with data collection group would make sure the posture was appropriate. Once the posture was approved, the subject would be asked to lift up the camera. The data collector would then collect the data for the appropriate amount of time (3 or 10 seconds). At the end of the 3 or 10 seconds, the data collector would shine a red laser pointer on a white page to indicate that it was time to take a photo, and also uttered the phrase "take the shot". Once the data collector confirmed that the photo was taken, the data collection was ended for that trial. Subjects were asked to refrain from breathing during the 3 second trial and during the 10 second trial they were queued when to refrain from breathing. The gripping posture can be seen below in Figure 3.

If one hand was used, this posture was implemented by the right hand. If the two hands were used, the same posture was used on both hands.

2.5 *Data analysis*

As part of the first step in the data analysis, the data collector would first make sure that all the markers were properly identified for the entire test. The data were then smoothed and converted to an excel file from a .trc file. The second step was to identify where the photo was taken. Using the Cortex software for each trial, the data collector would move the "time bar" forward and backward until he found the frame where the finger marker stopped moving by itself in relation to the other markers. A scatter plot of the X and Z movement data was plotted around that one point. The data collector would then find the data point where the scatter plot would abruptly change direction. The frame where this takes place would then be used as the point where the finger finished depressing the shutter button. This step was necessary since it was difficult to exactly determine when the finger finished depressing the shutter button. From this frame, 12 frames were counted to account for shutter lag (100 ms). An exposure time of 1/30 of a second was used to then select the next 4 frames as the time at which the exposure took place. One frame before and after the exposure frames were selected for all the markers and copied to excel for data analysis. Only the X and Z data were used since the X direction was left to right and the Z direction was in the up and down direction. The justification for removing the Y direction was that any movement in the front to back direction would be accounted for by the depth of field and would not result in motion blur. Two measures were calculated from the movement data. The first was called virtual marker. This marker would represent the movement the sensor saw during the exposure time. This was accomplished by halving the distance between the two middle markers in the X direction to get the X coordinates and halving the distance between the middle left and bottom marker in the Z direction. In both directions, the distance was added to the smaller of the markers to get the actual coordinates. An illustration of this can be seen in Figure 4.

From these calculations 6 frames would be generated for the X direction and 6 frames in the Z direction. The resultant movement was calculated using the X and Z coordinates of the virtual marker. This would result in 5 values. The middle three were used to represent the time of exposure and added up to generate the measure, virtual sum. The second measure was calculated by calculating the resultant for the four markers attached to the camera. For each marker the middle three values representing the exposure were added up and the average of these values across the four markers were classified as camera sum. These values were calculated for all of the subjects and for each test. A repeated measures ANOVA was used to test for significance. Means and standard deviations were also calculated.

Figure 3. Illustration of the correct posture (Koselj, 2008).

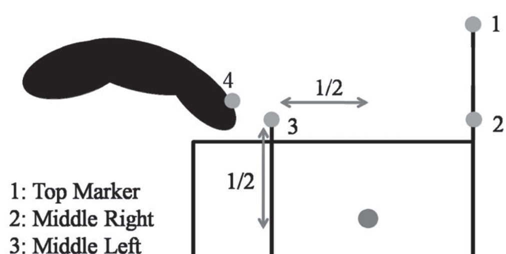


Figure 4. Marker setup and illustration of calculations of virtual sum measure.

3 RESULTS

3.1 *Means and main effects*

The means and standard deviations for the virtual and camera sum measures can be seen in Tables 2 and 3. Main effects for the virtual sum measure was found for handedness, $F(1,11) = 29.76$, $p < 0.001$, and duration, $F(1,11) = 9.21$, $p < 0.05$, for males and handedness, $F(1, 5) = 7.18$, $p < 0.05$, for females. The main effects for the camera sum measure was found for handedness, $F(1,11) = 29.67$, $p < 0.001$, and duration, $F(1,11) = 7.19$, $p < 0.05$, for males and no main effects were found for females. Even though no interactions were found, a graph of the Handedness and Posture condition was graphed below in Figure 5.

Table 2. Virtual sum measure means and standard deviations. (data in mm).

	One Hand Near 3 Sec Trial 1	2	10 sec Trial 1	2
Male: Mean	0.323	0.308	0.297	0.271
Male: St.dev	0.222	0.378	0.193	0.121
Female: mean	0.361	0.291	0.309	0.133
Female: St.dev	0.382	0.047	0.174	0.118
	Far 3 Sec Trial 1	**2**	**10 sec Trial 1**	**2**
Male: Mean	0.395	0.435	0.385	0.252
Male: St.dev	0.257	0.179	0.231	0.110
Female: Mean	0.246	0.301	0.245	0.140
Female: St.dev	0.171	0.190	0.195	0.052
	Two Hand Near 3 Sec Trial 1	**2**	**10 sec Trial 1**	**2**
Male: Mean	0.218	0.161	0.122	0.123
Male: St.dev	0.204	0.050	0.063	0.060
Female: Mean	0.084	0.130	0.151	0.080
Female: St.dev	0.020	0.068	0.135	0.041
	Far 3 Sec Trial 1	**2**	**10 sec Trial 1**	**2**
Male: Mean	0.249	0.259	0.137	0.143
Male: St.dev	0.123	0.097	0.054	0.052
Female: Mean	0.137	0.140	0.187	0.152
Female: St.dev	0.047	0.099	0.122	0.059

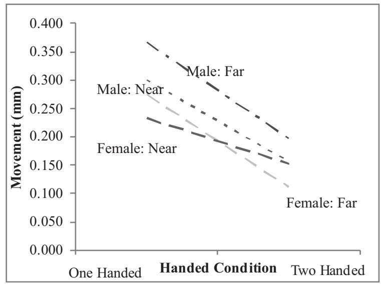


Figure 5. Graph of handedness and posture condition interaction for both males and females.

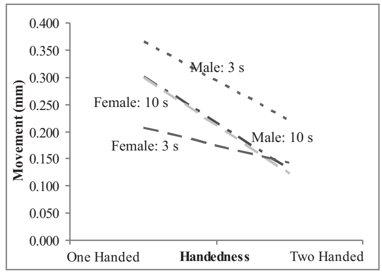


Figure 6. Graph of handedness and duration condition interaction for both males and females.

A similar graph was generated for handedness and holding time interactions, as can be seen in Figure 6.

4 DISCUSSION

Based on the results, virtual sum was able to distinguish significant effects where camera sum could not (especially for the smaller sample size females). The results also proved that holding a camera with two hands versus one camera will result in significantly less movement. Although duration was significant, the 10 second duration showed less movement than the 3 second duration. This was unexpected. A possible reason for this could be because the subjects were queued when to refrain from breathing. This could suggest that the subject did not care to hold the camera steady until after they were queued to hold their breath, reducing the effects of fatigue while holding the camera. A solution to this could have been either using instructions to notify the subjects that they should keep the camera steady at all times or refrain from breathing the entire 10 seconds of recording.

As can be seen from Figure 4, the female response to duration was not the same as the males. No significance was found for females other than the handedness. This could be due to the small sample size of 6 females.

Table 3. Camera sum measure means and standard deviations. (data in mm).

	One Hand Near 3 Sec Trial 1	2	10 sec Trial 1	2
Male: Mean	0.367	0.405	0.325	0.313
Male: St.dev	0.239	0.454	0.186	0.124
Female: Mean	0.390	0.326	0.345	0.184
Female: St.dev	0.435	0.125	0.149	0.145
	Far 3 Sec Trial 1	2	10 sec Trial 1	2
Male: Mean	0.410	0.439	0.396	0.274
Male: St.dev	0.241	0.192	0.217	0.111
Female: Mean	0.259	0.316	0.258	0.165
Female: St.dev	0.137	0.193	0.197	0.030
	Two Hand Near 3 Sec Trial 1	2	10 sec Trial 1	2
Male: Mean	0.221	0.165	0.135	0.133
Male: St.dev	0.207	0.036	0.050	0.049
Female: Mean	0.108	0.138	0.183	0.100
Female: St.dev	0.030	0.062	0.163	0.028
	Far 3 Sec Trial 1	2	10 sec Trial 1	2
Male: Mean	0.271	0.253	0.161	0.165
Male: St.dev	0.163	0.080	0.057	0.051
Female: Mean	0.147	0.167	0.179	0.163
Female: St.dev	0.055	0.067	0.113	0.062

Unfortunately, several limitations exist. The largest limitation was the selection of when the exposure was happening. Even though only one person looked through the data to find the exposure time, the lack of consistency was still present. A way to overcome this could be to only record data when the shutter is pressed instead of the entire time the subject is holding the camera. The second major limitation is the sample size, especially for the females.

5 CONCLUSION

The focus of this study was to create a measure that could be used to determine the movement after the shutter button was pressed. Main effects were found for handedness for both males and females. This suggests that holding a camera with one hand does in fact result in more camera movement than when two hands are used. The small sample sizes (especially for the females) and procedure could have resulted in the lack of significance for the other conditions.

The camera used for the experiment (Nikon Coolpix 5900) is typical of small digital cameras being sold today. The camera is small and light ($88 \times 60 \times 36.5$ mm and weighs only 180 g with the battery). The small sensor size (7.18 mm $\times$ 5.32 mm) implies that small camera movements can result in blurred images. Camera manufacturer recommendations offer only basic guidelines for holding camera. For example, Panasonic Lumix DMC ZS7 instruction manual states, "Stand with your arms close to your body and with your feet apart. Make sure that the camera does not move at the moment the shutter button is pushed."

The student subjects in this experiment were able to hold the camera reasonably steady (likely that image blur would not result in unacceptable photos enlarged to 20×30 cm or smaller). However, holding the camera with one hand did result in significantly more camera movement than when the camera was held with both hands.

Based on the results of the current experiment and general guidelines for camera holding, the following guidelines are offered for lcd type digital cameras:

Camera holding guidelines

- Keep feet shoulder width apart
- Keep arms as close to the body as possible
- Hold the camera with both hands
- Gently press the shutter release
- Hold shutter release button after depressing to allow for shutter lag
- Avoid holding the camera long distances away from the body, especially with only one hand
- Refrain from breathing when releasing the shutter.

REFERENCES

Chaffin, D.B., Andersson, G.B.J. & Martin, B.J. 2006. *Occupational Biomechanics.* Haboken, NJ: John Wiley & Sons, Inc.

Koselj, V. 2008. How to hold the camera with your hands for best support: Learning how to grip the camera. *Knol beta*. Retrieved May 27, 2010, from http://knol.google.com/k/how-to-hold-the-camera-with-your-hands-for-best-support#

Panasonic Lumix DMC ZS7 Instruction Manual, P31.

Tayyari, F. & Smith, J.L. 1997. *Occupational Ergonomics: Principles and applications.* Boston: Kluwer Academic Publishers.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Using customer perceived value for the evaluation of design concepts of physical user interfaces in CSCW

Youngbo Suh
Korea Institute for Defense Analyses, Seoul, Korea

Joobong Song & Myunghwan Yun
Department of Industrial Engineering, Seoul National University, Seoul, Korea

Taebum Ryu
Department of Industrial Engineering, Hanbat University, Daejun, South Korea

ABSTRACT: This study aims to evaluate a tangible interface for a new collaborative work support system at its conceptual design phase. Especially, we focused on how the Customer-Perceived Value (CPV), relating to the functionality of tangible interface, can be evaluated from integration of user-centered approaches. In this study, we elicited implicit user needs and functional requirements through scenario-based analysis and Hierarchical Task Analysis (HTA). By exploring the state-of-the-art technologies and deploying a relevant technology roadmap, we investigated technology alternatives that enable users to accomplish the required tasks effectively/efficiently and to satisfy their needs or functional requirements. For a quantifiable evaluation measure, CPV (Customer-Perceived Value) attributes were utilized to evaluate the benefits and the costs of current design concept compared to users' perceived alternatives. To evaluate the tangible interface, we developed a design concept evaluation framework consisting of 3 phases. In phase 1, users' functional requirements, considering human capabilities, limitations and characteristics were arranged through a series of analyses based on problem scenarios. In phase 2, technology alternatives derived from technology roadmap were correlated with users' tasks through relationship analysis, which made it possible to identify what technology would be appropriate for various users' tasks. In phase 3, design concepts are evaluated in views of whether it can support users' functional requirements. By utilizing the technology roadmap from phase 2, technologies applied to design concepts are also evaluated to tell whether they are appropriate to the trend of technological developments. Evaluation items are based on functional requirements and technology alternatives, and evaluation criteria are prepared by CPV attributes that were extracted from literature survey about measuring the value of a new product. Relative importance of each item and criterion is determined by conducting Analytic Hierarchy Process (AHP). Consequently, an evaluation checklist is developed.

Keywords: conceptual design, concept evaluation, scenario of use, technology trends analysis, CPV

1 INTRODUCTION

In the conceptual design phase, a design concept is developed and evaluated by defining and combining the product functional characteristics (Al-Salka, 1998). Conceptual design is important in a new product development cycle, because it determines success in the market, and it decreases the costs and the need for redesigning (Dhilon, 1996; Pugh, 1995; Huang, 2002; Ulrich & Eppinger, 2005). In the case of a new product, however, because the product is intangible, it is difficult to identify the users' requirements and related tasks (Suri & Monroe, 2003). It is also difficult to assess the value that users might perceive.

This study aims to suggest a systematic framework for evaluating design concepts of a new product in users' aspect. Especially, we focused on evaluating design concepts based on users' implicit requirements and tasks development trends of relevant technologies, and the users' perceived value related to the functionality of design concepts.

In this study, we elicited user needs and functional requirements through scenario-based analysis and Hierarchical Task Analysis (HTA). By exploring the state-of-the-art technologies and deploying a relevant technology roadmap, we investigated technology alternatives that enable users to accomplish the required tasks effectively/efficiently and to satisfy their needs or functional requirements. For a quantifiable evaluation measure, we utilized CPV (Customer-Perceived Value) attributes to evaluate the benefits and the costs of current design concept compared to users' perceived alternatives.

2 LITERATURE SURVEY

2.1 *Analyzing users' requirements*

In this paper, we proposed a procedure to identify and analyze users' requirements and their expected tasks by utilizing user scenarios and task analysis. Scenarios can be regarded as a set of use cases having common features that can be occurred in the middle of system or product usage (Carroll, 2000). The reason why scenario-based approaches are useful when it is necessary to identify users' requirement is that scenarios deals with specific situations (Maguire, 2001). Scenarios are also described in terms of how users use the system, which enables designers to focus on users' preferences and needs.

Task analysis includes all kinds of methods to collect, classify and analyze the data about user performances (Diaper & Stanton, 2004). Task analysis usually begins with deciding product development goals. After that, a task transcript including user performance based on the task hierarchy is prepared. Based on task transcript, an activity list, which shows each task that user performs as a list, is described. With that list, it is possible to identify correlated tasks, because components that the list includes are positioned under the correlation of each task. As a result, detailed analysis is iteratively performed to supplement task hierarchy and activity list, and thus a confirmed task model is obtained.

2.2 *Technology trends analysis*

There are various methodologies used for technology trends analysis (Bray & Garcia, 1997). For technology trends analysis in this study, we put forward the way of using technology roadmap. Technology roadmap is a method that helps selecting the appropriate technology alternatives based on nowadays' technology trends. It could be a technology strategy for unclear future, so it helps taking core technology earlier than others (Bray & Garcia, 1997; Albright & Kappel, 2003). Technology trends analysis is proper nowadays, where development and decline of technology are so rapid. It is a very efficient way to present vision and goal for future technology environment, and to make a plan to take a core technology for mid-to-long term by figuring out the structure and time relationship between technological elements (Albright & Kappel, 2003).

2.3 *CPV approaches in NPD*

Most of new product development specialists agree that it is important to reflect customers' needs at early stages of development. Understanding customers' intention as they pay for specific things improves acceptability and value of new product. In this study, we evaluate the value that customers assume about concept of new product by CPV (Customer Perceived Value). Developing new product based on CPV makes catch on attributes that determine whether customers buy or not (Woodruff, 1997).

There are many definitions of CPV in existing studies (Woodruff, 1997; Chen & Dubinsky, 2003; Monroe, 1990; Zeithaml, 1988; Zeithaml, 2000). However, most of them are based on a common concept. At first, CPV is inherent and related with using product or service. Second, CPV is not a routine, but is perceived by customers. Third, CPV is trade-off between paying for the value and customers' benefit by buying. Based on these things, CPV is defined in this study as "customers' cognitive evaluation for customers' perceived benefit by using product or service and cost to take it". As a result, CPV has two attributes, customers' benefit through product or service and cost for it. In this study, customers will evaluate the cost and benefit. It could help comprehending customers' acceptance of new product design concept and its value.

3 RESEARCH FRAMEWORK

In this chapter, we suggest a process of development design concepts evaluation framework consisting of 3 phases (see Figure 1).

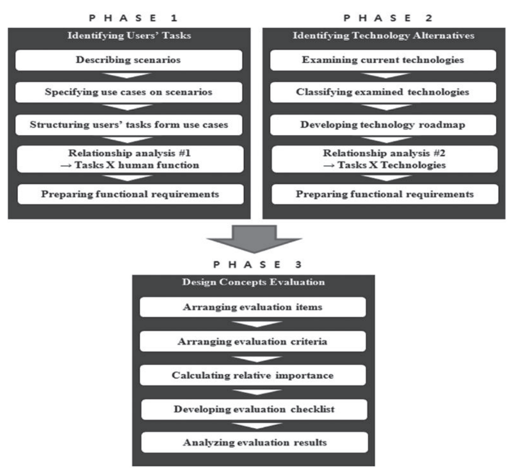


Figure 1. Design concepts evaluation framework.

3.1 *Identifying users' task*

Objective of phase 1 is eliciting use cases based on situations described scenarios that are available for applying the current design concept. First of all, a problem to be solved by design is described in scenarios. Then various use cases in the scenarios are identified, and a claims analysis is done to find the effect that use cases could influence to customers' tasks. Namely, the trade-off between positive and negative effects that could be made by function of new product are analyzed (Nielson, 1991). '+ (pros)' describes positive effect of related use cases to customers' behavior, and '– (cons)' describes negative effect. This way, subjective evaluation can be done about elements in scenario. Claims analysis is appropriate in developing scenarios in various ways based on every use cases in the scenarios, and it can be used to make a new idea for a product. Then, results of use cases are transformed as users' tasks, and then decomposed into sub-tasks and structured through HTA (Hierarchical Task Analysis). HTA is a methodology that divides one behavior into different tasks and sub-tasks to identify hierarchical structure of the behavior (Diaper & Stanton, 2004). These sub-tasks are correlated with various human functions through a relationships analysis. As a result, it is possible to arrange users' functional requirements considering human capabilities, limitations and characteristics.

3.2 *Identifying technology alternatives*

In phase 2, a comprehensive survey is conducted to investigate technologies relevant to specified users' tasks in a holistic perspective. It involves a survey of product or technology that already exists in market and also technology that could be developed in 4~5 years. Identifying technology trends that need more than 10 or 20 years to develop are very chancy and difficult for making specific technology implementation plan. After the technology survey, nowadays' technology developing trends are analyzed. Examined technologies are then classified into several categories based on the similarity between functional characteristics of each technology. Each of them are then further analyzed to estimate development trends. Based on the results, a technology roadmap is deployed by expert panel. Commonly, an axis of technology roadmap is market, product, technology or R&D, but in this study, we focus on technologies that could be developed in 4~5 years. Hence, we make a technology roadmap founded on technological characteristic, not market environment.

With help of the technology roadmap, it is possible to evaluate whether a technology applied to a design concept reflects the prospective trends. Conducting relationships analysis between users' tasks and technology alternatives derived from the technology roadmap, identification of the technology appropriate for various users' tasks is expected.

3.3 *Design concept evaluation*

In phase 3, functional requirements are identified to tell whether it can support users' functional requirements by a relationships analysis between users' task and human function. Then, effective technology alternatives to support specific user tasks are identified by relationship analysis between users' task and technology alternatives. The result of the first relationships analysis is used to evaluate design concept, and the second is used on technology alternative to support users' functional requirements. After that, sub-attributes of CPV are identified through a literature survey (Bitner, 1997; Curran, 2003; Parasuraman, 1998; Suri & Monroe, 2003). Then a list of them, are made to develop evaluation checklist. Criteria of sub-attributes for checklist are; 1) Selection criteria: whether it can used to evaluate on the stage of conceptual design or not, 2) Integration criteria: whether there is a similar or duplicated concept or not, and 3) Deletion criteria: whether it included users' subjective affection. According to these criteria, 24 attributes are selected, as evaluation criteria. Relative importance of each item and criterion is determined by conducting Analytic Hierarchy Process (AHP). Consequently, an evaluation checklist is developed.

4 CASE STUDY: VERIFICATION OF THE PROPOSED EVALUATION FRAMEWORK

A case study was conducted to evaluate the design concepts of a new CSCW-based tangible user interface, which was designed to support group decision making activities such as business meetings and ideations. The case study was conducted in accordance to each phase of proposed framework.

In phase 1, we developed 3 scenarios based on various types of meeting activities which were extracted from documents on group decision making processes. The 3 scenarios are 1) meeting on schedule planning, 2) meeting on proposing new ideas, and 3) meeting on deciding the optimal solution between various alternatives. Use cases implied by each scenario were identified, and their pros and cons were further analyzed. Users' tasks and subtasks such as 'show schedule', 'enumerate discussion issues', and 'record individual opinions' were then specified and structured by HTA method. As a result, 8 tasks and 47 sub-tasks were derived.

In phase 2, comprehensive document survey was conducted on state-of-the art technology developments in tangible interface technologies to support users' task based on the scenario analysis and task analysis. We examined and classified 124 tangible interface technologies into control-related technology group and display-related technology group, which were further organized each group by date of development.

Based on the results of classification, we analyzed development trends of each group in terms of functionality. For example, 'screen-based control' has improved its controllability by employing pucks in its execution. As a result, we produced a technology roadmap showing the prospective trends of tangible interface technology. Each technology alternative was listed and correlated with users' tasks from phase 1.

In phase 3, we evaluated the design concept of CSCW-based tangible interface using the evaluation checklist. The evaluation items were based on functional requirements that were derived from relationships analysis. Through a relationships analysis, each task was correlated with human functions which enabled us to specify users' functional requirements that are relevant to meeting activities.

To set up an evaluation checklist we grasped the degree in which the design concept of new product supports functional requirements, then selected tasks whose degree of support is above half. In this case, Task Understanding, Meeting summary, Explaining meeting main topics, Idea brainstormings, and Comparison of alternatives was selected.

A design concept evaluation was performed using the evaluation criteria derived from detailed CPV attributes. CPV attributes related to each tasks were chosen by pair-wise comparison between selected tasks and CPV attributes. As a result, 4 benefit attributes and 4 cost attributes were included. We used the following questionnaire sheet.

To evaluate CPV, 10 participants (8 male, 2 female participants) compared 2 types of meeting situations. Situation A was a typical meeting process without the support of the design concept of tangible interface. Situation B was the same situation

Year		~1999	2000	2001	2002	2003	2004	2005~
Screen-based Control	Finger/Pen/Eraser	Digital Desk/Magic Paint /Magic board	Magic Table	Finger tracking	Sketch Understanding	Finger tracking/Egaku	I/O brush/Finger tracking	
	Token/Puck/ Shuttle/Jog Wheel		Magic Table	Sense Board	Actuated work bench/Supply chain Visual Sensing/Sense Table/Tangible Query Interface/Audiopad	IP Network Design Workbench/Tangible Video Browser	AirportSim/circuiTUI/Tangible Business Process Analyzer/ Topobo Backpacks and duck	Disaster simulation
Input-Device-based Control	Mouse/ Keyboard	Digital Desk/Magic Paint			Intelligent Room Goals		Finger tracking	TheInfoCanvas
	Switch/ Button		Strata		ComTouch			Personal Audio
Action-based Control	Motion		Bottlegues	Lumitouch/Genie /Bottles/TellTale/ Pegblock	Doll talk/Dude's magic box & grandma's lapdesk		Topobo	Family Video Archive/Robo topobo/Technology Coach
	Gesture			Finger tracking				
Image-based Control	Facial expression			Digital family portraits	Dude's magic box & grandma's lapdesk			
Bio-signal-based Control	Bio-sensor					Plant Sensor	Finger Tracking	Grey Water Recycling
Sound-based Control	Voice				Doll talk Intelligent Room/Goals/Multilingual Speech Interface			Personal Audio
Object-based Control	Modeling		Luminous Room	CADcast	Illuminating clay		Sandscape	

Figure 2. Classification of display technology.

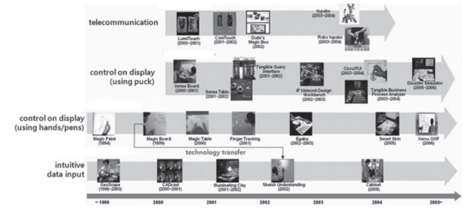


Figure 3. Tangible interface technology roadmap.

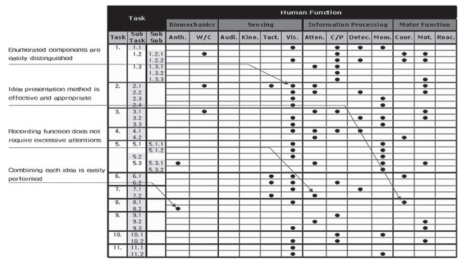


Figure 4. Human functions—task relationships analysis.

TITLE

Page Numbers

Date		Evaluator		Confirmation	

Functional Requirements	Users' functional requirements required to be satisfied	Rating	
		Relative Importance	
Current Functionality	Functionality provided to be intended to support the above requirement (required to be written by evaluator) * If not supported, remain as a blank	1 = not supported 2 = partially supported 3 = fully supported	/ 3

Evaluation Criteria		Weights	Rating Scale	Ratings
Benefit	Benefit Attributes I-1		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Benefit Attributes I-2		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Benefit Attributes I-3		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Benefit Attributes I-4		1 Low ←------- 5 Medium -------→ 9 High	/ 9
Cost	Cost Attributes II-1		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Cost Attributes II-2		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Cost Attributes II-3		1 Low ←------- 5 Medium -------→ 9 High	/ 9
	Cost Attributes II-4		1 Low ←------- 5 Medium -------→ 9 High	/ 9

Figure 5. Evaluation questionnaire sheet.

with the support of the design concept. They were asked to evaluate each situation and rate CPV attributes using a 9-point Likert scale.

Analysis results show that there are considerable differences between the detailed factors of benefit attributes. Especially, a considerable difference was frequently found between simplicity and other factors, which reveals that users cannot feel much convenience from the use of design concept. There is also a considerable difference between detail factors of cost attributes. Instructions tend to have a high score. As high score in cost attribute infers that more cost is required, this can be perceived as a negative factor of design concept. The following table shows the average scores of CPV attributes which includes the evaluation of 5 types of tasks.

We confirmed that the task of 'comparison of alternatives (1.266)' marked the highest score among total grades of CPV.

Table 1. CPV total score.

Task	Benefit score	Cost score	CPV score
Task understanding	6.070	5.117	1.186
Meeting summary	5.131	5.082	1.010
Explaining meeting main topics	4.845	4.514	1.073
Idea brainstorming	6.171	5.860	1.053
Comparison of alternatives	6.471	5.113	1.266

5 CONCLUSION AND DISCUSSIONS

In this study, we proposed a systematic framework for evaluating design concepts of new product based on user task and functional requirement through technology analysis. We identified use cases by using scenario-based analysis, and then classified and structured user tasks using HTA. By searching for state-of-the-art technology and deploying a technology roadmap, we were able to identify technological alternatives and fulfill users' functional requirement and tasks. Also, analysis of the value of new product design concept was conducted using CPV attributes with quantitative measures.

At the conceptual design phase, engineering specifications are not specified. Therefore, monetary cost, which may be crucial to product purchase, can not be precisely estimated. In the process of evaluating design concepts, we had no choice but to exclude cost attributes such as monetary expenditure. These tendencies are expected to surface throughout the conceptual design of new products. Therefore, a system of value evaluation taking aspects of evaluating market value of new products is necessary, and a continued effort in the future for improvement it is necessary. Also, the objective of our analysis of technological trend was to identify technology that will be developed within 3~5 years as a technological alternative. However, a systematical method of predicting long-term technological alternatives should be complemented within the current method of technological trend analysis in the aspect of developing of a genuine new product. However, it is still expected that our framework would be effective at the phase of evaluating design concepts of a new product from the user-centered design perspective.

REFERENCES

Albright, R.E. & Kappel, T.A. 2003. Roadmapping in the Corporation. *Research Technology Management* 46(2): 31–59.

Al-Salka, M.A., Cartmell, M.P. & Hardy, S.J. 1998. A framework for a generalized computer-based support environment for conceptual engineering design. *Journal of Engineering Design* 9(1): 57–88.

Bitner, M.J., Faranda, W.M., Hubbert, A.R. & Zeithaml, V.A. 1997. Customer contributions and roles in service delivery. *International Journal of Service Industry Management* 12(5): 423–450.

Bray, O.H. & Garcia, M.L. 1997. *Technology roadmapping: the integration of strategic and technology planning for competitiveness*. Sandia National Laboratories: 25–28.

Carroll, J.M. 2000. *Making Use—Scenario based design of human computer interaction*. MIT press.

Chen, Z. & Dubinsky, A.J. 2003. A conceptual model of perceived customer value in e-commerce: A preliminary investigation. *Psychology & Marketing* 20(4): 323–347.

Curran, J.M., Meuter, M.L. & Surprenant, C.F. 2003. Intentions to user self-service technologies: a confluence of multiple attitudes. *Journal of Service Research* 5(3): 209–224.

Dhillon, B.S. 1996. *Engineering design: a modern approach*. Time Mirror Higher Education Group Inc. Company.

Diaper, D. & Stanton, N.A. 2004. *The handbook of task analysis for human-computer interaction*. Lawrence Erlbaum Associates.

Huang, G.Q. 2002. Web-based Support for collaborative product design review. *Computers in Industry* 48: 71–88.

Maguire, M. 2001. Methods to support human-centered design. *International Journal of Human-Computer Studies* 55: 587–634.

Monroe, K.B. 1990. *Pricing: making profitable decisions*. McGraw-Hill.

Nielson, J. 1991. Paper versus computer implementations as mock up scenarios for heuristic evaluation. *Human-Computer Interaction Interact'90. Amsterdam: North-Holland*: 315–320.

Parasuraman, A. 1998. A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64(1): 12–40.

Pugh, S. 1995. *Total Design.* McGraw-Hill Companies Inc.

Schoonmaker, S.J. 1996. *ISO 9001 for Engineering and Designers.* McGraw-Hill Companies Inc.

Suri, R. & Monroe, K.B. 2003. The effects of time constraints on consumer's judgments of prices and products. *Journal of Consumer Research* 30(1): 92–104.

Ulrich, K.T. & Eppinger, S.D. 2005. *Product Design and Development.* McGraw-Hill Companies Inc.

Woodruff, R.B. 1997. Customer Value: the next source of competitive advantage. *Journal of the Academy of Marketing Science* 25(2): 139–153.

Zeithaml, V.A., Parasuraman, A. & Malhotra, A. 2000. *A conceptual framework for understanding e-service quality: Implications for future research and managerial practice*. Cambridge: Marketing Science Institute.

Zeithaml, V.A. 1998. Consumer perceptions of price, quality and value: A means-end model and synthesis of evidence. *Journal of Marketing* 52: 2–22.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

An evaluation metric on human-service interactivity of ubiquitous services

Joohwan Lee
Samsung S1 Co., Ltd., Seoul, Korea

Sangwoo Bahn, Ilsun Rhiu & Myung Hwan Yun
Department of Industrial Engineering, Seoul National University, Seoul, Korea

Hyosun Choi
Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, Korea

ABSTRACT: As many ubiquitous services have proliferated, the interactions between ubiquitous services and users have come to take an important position. Existing researches related to ubiquitous service evaluation systems have viewed the services as results of product use assuming ubiquitous services as products. Also, they were biased toward only ubiquitous computing technologies. It is necessary to develop a new evaluation method that evaluates interactivity with a user-centered perspective. The main objective of this study is the development of user-centered interactivity evaluation metrics of ubiquitous service attributes. Detailed goals of this study are as follows. First, existing ubiquitous service evaluation methodologies are reviewed to define the evaluation attributes for ubiquitous service evaluation. Second, user oriented interactive metrics that can overcome the limitations of those service evaluation methodologies, such as personal service evaluation techniques, usability evaluation techniques and psychometrics based evaluation techniques, are developed. As a validation study, suggested evaluation metrics are utilized to evaluate u-home service, and priority of each metrics is derived. The suggested evaluation metrics can be used to evaluate interactivity level of ubiquitous services and identify the potential users and their requirements at the stage of service development.

Keywords: user-centered design evaluation, ubiquitous service, interactivity, UEM (Usability Evaluation Method)

1 INTRODUCTION

The concepts of ubiquitous computing have been expanded to the concept of "Everyware", comprising the characteristics of computer hardware and software (Greenfield, 2006). The proliferation of ubiquitous computing suggests new paradigms of interaction inspired by constant access to information and computational capabilities (Abowd & Mynatt, 2000). The services in ubiquitous computing environments changed the subjects of marketing from developers to users. That is, "consumers" as "users" have come to emerge as principal actors in creating changes in today's societies along with their great influences. Consequently, the interactions between ubiquitous services and users have come to take an important position in ubiquitous industry. Also, new approaches to practical evaluations are required due to the emergences of unverified services.

The emergence of ubiquitous services, which has been transformed into the arena of state-of-the-art, is being fingered as a cause adding the economic burden to users due to temporary services neglecting the choices of users and one-sided services at the convenience of service providers. Commenced to overcome that vicious circle, the approach to large-scaled projects related to diverse ubiquitous services initially brought about communications among humans as well as between humans and computers. Moreover, the theoretical and technological advancements of service designs became to support collaborations between users (Resnick & Bowling, 2001; Takemoto et al. 2002; Blaine et al. 2005). They activated user studies relating to the selections of ubiquitous services (service matching/selection method) issues for the provision of the services complying with implicit needs of users focusing on the human factors such as usability, service awareness and trust (Lindenberg et al. 2007). On the other hand, it brought about visual advancements in complex viewpoints such as the performances of diverse components for the evaluation of the levels and qualities of ubiquitous services that have been progressing in many places, system usability and interactions between users and systems (Williams, 2004; Iqbal et al. 2005). This expansion of ubiquitous industry scope resulted in the need for development of techniques that will enable the evaluation of new and diversified ubiquitous services from the users–centered views.

Especially, as an approach to the evaluation of the levels and qualities of ubiquitous services, a small number of researchers attempted to apply an evaluation methodology. They used an extension of existing service quality evaluation methodology, which is an application of software quality evaluation methodology and usability among core concepts of UCD (User-Centered Design). Nevertheless, these approaches revealed their limitations by birth of a partial evaluation method that could not reflect the characteristics of ubiquitous services on appropriate places and evaluated only certain components of ubiquitous services. That is, excluding the attributes that need to be considered in ubiquitous such as experience, motivation and engagement, it grafted existing service evaluation methodologies to ubiquitous environments and thereby had the limitation that it could not evaluate the essence of ubiquitous services (Abowd, 1998).

The main objective of this study is the development of user-centered interactivity evaluation metrics reflecting ubiquitous service attribute. Detailed goals of this study are as follows. First, existing ubiquitous service evaluation methodologies are reviewed to define the evaluation attribute for ubiquitous service evaluation. Second, user-oriented interactive metrics that can overcome the limitations of those service evaluation methodologies such as personal service evaluation techniques, usability evaluation techniques and psychometrics-based evaluation techniques, are to be developed. As a validation study, suggested evaluation metrics are utilized to evaluate u-home service.

2 THE DEVELOPMENT OF INTERACTIVITY EVALUATION METRICS

2.1 *Development procedures and methodology*

Figure 1 shows the process of the development of interactivity metrics in this study. The approach to this study was composed in two parts: the development of ubiquitous service interactivity attribute index and the development of interactivity performance measurements.

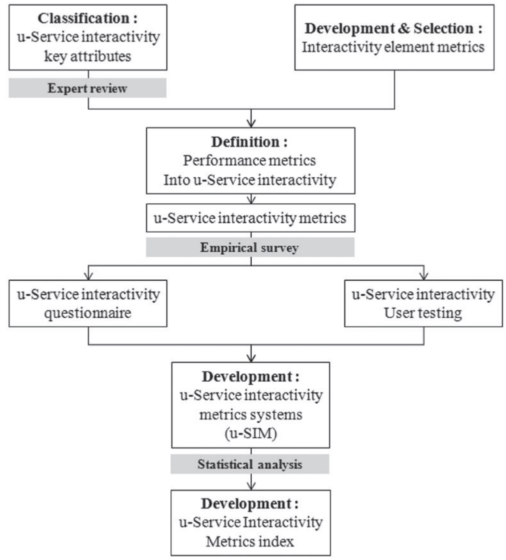


Figure 1. Development procedures of interactivity evaluation metrics.

The models of appropriate reference examples were too insufficient to define component of ubiquitous services. Thus this study analyzed the components of the services to constitute ubiquitous services such as IT services, inter-personal services, mobile services and system services.

The indices collected from literature review were integrated to define ubiquitous attribute indices. Expert group (policy maker, order placer, designer, developer, standard researcher) reviews were executed to select and structuralize the indices. Interactivity measurement definitions were made by classifying related indices with an analysis on the researches related to usability measurements and ubiquitous service evaluation methods.

The indices and measurements constituted the Evaluation Areas (EA) of ubiquitous service components to finalize metrics indices. Large scaled user surveys were executed on U-Home services for validation and priority weight estimation. The final metrics were developed as evaluation systems so that they can be readily applied to practices by designers to maximize utility.

2.2 *Interactivity attributes in ubiquitous service environments*

The interactivity attributes in ubiquitous service environments can be defined as the concept of "How easily the tasks for ubiquitous service can be executed and the user can be satisfied with results in user environments (Siewiorek, 2002)". The concept of interactivity attributes can be explained in the objective aspects related to the works executed within the integrated service consisting of ubiquitous service environments, relevant services and users. The results of the researches related to users' subjective aspects were implemented in the areas such as HCI and human factors.

In order to define the interactivity attributes of ubiquitous services, extraction of interactivity attributes through literature review, selection and integration of interactivity attributes and hierarchical classification of interactivity attributes were executed.

Thirty-one ubiquitous service interactivity attribute indices selected and integrated were structuralized through expert survey method. They were mathematically interpreted through the results of Principal Component Analysis (PCA) and Correspondence analysis. The surveys were executed by the same evaluators in order to maintain consistency in indices deducing and relation analyses. By analyzing the major components, 4 attribute indices with the explanation ability of 79.3% were extracted as in Table 1.

Figure 2 shows the structure for the proposed interactivity evaluation metrics in this study. Each sub-construct(Service capability, User experience, Ubiquity, Contextualization) consists of detail measure metrics, and each measure metric is comprised of 10 subjective metrics and 40 objective metrics.

Table 1. The attribute indices deduced with Principal Component Analysis (PCA).

Attribute	Description
Contextualization support	The level of the interactions considering the relations between services and users The level of the provision of customized services through status perceptions
Service capability support	The level of ubiquitous service user protection and error prevention The level of the performance, speed, security and storage ability of service systems
Ubiquity support	The extent of the ubiquity of ubiquitous services Ubiquitous connectivity of services and convenience in carrying devices etc
User-experience support	User participation and the degree of effort for use of ubiquitous services The communication direction/response level of ubiquitous service users

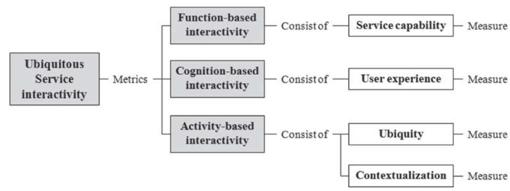


Figure 2. Structure for ubiquitous interactivity evaluation metrics.

2.3 *Development of metrics for ubiquitous service interactivity*

Since interactivity attributes of ubiquitous services have been defined, the measurement values that can evaluate the interactivity of ubiquitous services are required. Interactivity metrics are the measures to determine the levels of interactivity factors comprising the indices indicating what measurement values should be used to evaluate the extents of interactivity. Table 2 show an example of interactivity evaluation metrics suggested in this study.

The most important consideration in extracting the measurement values of interactivity is whether the measurement values are suitable to evaluate the levels of relevant factors because if uncorrelated measurement values are selected, the evaluations of interactivity factors become insignificant. This study was executed in the fashion of existing studied where measurement values were analyzed to extract all the lists selectable as measurement values and their correlations with factors were subsequently determined. Table 3 shows the correlations between evaluation factors and ubiquitous service interactivity attribute.

3 CASE STUDY

The experimental subject consists in the dimension of 125.6 m^2 and provides 17 ubiquitous service items. The items consist of 3 parts, i.e. facial recognitions for entrance, the door of future, UPIS (Ubiquitous Parking Information System), etc., media table relating to home network control, bidirectional TV relating to multi-media, medical camera relating to health, u-Health, intelligence healthy menu system, therapy music, light control, custom short message relating to living, magic mirror, digital frame, product information built-in OLED (Organic Light Emitting Diodes), Family collaborated learning system relating to education, etc. (Fig. 3). One hundred and eighteen participants (59 males, 49 females) evaluated the service. The evaluation was conducted by interactivity evaluation metrics for PDA.

Table 2. An example of interactivity evaluation metrics.

Measurable criteria			Metrics	Derived measure	Source (modified/adopted)
Component	Attribute	Indicator			
u-Service user	User expectation level (Perception) in relation with u-service functions	Agreement degree of service functions (SC: service concordance): ratio of expected service functions and the results. (u_1)	SC (Service Concordance) = A/B	A=Number of expected and comprehended service functions B=Number of functions provided from ubiquitous service	Constantine and Lockwood (1999)
		Participation degree of bidirectional communication, while using service. (u_2)	degree (questionnaire & user testing)	User's participating degree for bidirectional communication of ubiquitous service.	-
		Immersion degree in the service without own location-awareness, while using service. (u_3)	degree (questionnaire & user testing)	User's immersion degree in ubiquitous service	Scholtz (2006)
		The degree of understanding input data and expecting output at service request. (u_4)	A/B or degree (questionnaire & user testing)	A=Number of expectable and performable I/O B=Number of I/O provided from ubiquitous service	ISO 25000 (2005)
		The degree of receiving unexpected service by providing unrequested service functions. (u_5)	degree (questionnaire & user testing)	Whether the implicit needs of ubiquitous service user is to be provided or not	-
		Recognized degree of wasting time while service use. (u_6)	degree (questionnaire & user testing)	Recognized degree of wasting time during using ubiquitous service	-

Table 3. An example for evaluation factors of ubiquitous service interactivity attribute.

Measurable criteria \ The characteristics indicator of ubiquitous service interactivity (characteristic factor)			y_1 (user experience) Participation or effort level of user to use the service	y_2 (contextualization) Provides services adequate to the situation, considered the relation of service and user	y_3 (ubiquity) Portability and pervasive connectedness	y_4 (service capability) Level of user protection and error prevention
u-Service user	User expectation level (Perception) in relation with u-service functions	Agreement degree of service functions (SC: service concordance) : ratio of expected service functions and the results. (u_1)	O	-	-	-
		Participation degree of bidirectional communication, while using service. (u_2)	S	-	-	-
		Immersion degree in the service without own location-awareness, while using service. (u_3)	S	-	-	-
		The degree of understanding input data and expecting output at service request. (u_4)	S/O	-	-	-
		The degree of receiving unexpected service by providing unrequested service functions. (u_5)	-	S	-	-
		Recognized degree of wasting time while service use. (u_6)	-	S	-	-
	User's efforts(Perception) required within u-service use	User's approved time before using the service. (u_7)	O	-	-	-
		Learning time to use new service functions. (u_8)	O	-	-	-
		Time of user spending in hesitation or hold to use the service. (u_9)	O	-	-	-
		Number of user out-of-controls during service use. (u_{10})	O	-	-	-
		Ratio to change the service contents properly to user's preference or habit automatically. (u_{11})	-	O	-	-
		Required time / degree for user to modify the service function procedure in user's convenience. (u_{12})	-	S/O	-	-
	Cognizance degree of economical response to u-service (Perception)	Distance degree that user must move additionally to receive the service at proper place and time. (u_{13})	-	S/O	-	-
		Range degree of physical spaces to recognize the status of the service. (u_{14})	-	S	-	-
		Ratio of error occurrence during service use. (u_{15})	-	O	-	-

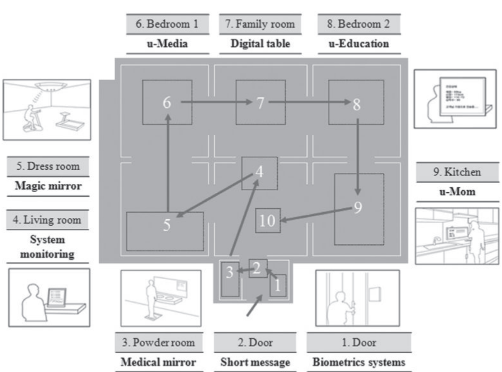


Figure 3. U-Home service evaluation process (10 stages).

The process of reasoning influence variable and deducing the model for analyzing data gathered by interactivity evaluation of ubiquitous service is as Figure 4. First, the basic statistic characteristics of the data are analyzed through descriptive statistic analysis. In addition, the characteristics of metrics are investigated based on the result of objective and subjective evaluation. Second, using Cronbach's alpha value and ANOVA, it is checked whether measured data of design value and subjective measured factors gathered for interactivity evaluation provide meaningful result or not. Third, establishing design variable level by K-mean grouping method, it is analyzed whether each group has statistically meaningful influence. Fourth, result and model of design variable defined by using I-type quantification analysis. It is figurable how

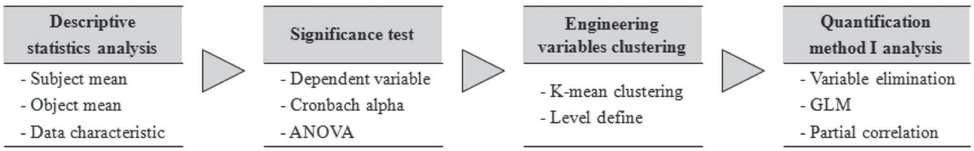


Figure 4. Analysis process of this study.

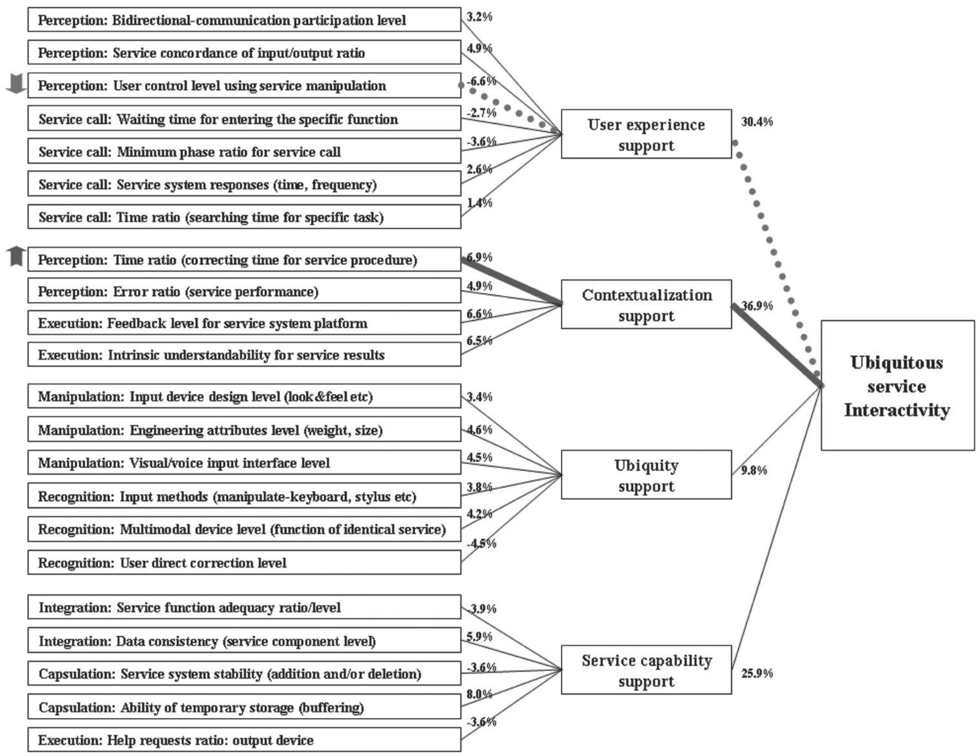


Figure 5. U-home service case study results.

each design variable level influence interactivity by analysis, and the size of influence power between design variables which affect interactivity as deducing partial correlation is determined.

Figure 5 is the result of computing the importance of the factors deduced by the experiment. Since the results of the importance of each dependent variable are different from each other, estimated value of each descriptive variables were converted into normalized value. By the result of the modeling, the most important variable of the ubiquitous service interactivity properties was contextualization support attributes. 'Perception: time ratio (correcting time for service procedure)' metrics had the largest importance as a dependent variable, and 'perception: user control level using service manipulation' metrics was a dependent variable containing negative effects.

4 CONCLUSION AND DISCUSSION

This study has suggested metrics for ubiquitous services that are experienced in real life, and introduced the concept of ubiquitous service interactivity which consist of contextualization, ubiquity, user experience and service capability. For validation of the developed metrics, we applied the proposed metrics to evaluation of the interactivity of U-home service. Through statistical analysis of evaluation data, a subjective satisfaction model of interactivity was developed with the various components of interactivity of ubiquitous service using I-type quantification analysis. As a result, the 22 main components and the relationships between each main component and their sub-components were identified.

As referred to previously, there have been many attempts to develop design guideline, metrics, and evaluation techniques for ubiquitous computing, but there were no contents for user's evaluation on the service. Most of research methods concerning the ubiquitous environment are the same as conventional ones in the IT era. Therefore, they cannot fully reflect the characteristics of ubiquitous services. To select a space-time, that is necessary for the realization of ubiquitous services among many space-times that users come to confront and to provide services appropriate to the characteristics of each space-time, we need an evaluation methodology that comprehends the importance of space-time and reflects its characteristics. The conventional user evaluation techniques cannot satisfy this condition. It is because the conventional techniques have focused on evaluating the quality of specific applications (Bing, 2001).

The suggested model can be used to analyze the potential users of the service by evaluating the interactive level of ubiquitous service. Its application to the stage of service development can be helpful in collecting the demands of potential service users and apprehending their levels. In a situation where the boundary between product and service is disappearing, we may extend the utility evaluation concept of "everyday product" to the evaluation concept of interactivity.

REFERENCES

Aarts, E. & Marzano, S. 2003. *The new everyday: views on ambient intelligence*. Rotterdam: Uitgeverij 010 Publishers.

Abowd, G.D. 1998. Software design issues for ubiquitous computing. In *Proceedings of VLSI (Very Large Scale Integration) System Level Design, IEEE Computer Society Workshop*. USA.

Abowd, G.D. & Mynatt, E.D. 2000. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction* 7(1): 29–58.

Bing, J.W. 2001. Developing a consulting tool to measure process change on global teams: the global team process questionnaire TM. In *Proceedings of Academy of Human Re-source Development Conference*. U.K.

Blaine, A.P., Karim, A. & Bashar, N. 2005. Keeping ubiquitous computing to yourself: A practical model for user control of privacy. *International Journal of Human-Computer Studies* 63: 228–253.

Greenfield, A. 2006. *Everyware: the dawning age of ubiquitous computing*. Berkeley: New Riders.

Hull, R., Reid, J. & Kidd, A. 2002. Experience design in ubiquitous computing. *HP Laboratories Bristol*: HPL-2002–115.

Iqbal, R., Sturm, J., Kulyk, O., Wang, J. & Terken, J. 2005. User-centered design and evaluation of ubiquitous services. In *Proceedings of the 23rd annual international conference on Design of communication: documenting & designing for pervasive information*, 138–145.

Jo, W.D., Lee, K.J., Lee, H.K., Kwon, O.B., Kim, K.G. & Lee, E.J. 2006. *Ubiquitous paradigm and u-society*. Seoul: Jin-han M&B Inc.

Langheinrich, M. 2002. A privacy awareness system for ubiquitous computing environments. In *Proceedings of the 4th international conference on Ubiquitous Computing*. 237–245.

Lindenberg, J., Pasman, W., Kranenborg, K., Stegeman, J. & Neerincx, M.A. 2007. Improving service matching and selection in ubiquitous computing environments: a user study. *Personal and Ubiquitous Computing* 11: 59–68.

Lyytinen, K. & Yoo, Y. 2002. Issues and challenges in ubiquitous computing. *Communications of the ACM* 45(12): 63–65.

Resnick, P. 2002. Beyond bowling together: Sociotechnical capital. In John M. Carroll (ed.), *Human-computer interaction in the new millennium* 2002. NY: ACM Press.

Siewiorek, D.P. 2002. New frontiers of application design, *Communication of the ACM(CACM)* 45: 79–82.

Takemoto, M., Sunage, H., Tanaka, K., Matsumura, H. & Shinohara, E. 2002. The Ubiquitous Service-Oriented Network (USON) an approach for a ubiquitous world based on P2P technology. In *Proceedings 2nd International Conference on Peer-to-Peer Computing*. 17.

Williams, J.R. 2004. *Developing performance support for computer systems: a strategy for maximizing usability and learnability*. Boca Raton: CRC Press.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluating the effectiveness of haptic feedback on a steering wheel for FCW

Jaemin Chun, Seunghwan Oh & Sung H. Han
Department of Industrial and Management Engineering, POSTECH, Pohang, Republic of Korea

Gunhyuk Park, Jongman Seo & Seungmoon Choi
Department of Computer Science and Engineering, POSTECH, Pohang, Republic of Korea

Kyunghoon Han & Woochul Park
Ergonomics Research Team, Hyundai Motor Company, Seoul, Republic of Korea

ABSTRACT: The effectiveness of a haptic warning on a steering wheel in an automobile was evaluated in terms of Forward Collision Warning (FCW). An effective warning system is critical for safe driving, especially when the driver is not paying attention to the road. Haptic feedback warning systems have attracted interest because they do not require any extra visual load; warnings are transmitted to the driver using tactile sensation and are not interfered by ambient noise around the vehicle. Twenty-four subjects participated in the study. Experiment was conducted in a virtual reality environment that simulated collision situations involving FCWs. Performance measures that were evaluated in the FCW experiment included the driver's reaction time to the warning signal. The results show that a haptic FCW on the steering wheel can play a supportive role in preventing a collision. We conclude that haptic warning on a steering wheel could be useful for FCWs.

Keywords: haptic feedback system, forward collision warning, blind spot warning

1 INTRODUCTION

Constant visual attention on the road is important for safe driving. However, a driver's attention is often distracted by audio, navigation, and driver information systems. The additional visual load incurred while using such systems can increase the risk of collision (Ranney et al. 2000).

Issuing a warning signal via a visual display can be problematic as this can simply serve as an additional visual distraction. Audio is commonly used in warning signals; however, warnings issued in this way cannot be transferred effectively when the driver is surrounded by ambient noise (Ryu et al. 2010). Moreover, when audio signals are already in use in the vehicle, the meaning of the warning signal can be confusing in an urgent situation.

A warning signal that utilizes haptic feedback can offer an effective alternative. Via direct contact, haptic feedback can deliver information effectively and in a manner that the driver can respond to intuitively. In addition, the information is transmitted only to the driver; the passengers remain unaware of the warning.

Recently, the automobile industry has begun using haptic feedback in warning systems. Examples include warning drivers of a potential forward collision, as well as blind spots and lane departure. In most instances, the haptic feedback is transmitted via the steering wheel, seat or pedals. Researchers have evaluated the effectiveness of haptic warning signals in various situations such as FCW, LDW (Lane Departure Warning), and BSW (Blind spot warning) (Lerner et al. 1997, Lee et al. 2004, Griffiths et al. 2005, Navarro et al. 2007). Campbell et al. (2007) summarized the results of existing studies and empirically evaluated the effectiveness of various proposed warning systems. However, some combinations of warning location and warning situation have not yet been fully examined (Table 1). We do not know whether it is effective to provide a signal on the steering wheel for FCW and BSW.

Table 1. Results of evaluation of haptic feedbacks on collision warning systems.

	FCW	LDW	BSW
Seat	Good	Fair	Fair
Steering wheel	–	Good	–

Haptic feedback on pedals was excluded due to the potential disturbance that might occur while manipulating the pedals (Campbell et al. 2007). An additional problem with haptic feedback on pedals is that the amplitude of the signal can be seriously affected by the thickness and material of the sole of the driver's shoe. By evaluating the effectiveness of a combination of haptic feedbacks in warning situations not previously studied, the effectiveness of warning signals can be compared qualitatively.

In our study, we assessed haptic feedback on the steering wheel in FCW situation.

Figure 1. Example of secondary task device.

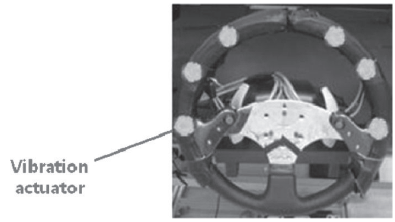


Figure 2. Haptic steering wheel (Circles are locations of vibration actuators for haptic warning).

2 METHODS

2.1 *Apparatus*

A driving simulator was developed to represent events surrounding FCW in a virtual environment. A 50-inch monitor and two 23-inch monitors were used to ensure a seamless representation of the driving environment. A 7-inch touch screen was used for a secondary task (Fig. 1).

Six vibration actuators (3 on the left, 3 on the right) were embedded in the steering wheel to deliver sufficient amplitude of haptic feedback (Fig. 2). All 6 vibration actuators were activated at the same time (frequency 100 Hz; amplitude, 0.025 mm; and periodic envelope with 0.4 sec on-time and 0.1 sec off-time). Pilot tests were conducted to determine the specification of the vibration feedback.

Participants wore headphones that played white noise to prevent sound cues being generated from haptic feedback. The actual size and layout of the seat and steering wheel were considered when designing the seating buck. The distance between the seating buck and steering wheel was adjustable to reduce driver discomfort.

2.2 *Participants*

Briefly, two different age groups, 30–40 yrs and 50–60 yrs, were assessed in each warning system. Since the ability to detect haptic sensations reduces with age (Verrillo et al. 2002), we aimed to determine the effect of age on driver effectiveness in the FCW experiment. For the age group of 30–40 yrs, the participants consisted of 12 males, with an average age of 39.3. The average age of the older group was 56.0 (12 male participants).

2.3 *Experiment design and scenario for FCW*

The independent variables were haptic warning and age group. Dependent variables evaluated were performance and preference measures.

The performance measures were the rate of collision prevented and reaction time. We divided the number of successful decelerations (without collision) by the total number of collision events to calculate the rate of collision prevented. The reaction time was defined as the time that elapsed between deceleration of a preceding vehicle (with a brake light on) and manipulation of the brake pedal.

The preference measures examined were usefulness of the proposed warning signal and overall driver satisfaction. The participants were also asked to consider the inconvenience of the proposed haptic warning when evaluating overall driver satisfaction.

In the experiment, a dual-task paradigm was used to distract participants while driving. As a primary task, the participants were asked to follow a car that maintained a speed of 80 km/sec. In the secondary task, they were required to continuously enter 7 digits that appeared on a touch screen (Fig. 1). From pilot test, we found that the participants were distracted for about 10 sec while completing the secondary task. Whenever a participant makes consecutive errors while entering digits, we determined that the participant was not concentrating and hence excluded the data.

While conducting the secondary task, the vehicle in front randomly started to decelerate and the participants were instructed to manipulate the brake pedal to avoid possible collision. Each participant repeated the FCW event 20 times. The warning signal was transmitted 4 sec prior to collision. Once the warning signal was transmitted, it maintained until TTC (Time To Collision) exceeded 4 sec.

3 RESULTS & DISCUSSION

The initial speed of participants' vehicles and the average distance between vehicles did not show a significant difference based on independent variables. Therefore, the forward collision events can be assumed to have occurred under similar driving conditions.

In the 30–40 yrs age group, the reaction time was significantly shorter in the presence of the haptic warning. Similar results were observed in the older age group (Fig. 3). When the haptic warning was present, participants were more quickly able to recognize that the car in front was decelerating. A shorter reaction time was associated with a decrease in the braking distance (7.36 m, 30–40 yrs; 4.89 m on average, 50–60 yrs).

The number of collisions significantly decreased with haptic warning (Fig. 4). However, the effectiveness of the haptic warning was relatively small in the older age group. Generally, the absolute threshold of haptic sensation increased with age. Overall, this resulted in a lower detectability of haptic warning in the older age group.

The results of preference evaluation showed that participants found the haptic warning useful and satisfactory (Fig. 5).

Participants in the younger group favored haptic feedback more than the older group (the score of preference measures greatly increased when haptic warning is given). Some participants commented that although they found haptic feedback strange at first, they soon became familiar with the stimuli.

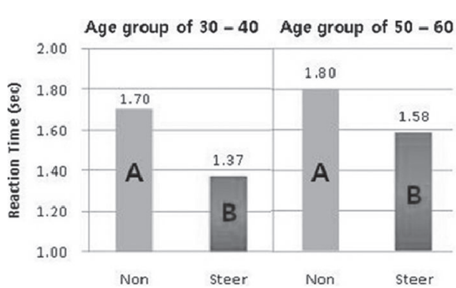


Figure 3. Reaction time.

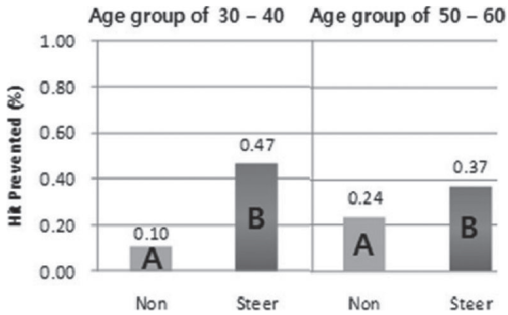


Figure 4. The rate of hit prevented.

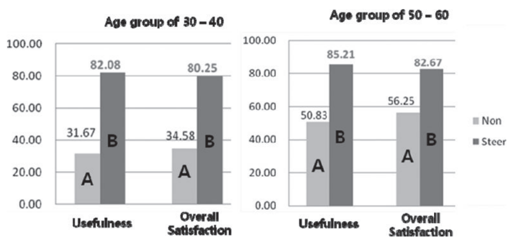


Figure 5. Results of evaluation of preference measures (FCW).

4 CONCLUSION

In this paper, we evaluated haptic warnings on the steering wheel for FCW, which has been studied in previous studies. The risk of collision was significantly decreased in the presence of the haptic warning, and the effectiveness of the warning was consistent across the entire age band evaluated in this study. Moreover, participants felt the warning useful and were satisfied with its performance. The performance of the younger group in particular was better when haptic warnings were given.

Based on our results, we conclude that haptic feedback can be effectively used in FCW. The methods proposed in this study could be used to evaluate other types of warnings for FCW. Further study is needed to find a robust specification of haptic feedback suitable for older age groups. We plan to conduct additional research on the value of haptic feedback on BSW. Also, we will consider other driving situations such as heavy or speeding traffic.

ACKNOWLEDGEMENT

This work was supported by Mid-career Researcher Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (MEST) (No. 2010-0000364).

REFERENCES

Campbell, J.L., Richard, C.M., Brown, J.L. & Marvin. 2007. *Crash warning system interfaces: Human factors insights and lessons learned*. NHTSA technical report.

Griffiths, P.G. & Gillespie, R.B. 2005. Sharing control between humans and automation using haptic interface: Primary and secondary task performance benefits. *Human Factors* 47(3): 574–590.

Lee, J.D., Hoffman, J.D. & Hayes, E. 2004. Collision warning design to mitigate driver distraction, *CHI 2004*, April 24–29 (2004).

Lerner, N.D., Kotwal, B.M., Lyons, R.D. & Gardner-Bonneau, D.J. 1997. *Preliminary human factors guidelines for crash avoidance warning devices*. NHTSA.

Navarro, J., Mars, F. & Hoc, J.M. 2007. Lateral control assistance for car drivers: A Comparison of motor priming and warning systems. *Human Factor* 49(5): 950–960.

Ranney, T., Mazzae, E., Garrott, R. & Goodman, M. 2000. *Driver distraction research: Past, present, and future*. NHTSA.

Ryu, J., Chun, J., Park, G., Choi, S. & Han, S.H. 2010 *Vibrotactile feedback for information delivery in the vehicle. IEEE Transactions on Haptics*, 06 Jan. 2010.

Verrillo, R.T. 2002. Effect of aging on the subjective magnitude of vibration. *Somatosensory & Motor Research* 19(3): 238–244.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluation of human factors in the modification of video display units in the main control room

Chih-Wei Yang, Tzu-Chung Yenn, Tsung-Chieh Cheng & Ming-Huei Chen
Nuclear Instrumentation Division, Institute of Nuclear Energy Research, Taoyuan, Taiwan, ROC

Tung-Ming Wu
Department of Nuclear Engineering, Taiwan Power Company, Taipei City, Taiwan, ROC

Chiuhsiang Joe Lin
Department of Industrial and Systems Engineering, Chung-Yuan Christian University, Taoyuan, Taiwan, ROC

ABSTRACT: This study evaluated the influences of the Video Display Unit (VDU) design change on the performance of operator in the Main Control Room (MCR). The VDU in the Nuclear Power Plant (NPP) provides operators with information supporting situation assessment, monitoring and detection of disturbances, response planning, and response implementation. These characteristics of design represent a trend toward making displays more immediately meaningful to personnel by mapping display representations to important aspects of plant processes and to underlying cognitive mechanisms, such as perceptual processes and mental models. In general, poor implementation of VDU can lead to human performance problems. Some human performance issues associated with display design include the following: information overload; limited display area; lack of overview of plant status, appropriateness of display formats and page arrangements to operator tasks; and improper integration and presentation of numerical data. To validate the effect of VDUs design change on operating errors in the MCR, this present research adopted an experimental design to develop different experimental treatments. According to the experimental results, information requirements are identified; that is, display formats are developed to communicate the information to the operators. The success of representational aids was discussed as being a function of the characteristics of the plant, the operators, and the representation through which the operators view and act on the plant. All three are essential parts of the overall system. Some research support was found for the representation principles; however, most of the studies could not be generalized due to the use of simple systems and tasks. The conclusions of this study could not only be implemented by the HSI designers of a MCR in the advanced NPP but also could be generalized to the extent that other digital workstation similar to the MCR.

Keywords: video display unit, main control room, nuclear power plant, human system interface, human performance

1 INTRODUCTION

This study evaluated the influences on operators' signal-detection performance in the modification of Video Display Unit (VDU) monitoring tasks within the Nuclear Power Plant (NPP) control room. The NPP Main Control Room (MCR) designs have progressed in the last thirty years. One of the most significant changes in MCR designs has been the increasing use of computers in plant monitoring and control (Chuang & Chou, 2005).

A control room uses computer driven displays for information presentation and plant control, as opposed to using hardware components. Increased computerization has made it possible to provide an interactive Human System Interface (HSI) through VDUs to monitor and control most operations (Stubler, O'Hara & Kramer, 2000).

With these computerization increasingly taking over routine plant control tasks, operator actions will be less skill-based and rule-based but more knowledge-based. Operator training will require less focus on building skills and rules for actions but greater focus on general ability to understand how systems function and the ability to respond flexibly when faced with system problems.

Human operators in a hazardous and complex automated system are required to monitor a large number of information sources and to reach a decision relevant to achieving a system goal efficiently and safely (Parasuraman et al., 2000; Hollnagel & Bye, 2000). When everything works smoothly, human monitors are just spare parts, but when there is anything wrong, the subsequent diagnosing and decision tasks impose a large amount of mental workload and stress on the human operator during abnormal operating procedures.

In the last decades, VDUs have become widespread for routine use, in almost all work spaces (Aarås, Horgen, Bjørset, Ro & Thoresen, 1998). Punnet & Bergquist (1997), in their review of epidemiological studies of VDU work, found that VDU work indicated higher risk of neck, shoulder, arm, wrist and hand musculoskeletal illness compared with non-VDU work. Of all the hazards associated with VDU work, visual problems are the most widespread. During the last twenty years in which VDUs have been in general use there have been several hundred studies which have confirmed the link between VDU work and the visual problems reported by users (London Hazards Centre, 1993).

Supervisory control by utilizing VDU is widely implemented in many industries. Human operators are usually assigned to monitor several information sources continuously or periodically to looking for signals that require further actions. For each information source, there may be different types of information associated with a different signal frequency. Supervisory controllers must allocate their attention over these sources with appropriate sampling strategies and frequencies followed by state identification that will suggest appropriate action (Moray, 2003).

Early studies (Ware, Baker & Shelon, 1964; Tulga & Sheridan, 1980; Yeh & Wickens, 1988) tested the relationship between overload condition of workload and task performance. Their results showed that subjective perception of mental workload increased with amounts of resource investment and with demands of working memory.

Tullis (1988) found that the number of character groups and the number of items influenced the controllers' search time to extract information from a static alphanumeric display. As the average size of the groups and the number of items increased, search time tended to increase. Other studies suggested that there existed an optimal display design in which appropriate arrangement of information and amount could improve controller's performance and reduce workload (Moray, 1979, 1988; Tsang & Wilson, 1997; Bujas et al., 1999).

Most of these studies were performed within a single display framework. A particular case of supervisory control is the building monitoring system where the operator is brought to watch several displays containing video camera information. In this situation, supervisory controllers monitor every nook and cranny of the building from displays without any auditory or visual cues.

The advancement of technology has made multiple screens on a single display available to accommodate several camera views. It seems that the more camera views, the safer the system, assuming that the human operator is able to watch as many camera views as provided. However, it is anticipated that the monitoring performance under these situations may degrade, but it is not clear what that might entail?

To investigate the influences on operators' signal-detection performance in the modification of VDU monitoring tasks within the NPP control room, this present research adopted a within subject experimental design to develop different experimental treatments based on two types of display modes included the consistent mode and the mixed mode. The next section introduced the experimental design for reducing human errors by enhancing the VDU design strategy.

2 METHOD

2.1 *Experimental design*

To investigate the influences on operators' signal-detection performance in the modification of VDU monitoring tasks within the NPP control room, an experiment was conducted using a within-participants design with thirteen participants participating in the simulated scenario.

All of them participated in two types of display modes included consistent mode and mixed mode. The display of consistent mode was constituted by two the same VDUs which represented normal size message (normal size icon and 12 pt font size). Moreover, the mixed mode VDUs was constituted by two different VDUs, one represented normal size message, another represented small size message (0.6 times of normal size icon and 9 pt font size).

2.2 *Tasks*

During the experiment, participants would execute the detect-and-hit operation with the VDUs to deal with the monitor tasks. The icons on the VDUs light in blue or red light represent system status. While the icon is lighting red light, meaning that an alarm has occurred, the participants must hit this icon by mouse.

Moreover, icon lighting in blue means the system status is normal; therefore, the participants have to detect this icon but do not need to hit it.

In each VDU, the icons light in blue and red both flash 90 times, that is, the participants should hit 180 times totally while detecting the icons in red during a specific display modes. The icon flash three seconds one time, hence the participant might miss to hit if they did not detect the icon lighting in red timely.

2.3 *Apparatus*

The display of VDUs in advanced MCR was simulated using a computer program developed using Macromedia Flash Player 8 and Microsoft Office Access. Two 19-inches LCD provide a graphical interface for monitoring the reactor system status. Each subject could execute his hitting actions using a mouse. Subject's Critical Flicker Fusion frequency (CFF) was measured with Lafayette Flick Fusion Control 12021.

2.4 *Dependent variables*

There are five dependent variables in the experiment:

Frequency of miss: the frequency of miss when monitor the displays as an indicator of objective performance accuracy. The frequency of miss meant the number of icons lighting in red that the participants must hit but they miss.

Reaction time: The start time of the icon lighting in red to the end of the participant hitting the icon was recorded by the experimental system. Therefore, the value of the reaction time is between 0 and 3000 millisecond.

Situation awareness: Situational awareness was assessed using the 3-D situational awareness rating technique (SART; Taylor, 1990). The 3-D SART is an experimentally validated, retrospective measure which requires participants to rate themselves on three dimensions that included attentional Demands (D), attentional Supply (S), and Understanding (U) immediately following task performance. The ratings on each of the three dimensions are combined into a single SART value according to a formula (Selcon et al., 1992): Situation Awareness = U− (D−S).

Subjective performance: Subjective performance was rated on a scale of 1–10, with 1 being the worst performance, and 10 being optimal performance.

Critical flicker fusion: CFF measures the minimal number of flashes of light per second at which an intermittent light stimulus no longer stimulates a continuous sensation. It is an effective measure of visual fatigue (Knulst & Kraaykamp, 1998). CFF change was the CFF difference before and after the experiment. A drop in CFF reflects a drop in the sensory perception function, attributable to a decrease in alertness (Hashimoto, 1963).

Subjective visual fatigue: The subjective visual fatigue was measured via a visual discomfort questionnaire. The questionnaire contains nine questions: (1) I have difficulties in seeing. (2) My eyelids are heavy. (3) I feel eye strain. (4) I have burning eyes. (5) I have a strange feeling around the eyes. (6) I have itching eyes. (7) I feel numb. (8) I feel dizzy. (9) I have a headache. Response was rated on a 7-point scale, with 1 representing "extremely disagree", and 7 representing "extremely agree".

2.5 *Participants*

Thirteen engineering undergraduate and graduate students with normal or corrected to 20/20 visual acuity and full-color vision participated for monetary compensation on a voluntary basis. Participant age range was 23–25 years. Most participants were right-handed. These participants were required to have experience using a personal computer. The participants have received different periods of training for the experimental task and necessary operating procedures. To minimize the effects of individual differences, participants were randomly assigned to one display mode.

2.6 *Experimental procedure*

The experiment was conducted in a secluded and quiet room. The experimental procedure included seven steps: (1) A document material used in the experiment was given to the participants and observers explain the experiment content and related special note. (2) Let the participants acquaint the simulated display of VDUs in NPP and practice the operation until skilled.

(3) After practicing, let participants take a break until the visual fatigue restoring. (4) To measure the CFF of the participant before experiment, and then require monitoring the VDUs and starting to perform the detecting tasks in the formal experiment. (5) To measure the CFF of the participant after experiment and fill out the first SART, subjective performance and visual fatigue questionnaire, then take a break about 30 minutes.

(6) Repeat step (4) and (5) but monitor another display mode of VDUs. (7) To finalize the experiment. The participants were required to try their best to perform the tasks in the required time and avoid any mistakes during the formal experiment. The entire experiment for a subject took 70–80 minutes.

2.7 *Hypotheses*

It was hypothesized that participants will have lower situation awareness in monitoring mixed mode VDUs than in monitoring consistent mode VDUs.

Furthermore, subjective and objective visual fatigue will significantly increase while monitoring with mixed mode VDUs. Corresponding with the visual fatigue increasing, it was expected that participants' frequency of miss and reaction time will get increases. It was also hypothesized that subjective performance will decrease while monitoring with mixed mode VDUs.

3 RESULTS

The findings indicated significant relationships between frequency of miss, reaction time, and subjective performance. Frequency of miss was positively correlated with reaction time ($r = 0.77$, $p < 0.01$). Subjective performance was negatively correlated with frequency of miss ($r = –0.46$, $p < 0.05$) and reaction time ($r = –0.52$, $p < 0.01$). These results demonstrated that there is a significant correlation between objective performance (frequency of miss and reaction time) and subjective performance.

To examine the hypothesis, the study conducted a one-way ANOVA. The analytic results are shown in Table 2 and Table 3. The ANOVA results revealed that there were significant group differences. First, there was a significantly different frequency of miss between the consistent mode VDUs and mixed mode VDUs ($F_{1, 12} = 26.67$, $p < 0.01$).

Table 1. Descriptive statistics and correlations.

Variable	Mean	Std.	Miss	RT	SA
Frequency of miss	20.81	10.28			
Reaction time	1798.62	137.72	0.77***		
Situation awareness	5.14	1.80	0.14	0.17	
Subjective performance	5.35	2.15	–0.46**	–0.52***	–0.03

Note: $^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$.

Table 2. The ANOVA results of dependent variables.

	Display mode		
Dependent variable	Consistent mode	Mixed mode	F-value
Frequency of miss	15.69	25.92	26.67***
Reaction time	1749.53	1847.72	16.52***
Situation awareness	5.60	4.69	3.42*
Subjective performance	6.00	4.69	4.00*
CFF	0.83	1.74	3.04

Note: $^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$.

Table 3. The ANOVA results of subject visual fatigue.

	Display mode		
Subjective visual fatigue	Consistent mode	Mixed mode	F-value
Difficulties in seeing	3.08	3.77	5.86**
Eyelids are heavy	4.08	4.77	6.94**
Eye strain	5.00	5.62	4.52*
Burning eyes	3.00	3.31	0.60
Strange feeling	3.69	3.85	0.10
Itching eyes	3.23	3.39	0.12
Numb	3.62	3.23	2.54
Dizzy	3.00	3.00	0.00
Headache	2.46	2.77	0.45

Note: $^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$.

The frequency of miss in operating mixed mode VDUs ($M = 25.92$) was significantly higher than that in consistent mode VDUs ($M = 15.69$). Second, there was a significantly different reaction time between the consistent mode VDUs and mixed mode VDUs ($F_{1, 12} = 16.52$, $p < 0.01$). The reaction time of mixed mode VDUs ($M = 1847.72$) was significantly higher than that of consistent mode VDUs ($M = 1749.53$).

Third, there was a significantly different situation awareness between the consistent mode VDUs and mixed mode VDUs ($F_{1, 12} = 3.42$, $p < 0.1$). The situation awareness of consistent mode VDUs ($M = 5.60$) was significantly higher than that of mixed mode VDUs ($M = 4.69$).

Fourth, there was a significantly different subjective performance between the consistent mode VDUs and mixed mode VDUs ($F_{1, 12} = 4.00$, $p < 0.1$). The subjective performance of consistent mode VDUs ($M = 6.00$) was significantly higher than that of mixed mode VDUs ($M = 4.69$). Finally, although display mode did not have a significant influence on CFF ($F_{1, 12} = 3.04$, $p = 0.11$), the participants who monitor mixed mode VDUs seemed to have higher average reduction of CFF than the participants who monitor consistent mode VDUs.

Table 3 shows that display modes of VDUs were significant on some questions of subjective visual fatigue, they are 'Difficulties in seeing', 'Eyelids are heavy', and 'Eye strain'. The participants who monitor mixed mode VDUs have higher subjective visual fatigue than the participants who monitor consistent mode VDUs.

4 DISCUSSION

From these results, the mixed mode VDUs are found significantly negative impacts on dependent variables while performing a monitoring task. First, supervisory controllers will have lower situation awareness while monitoring the mixed mode VDUs than monitoring the consistent mode VDUs. There are some differences between this study and Huang, Hwang, Yenn, Yu, Hsu & Huang (2006).

In Huang et al. (2006), the simplified VDU design will enhance operators' situation awareness. The VDU layout designs in this study are fixed. The critical point in the present study is to compare the consistency of icon size and font size between two modes VDUs. The results provide evidences to VDU designers that utilizing consistent VDU designs could maintain operators' situation awareness.

Second, supervisory controllers will have significantly more frequency of miss and longer reaction time while monitoring the mixed mode VDUs than monitoring the consistent mode VDUs. No operation errors are allowable in NPPs.

In the past studies (Lin, Yenn & Yang, 2010a, 2010b), the reaction time are found getting increases in complex design. However, there are no significant differences between complex and simple designs in these two studies. In this study, the reaction time is also found getting increases in the mixed modes VDUs. Furthermore, this study found that there is significantly more frequency of miss in the study. The analytic results of subjective performance are found to support the above results gathered from objective performance.

Despite there are no significant differences in CFF, there were significant differences on some questions of subjective visual fatigue between different display modes of VDUs, they are 'Difficulties in seeing', 'Eyelids are heavy', and 'Eye strain'. This implies that the mixed mode VDUs could substantially degrade monitoring performance.

There is a risk associated with this performance decrement if human operators are put responsible for monitoring information on the VDUs, especially if such application is to be used in high risk and critical environment. Of all the hazards associated with VDU work, visual problems are the most widespread. The results of visual fatigue gathered in this study are similar to several hundred studies which have confirmed the link between VDU work and the visual problems reported by users (London Hazards Centre, 1993).

5 CONCLUSIONS

The advanced NPPs adopt the modernized fully integrated digital design, the instrumentation and control (I&C) systems are digitized, and multiplexing network techniques and soft control are extensively adopted. Therefore, I&C systems have the great potentials for enhancing plant safety, reliability, operability, availability, and maintainability.

Currently digital I&C system design are usually implemented at various vendor facilities. From the analytic results in this study, the mixed mode VDUs are found significantly negative impacts on frequency of miss, reaction time, situation awareness, subjective performance, and subjective visual fatigue while performing a monitoring task.

No operation errors are allowable in NPPs, so this study strongly suggested that a configuration management plan and a vendor integration plan are necessary to resolve the complexity and inconsistency involved in the I&C system implementation, such as the consistent icon size and font size. A well defined configuration management plan and vendor integration plan could be beneficial to the safe operation of NPPs. The conclusions of this study could not only be implemented by the HSI designers of a MCR in the advanced NPP but also could be generalized to the extent that other digital workstation similar to the MCR.

REFERENCES

Aarås, A., Horgen, G., Bjørset, H.-H., Ro, O. & Thoresen, M. 1998. Musculoskeletal, visual and psychosocial stress in VDU operators before and after multidisciplinary ergonomic interventions. *Applied Ergonomics* 29(5), 335–354.

Bujas, Z., Ajdukovic, D., Szabo, S., Mayer, D. & Matutinovic, Z. 1999. Perception of the form of stimulus increment as a method in assessment of the psychophysical relationship. *Perceptual and Motor Skills*. 89: 509–531.

Chuang, C.-F. & Chou, H.-P. 2005. Investigation of potential operation issues of human-system interface in Lungmen nuclear power project. *IEEE transactions on nuclear science*. 52(4): 1004–1008.

Hashimoto, K. 1963. Physiological meaning of the Critical Flicker Frequency (CFF) and some problems in their measurement—theory and practices of the flicker test. *Japanese Journal of Industrial Health* 5: 563–578.

Hollnagel, E. & Bye, A. 2000. Principles for modeling function allocation. *International Journal of Human-Computer Studies* 52(2): 253–265.

Huang, F.-H., Hwang, S.-L., Yenn, T.-C., Yu, Y.-C., Hsu, C.-C. & Huang, H.-W. 2006. Evaluation and comparison of alarm reset modes in advanced control room of nuclear power plants. *Safety Science* 44(10): 935–946.

Knulst, W. & Kraaykamp, G. 1998. Trend in leisure time: forty years of research on reading in the Netherlands. *Poetics* 26(1): 21–42.

Lin, C.J., Yenn, T.-C. & Yang, C.-W. 2010a. Automation design in advanced control rooms of the modernized nuclear power plants. *Safety Science* 48(1), 63–71.

Lin, C.J., Yenn, T.-C. & Yang, C.-W. 2010b. Evaluation of operators' performance for automation design in the fully digital control room of nuclear power plants. *Human Factors and Ergonomics in Manufacturing & Service Industries* 20(1): 10–23.

London Hazards Centre 1993. *VDU work and the hazards to health.* London Hazards Centre Trust Ltd.

Moray, N. 1979. *Mental workload: Its theory and measurement*. New York: Plenum Press.

Moray, N. 1988. Mental workload since 1979. *International Reviews of Ergonomics* 2(2): 123–150.

Moray, N. 2003. Monitoring, complacency, scepticism and eutactic behaviour. *International Journal of Industrial Ergonomics* 31(3): 175–178.

Parasuraman, R., Sheridan, T.B. & Wickens, C.D. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans* 30(3): 286–297.

Punnet, L. & Bergquist, U. 1997. Visual display unit work and upper extremity muskuloskeletal disorders. A review of epidemiological findings. In: *Sweden-Ergonomic Expert Commite Document No. 1*, National Institute for Working Life, Solna, Sweden.

Selcon, S.J., Taylor, R.M. & Shadrake, R.M. 1992. Multi-modal cockpit warnings: Pictures, words, or both? In *Proceedings of the Human Factors Society 36th Annual Meeting, Human Factors Society*, Santa Monica, CA, pp. 57–61.

Stubler, W.F., O'Hara, J.M. & Kramer, J. 2000. *Soft controls: technical basis and human factors review guidance* [NUREG/CR-6635]. Washington, DC: U.S. Nuclear Regulatory Commission.

Taylor, R.M. 1990. Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations,* AGARD-CP-478, Advisory Group for Aerospace Research and Development, Neuilly-sur-Seine, France, pp. 3–1 to 3–17.

Tsang, P. & Wilson, G. 1997. Mental workload. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*. New York: Wiley.

Tulga, M.K. & Sheridan, T.B. 1980. Dynamics decisions and workload in multitask supervisory control. *IEEE transactions on Systems, Man, and Cybernetics* 10(5): 217–232.

Tullis, T.S. 1988. A system for evaluating screen formats: Research and application. In H.R. Hartson & D. Hix (Eds.), *Handbook of humancomputer interaction*. Norwood, NJ: Ablex Publishing. pp. 214–286.

Ware, R.J., Baker, R.A. & Shelon, R.W. 1964. Effect of increasing signal load on detection performance in a vigilance task. *Perceptual and Motor Skills* 18: 105–106.

Yeh, Y.Y. & Wickens, C.D. 1988. Dissociation of performance and subjective measures of workload. *Human Factors* 30(1): 111–120.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The link between systemizing, one characteristic of the male brain, and persistence in acquiring skills in sports

Yujiro Kawata & Kazusa Oki
Juntendo University Graduate School of Health and Sports Science, Chiba, Japan

Yasuyuki Yamada & Ikuyo Kaneko
Juntendo University School of Health and Sports Science, Chiba, Japan

Motoki Mizuno, Sumio Tanaka & Masataka Hirosawa
Juntendo University Graduate School of Health and Sports Science, Chiba, Japan
Juntendo University School of Health and Sports Science, Chiba, Japan

ABSTRACT: Persistence in the acquisition of sport skills is one of the essential psychological factors for becoming high-level players or athletes. In connection with this theme, many researchers have paid much attention to their drive to repetitive and long-term practice, and have tried to find out their psycho-social background. However, few studies have dealt innate (inherent) psychological factors or interactions of inherent and acquired psycho-social factors. Here, we paid attention to "empathizing and systemizing", which are known as inherent psychological driving forces. The purpose of this study is to examine whether the driving forces of "empathizing and systemizing" are linked to persistence in acquiring skills in sports or not. We conducted a questionnaire investigation with 166 (114 male and 52 female) Japanese university students who belong to competitive sports clubs. The results showed that systemizing is strongly linked to persistence in the acquisition of sports skills compared with empathizing.

Keywords: persistence in the acquisition of sports skills, male and female brain, systemizing, empathizing

1 INTRODUCTION

Persistence in the acquisition of sport skills is one of the essential psychological factors for becoming high-level players or athletes. Many researchers have paid much attention to their drive to repetitive and long-term practice in terms of various perspectives and theories (Duda & Hall, 2001; Roberts, 2001). However, there are few studies considering innate (inherent) psychological factors or interactions of inherent and acquired psycho-social factors which make athletes to persist acquiring sport skills.

Reviewing previous studies, developmental neuropsychologists insisted that there are two primary states of cognition which emerge in infancy: cognitive style of understanding social-psychological agency, and that of understanding physical events (Cosmides & Tooby, 1994; Wellman & Inagaki, 1997). These two cognitive styles were named based on the corresponding academic field; "folk psychology" (an interest in how people work) and "folk physics" respectively (an interest in how things work) (Dannett, 1987; Baron-Cohen, 1997; Baron-Cohen & Wheelwright 1999).

Baron-Cohen (2002) proposed the Empathizing-Systemizing theory of sex difference in cognition. These concepts are derived from folk psychology and folk physics. Empathizing is defined as the drive to identify emotions and thoughts in others and to respond to these appropriately. This provides understanding another's intent or mental state. Systemizing is defined as the drive to analyze and construct systems (Baron-Cohen et al., 2002). This provides understanding rules in the systems and shows how systems behave.

Baron-Cohen et al. (2003) proposed some types of systemizing including sports techniques (motoric systems), in which sport skills were regarded as a system. If a player had high systemizing, the player would explore the rule or principals in the physical motion or structure of motor skills. As an example, a tennis player might realize that, if the player changes stroke with the top-spin, the ball bounces to the right. This brings the thought: systemizing may be a drive to persistence in skill-acquisition.

According to the general population studies, there is evidence of sex differences in drives for cognition. Females show superiority in empathizing and males show superiority in systemizing. Thus, empathizing is one of the characteristics of so-called "female brain" and systemizing is one of the characteristics of so-called "male brain". Moreover, systemizing and empathizing are related to pre-natal testosterone exposure (Auyeung et al., 2006; Champman, E et al., 2006). Therefore, they are considered inherent psychological factors.

This study examined the hypothesis that empathizing and systemizing are linked to persistence in acquiring skills in sports.

2 METHOD

2.1 *Participants*

The participants comprised of 166 Japanese university students, including 114 males and 52 females, who belong to sports clubs of badminton, baseball, basketball, futsal, gymnastics, handball, judo, kendo, rugby, soccer, softball, squash, swimming, tennis, track and field, volleyball, and so on. The mean age of the participants in this study was 20.1 (SD = 0.64, range 19–22), the mean duration of sports events was 8.9 (SD = 3.93, range 1–18). The number and distribution of their competitive level of sports were as follows: municipality level (N = 38, 23%), prefectural level (N = 42, 25%), district level (N = 43, 26%), national level (N = 46, 25%), international level (N = 2, 1%).

2.2 *Measure*

2.2.1 *Individual profiles*

We examined such participant's individual profiles as sex, chronological age, main sports activities and duration of sports events.

2.2.2 *Systemizing Quotient (SQ)*

We used the Japanese adult version of the Systemizing Quotient (SQ) to assess individual trait for systemizing (Baron-Cohen, 2002; Wakabayashi et al., 2006). Systemizing is defined as the drive to analyze and construct systems (Baron-Cohen, 2002). The SQ was designed to be a self-reporting 60-item questionnaire with four alternative answers for each question. The subjects indicate how strongly they agree with each statement indicating if they (a) definitely agree, 4ghtly agree, (c) slightly disagree, or (d) definitely disagree. The scoring for each item gives a value of +2 to (a), +1 to (b), 0 to (c) and (d). A value of +2 indicates that participants have strong trait for systemizing drive.

2.2.3 *Empathizing Quotient (EQ)*

We used the Japanese adult version of the Empathizing Quotient (EQ) to assess individual trait for empathizing (Baron-Cohen, 2002; Wakabayashi et al., 2006). Empathizing is defined as the drive to identify emotions and thoughts in others and to respond to these appropriately (Baron-Cohen, 2002). The EQ was designed as a self-reporting 60-item questionnaire with four alternative answers for each question. The subjects indicate how strongly they agree with each statement indicating if they (a) definitely agree, (b) slightly agree, (c) slightly disagree, or (d) definitely disagree. The scoring for each item gives a value of +2 to (a), +1 to (b), 0 to (c) and (d). A value of +2 indicates that participants have strong trait for empathizing drive.

2.2.4 *Persistence in the Acquisition of Skills in Sports Questionnaire (PASSQ)*

We used the "Persistence in the Acquisition of Skills in Sports Questionnaire (PASSQ)" to assess individual persistence in sports skills which we made in this study using frequently-used psychological methods in the development of the questionnaire. The PASSQ was designed to be a self-report questionnaire. The PASSQ is a 10-item questionnaire with five alternative answers for each question. The subjects indicate how frequently they agree with each statement indicating if they (a) always (100%), (b) often (75%), (c) sometimes (50%), (d) rarely (25%), or (e) never (0%) persist. The scoring for each item gives a value of +5, +4, +3, +2, +1 respectively. A value of +5 indicates that they frequently persist in the acquisition of skills in sports.

2.3 *Procedure*

We carried out a questionnaire investigation into the participants in January 2010 after informed consent from the participants and permission of the coaches. The participants were voluntarily asked to complete the survey in quiet classrooms at the university. The questionnaire was distributed to the participants and collected by authors and trained research assistants at that time. The questionnaire took about 30 minutes to complete.

All statistical analysis in this study was carried out using the Predictive Analytics Software (PASW) 18.0 for Windows.

3 RESULTS

3.1 *Development of the PASSQ*

Firstly, we conducted a pilot test with an open-response question, "What kind of persistence do you have in your sports domain?" to 80 Japanese university students who belong to sports clubs in order to clarify the existence of persistence in the acquisition of skills in sports. As a result, we identified various descriptions which are related to athletes' persistence in the acquisition of skills in sports.

Next, we made up items which reflected the persistence in the acquisition of skills in sports with two sports psychologists, one psychiatrist who was well-experienced in counseling for athletes, and two graduate students who were majoring in sports psychology. Then we asked 166 participants in this study to answer the questions.

Then, we analyzed the responses using the item-total correlation analysis in order to select relevant items. As a result, we confirmed that two items were not appropriate in evaluating persistence in acquiring sports skills because they did not show significant correlation coefficients. Therefore, we excluded them from the questionnaire.

Then, we carried out the factor analysis of the selected items in order to evaluate the validity of constructive concepts. The results showed that there are almost the same two factor structures in both methods of a varimax and a promax rotation method. We came to a judgment that the varimax rotation method was better in this study because there was no significant correlation between elucidated factors. In this study, we adopted the following frequently-used common criteria: eigenvalues (first solution) is over 0.10, factor loading is over 0.40.

Table 1 shows the results of the factor analysis using the principal method with varimax rotation. Factor 1 included items which were related to behavioral persistence, and the factor 2 included items which were related to analytic persistence. Thus, we named them "Persistence in action (PAC)" and "Persistence in analyzing (PAN)" respectively.

Moreover, we evaluated the internal consistency of above two factors using the alpha coefficient analysis. The results showed a moderate value of the alpha coefficient, $\alpha = .75$ and $\alpha = .67$ respectively (Table 1).

3.2 *Effects of systemizing and empathizing on persistence in the acquisition of skills in sport*

First, we examined the relationship between EQ, SQ and PASSQ using a correlation on analysis (Table 2). The results showed that there was significant correlation between EQ and PASSQ ($r = .21$, $p < .05$), and there was significant correlation between SQ and PASSQ ($r = .28$, $p < .001$).

Next, we examined the relationship between EQ, SQ and the subscales of PASSQ. The results showed that there was no significant correlation between EQ and PAC ($r = .15$, $p <$ n.s.), while there was significant correlation between EQ and PAN ($r = .21$, $p < .05$). Also, there was significant

Table 1. Results of the factor analysis of the PASSQ.

		Factor loadings	
		PAC	PAN
4. I practice to acquire a motor skill I pursue on holidays.		0.75	0.15
2. I forget all about the time when I practice a motor skill.		0.73	0.14
1. I repeat practicing until I acquire a motor skill.		0.63	0.13
5. I look out for kinematic structure of a motor skill.		0.52	0.31
10. I focus on whether or not I can do the motor skills.		0.08	0.73
8. When I see a motor skill among top athletes or players, I want to know their motor structure.		0.17	0.64
6. I realize the difference between my good motor skill and my bad motor skill.		0.20	0.50
7. I pay attention to a smaller movements rather than a larger movements.		0.10	0.41
	Eigen value	3.01	1.45
	Cumulative contribution ratio	23.11	42.09
	Alpha coefficients	0.75	0.67

PASSQ: Persistence in the Acquisition of Skills in Sports Questionnaire.
PAC: Persistence in Action, PAN: Persistence in Analysis.

Table 2. Results of the correlation analysis between EQ, SQ and PASSQ.

	PAC	PAN	PASSQ
EQ	.15	.21 *	.21 *
SQ	.29 ***	.12	.28 ***

EQ: Empathizing Quotient, SQ: Systemizing Quotient.
PAC: Persistence in Action, PAN: Persistence in Analysis.
PASSQ: Persistence in the Acquisiion of Skills in Sports Questionnaire*p < .05, **p < .01, ***p < .001.

correlation between SQ and PAC ($r = .29$, $p < .001$), while there no was significant correlation between SQ and PAN ($r = .12$, p = n.s.).

4 DISCUSSION

In this study, we examined the hypothesis that drives for empathizing and systemizing, which are inherent psychological factors, are linked to persistence in acquiring skills in sports.

Firstly, we formed a new way to measure and assess persistence in acquiring sports skills using frequently-used psychological methods in the development of the questionnaire. As the results show, we clarified that persistence in acquiring skills in sports was structured following at least two components: "Persistence in Action (PAC)" and "Persistence in Analyzing (PAN)". This made it possible for measuring persistence in acquiring skills in sports.

Next, we examined the relationship between systemizing, empathizing and persistence in acquiring sports skills using the correlation analysis.

The results indicated that empathizing and systemizing may be linked to persistence in acquiring sports skills. Moreover empathizing was related to Persistence in Analyzing (PAN), and systemizing was related to Persistence in Action (PAC).

According to the value of correlation coefficients, systemizing may be a stronger predictor than empathizing to persistence in acquisition of sports skills. Of course, the interpersonal quantitative balance of systemizing and empathizing should also be considered in order to understand persistence in acquiring sports skills in details, using five brain types framework (Extreme empathizing type, Empathizing type, Balance type, Systemizing type and Extreme systemizing type) of E-S theory that previous study presented (Baron-Cohen et al., 2005). In addition, empathizing and systemizing are one of the characteristics of the male brain and is related to pre-natal testosterone exposure (Auyeung et al., 2006). Therefore, this is the first evidence that some inherent psychological factors may have an effect on persistence in acquiring sports skills.

We focused on only persistence in sports skill acquisition in this study and presented systemizing as one predictor of persistence in sports skill acquisition. However, empathizing also plays an important role in the sports domain, such as understanding another's intent and mental state which are required, especially in team sports. Therefore, the role of empathizing should also be considered in relation to various phenomena which occur in the sports domain.

5 CONCLUSION

We came to the following two conclusions from the results of this study:

1. Persistence in acquiring skills in sports was structured at least following two components: "Persistence in Action" and "Persistence in Analyzing".
2. Systemizing is strongly linked to persistence in acquiring skills in sports, compared with empathizing.

ACKNOWLEDGEMENT

This work was funded by the grant from the Special Research Funds for Doctoral Students at Juntendo University Graduate School of Health and Sports Science. We would like to thank all the participants who took part in the study.

REFERENCES

Auyeung, B., Baron-Cohen, S., Champman, E., Knickmeyer, R., Taylor, K. & Hackett, G. 2006 Fetal testosterone and the Child Systemizing Quotient (SQ-C). *European Journal of Endocrinology* 155: 123–130.

Baron-Cohen, S. 1997. Are children with autism superior at folk physics? In H.M.Wellman & K. Inagaki (eds.), *Children's theories. New direction for child development series*. Jossey-Bass Inc.

Baron-Cohen, S. 2002. The extreme male brain theory of autism. *Trends in Cognitive Science* 6: 248–254.

Baron-Cohen, S. & Wheelwright, S. 1999. 'Obsessions' in children with autism or Asperger syndrome. Content analysis in terms of core domain of cognition. *British Journal of Psychiatry* 175: 484–490.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N. & Wheelwright, S. 2003. The systemizing quotient: an investigation of adult with Asperger syndrome or high functioning autism, and normal sex

differences. *Philosophical Transactions of the Royal Society, Series B. Special issue on "Autism: Mind and Brain"* 358: 361–374.
Baron-Cohen, S., Knickmeyer, R. & Belmonte, M. 2008. Sex deferences in the brain: Inplications for explaining autism. *Science* 310 (4): 819–823.
Champman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K. & Hackett, G. 2006. Fetal testosterone and empathy: Evidence from the 'Reading Mind in the Eyes' Test. *Journal of Social Neuroscience* 1(2): 135–148.
Cosmides, L. & Tooby, J. 1994. Origins of domain specificity: The evolution of functional organization. In L.A. Hirschfeld & S.A. Gelman (eds.), *Mapping the Mind: Domain specificity in cognition and culture*: 85–116. Cambridge: Cambridge University Press.
Dannett, D. 1987. *The intentional stance*. Cambridge, Massachusetts: The MIT Press/Bradford Books.
Duda, J.L. & Hall, H. 2001. Achievement theory in sport: Recent extension and future directions. In R.N. Singer, H.A. Hausenblas, and C.M. Janelle (eds.), *Handbook of sport psychology* 2: 417–443. New York: Wiley.
Roberts, G.C. 2001. Understanding the dynamics of motivation in physical activity: The influence of achievement goals on motivational processes, In Roberts, G.C. (eds.), *Advances in Motivation in Sport and Exercise*: 1–50. Champaign, IL: Human Kinetics.
Wakabayashi, A., Baron-Cohen, S. & Wheelwright, S. 2006. Individual and gender differences in empathizing and systemizing: Measurement of individual differences by Empathy Quotient (EQ) and Systemizing Quotient (SQ). *Journal of Japanese Psychology* 77(3): 271–277 (in Japanese).
Wellman, H.M. & Inagaki, K. 1997. The emergence of core domains of thoughts: Children's reasoning about physical, psychological and biological phenomena. Jossey-Bass Inc.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Development of the multiple roles map program for the enhancement of self-understanding

Yasuyuki Yamada
Juntendo University School of Health and Sports Science, Chiba, Japan

Yasuyuki Hochi, Yujiro Kawata & Kazusa Oki
Juntendo University Graduate School of Health and Sports Science, Chiba, Japan

Motoki Mizuno & Masataka Hirosawa
Juntendo University School of Health and Sports Science, Chiba, Japan
Juntendo University Graduate School of Health and Sports Science, Chiba, Japan

ABSTRACT: The purpose of this study was to examine the validity of the Multiple Roles Map (MRM) program as new methods to intervene in the enhancement of self-understanding among Japanese university students. Through this study, three conclusions were confirmed. 1) Japanese university students had mainly five kinds of multiple roles and the spillover, compensation, and segmentation were observed in the interfaces between these roles; 2) Participants of the MRM program reported mainly eight kinds of awareness about their multiple roles; 3) The validity of the MRM program as interventional methods to enhance university athletes' self-understanding was confirmed.

Keywords: multiple roles, university students, MRM, role theory, self-understanding

1 INTRODUCTION

1.1 *Necessity of the deepening self-understanding among university students*

Deepening of one's own understanding is an effective approach to cope with stress, to carve a career by oneself, and to improve social skills for university students. Actually, enhancement of self-understanding is becoming one of the aims of methods and/or programs in the scene of clinical counseling, career counseling, and organizational development for university students in Japan (e.g. Hochi et al., 2009; Yamada et al., 2009). Hence, in order to accomplish a more effective approach to analyze and discuss in their self-understanding, developing new methods were needed.

1.2 *Awareness of the multiple roles*

As many university student-athletes have an athletic, an academic, and a social roles (Adler & Adler, 1991), and as many workers are required to perform both roles in work and family (e.g. Edwards & Rothbard, 2000; Hanson et al., 2006; Staines, 1980; Yamada et al., 2009), people usually have some multiple roles in their daily life. Generally, such people may burden from multiple roles and conflicts in roles included negative effects from role to role through playing multiple roles. As an approach to cope with these problems, enhancement of awareness toward multiple roles and transactions between them will be effective. If we get some awareness toward features of multiple roles and their function, more effective mental management will be possible through controlling them. Therefore, this study developed the Multiple Roles Map (MRM) program which enables them to understand their self and condition.

1.3 *Multiple Roles Map (MRM) program for the new interventional methods*

The aim of the MRM program is supporting people to take aware of their condition. Awareness of having multiple roles will be necessary and developing the way of analyzing and discussing condition will make it possible to deepen peoples' self-understanding around their roles. For the accomplishment of this purpose, participants of the MRM program were needed to think and draw a number, names, contents, interactions, and priority of their multiple roles. Especially, drawing the spillover, compensation, and segmentation with detailed episodes is main task in the MRM program.

In the multiple role theories, Staines (1980) originally proposed three competing mechanisms for understanding the relationship between multiple roles: spillover, compensation, and segmentation. Concretely, spillover is the transfer of characteristics of the affects, values, skills, and behaviors from one domain to the other domain (Edwards & Rothbard, 2000; Staines, 1980). Compensation was what is lacking in one role is sought after and obtained in the other role (Staines, 1980; Hanson et al., 2006). For example, sometimes we feel dissatisfaction through playing one role and try to compensate it by participation in the other role. Segmentation refers to the separation of one role and the other role (Staines, 1980; Hanson et al., 2006). In this pattern, both roles were independent from each other and there were no interactions between them. If we can become aware and control them below the level of consciousness through the MRM program, more practical self-mental management will be possible.

1.4 *Guideline of the MRM program*

Therefore, this study established the procedure of the MRM program as follows experimentally. In our MRM program, firstly, the facilitator had to get informed consent from all participants. Secondly, the participants considered their multiple roles on the basis of the MRM sheet (Figure 1). In this A3 sheet, roles in university (academic role), club activity (club role), their part time job (part time job role), family (family role), human relationships (human relationship role), and private time (no role) were expressed as faces. If they don't have some roles, they delete them. If they have other roles, they add faces on the sheet. Thirdly, participants described the contents of role in detail beside each face. Fourthly, they received explanations about the spillover, compensation, and segmentation and considered about them. Next, they showed a negative spillover by red path between roles and write detailed episodes beside it. Similarly, they expressed a positive spillover by blue path, the compensation by green dot-path, and the separation by black double-headed path with their episodes. Fifthly, they prioritized these roles. Finally, they considered the resulting features of MRM sheet by themselves and described their awareness in white space.

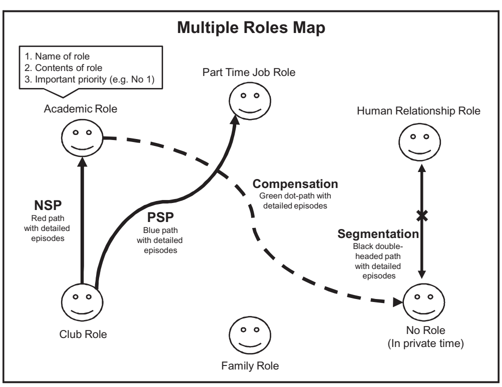


Figure 1. Guideline for drawing the MRM sheet.

1.5 *Purpose and hypotheses*

The purpose of this study was to examine the validity of the MRM program as new methods to intervene in the enhancement of self-understanding among Japanese university students. For this purpose, this study established and verified the following three hypotheses:

Hypothesis 1: It is appropriate for university students to deal with five kinds of roles and no role in private time in the original MRM sheet.

Hypothesis 2: Every perspective in role interface (spillover, compensation, and segmentation) is supported.

Hypothesis 3: Participation in the MRM program contributes to enhancement of self-understanding.

2 METHODS

2.1 *Participants*

This study was carried out with 70 (male = 41, female = 29) Japanese university students in 2009. They described their awareness on the A4 sheet after the intensive program experience of the MRM. The valid data for analysis were 64 (male = 38, female = 26). The average age of respondents was 20.27 (SD = 0.94) years old.

2.2 *Procedures of investigation*

All the participants of this study agreed to the informed consent. This study was carried out in the lecture room of the university. Participants were required to give information and futures relating to their multiple roles in the MRM sheets on the basis of above guideline.

2.3 *Procedure of analysis*

Firstly, in order to examine the hypothesis 1, this study counted the number of samples that added other roles to the original five roles in the MRM sheet. Next, in order to examine the hypothesis 2, this study counted the number of the paths of the

Table 1. The number of samples with PSP, NSP, compensation, and segmentation.

Interface between mutiple roles and dynamic direction	PSP		NSP		Compensation		Segmentation	
	n	%	n	%	n	%	n	%
Academic→Club	9	14.3	5	7.9	3	4.8	10	15.9
Academic→Family	2	3.2	3	4.8	4	6.3	15	23.8
Academic→Human relationship	1	1.6	0	0.0	6	9.5	6	9.5
Academic→No role	5	7.9	8	12.7	9	14.3	7	11.1
Academic→Part time job	12	19.0	6	9.5	1	1.6	13	20.6
Club →Academic	6	9.5	12	19.0	2	3.2	10	15.9
Club →Family	6	9.5	12	19.0	5	7.9	11	17.5
Club →Human relationship	5	7.9	3	4.8	6	9.5	11	17.5
Club →No role	5	7.9	6	9.5	7	11.1	0	0.0
Club →Part time job	11	17.5	8	12.7	1	1.6	8	12.7
Family→Academic	1	1.6	3	4.8	0	0.0	15	23.8
Family→Club	1	1.6	1	1.6	0	0.0	11	17.5
Family→Human relationship	1	1.6	1	1.6	2	3.2	22	34.9
Family→No role	2	3.2	5	7.9	2	3.2	18	28.6
Family→Part time job	0	0.0	3	4.8	1	1.6	15	23.8
Human relationship→Academic	2	3.2	3	4.8	0	0.0	6	9.5
Human relationship→Club	6	9.5	2	3.2	0	0.0	11	17.5
Human relationship→Family	4	6.3	1	1.6	3	4.8	22	34.9
Human relationship→No role	5	7.9	2	3.2	1	1.6	16	25.4
Human relationship→Part time job	6	9.5	4	6.3	1	1.6	15	23.8
No role→Academic	7	11.1	4	6.3	1	1.6	7	11.1
No role→Club	4	6.3	4	6.3	0	0.0	0	0.0
No role→Family	4	6.3	4	6.3	0	0.0	18	28.6
No role→Human relationship	11	17.5	5	7.9	3	4.8	16	25.4
No role→Part time job	6	9.5	9	14.3	0	0.0	3	4.8
Part time job→Academic	4	6.3	11	17.5	0	0.0	13	20.6
Part time job→Club	5	7.9	5	7.9	2	3.2	8	12.7
Part time job→Family	10	15.9	7	11.1	5	7.9	15	23.8
Part time job→Human relationship	11	17.5	5	7.9	4	6.3	15	23.8
Part time job→No role	10	15.9	12	19.0	12	19.0	3	4.8

positive spillover, negative spillover, compensation, and segmentation in the whole sample. Finally, in order to examine the hypothesis 3, this study analyzed free descriptions about considerations toward the MRM sheet and awareness through the MRM program.

3 RESULTS

3.1 *What kinds of roles do the university students have?*

In order to examine the hypothesis 1 that it is appropriate for university students to deal with five kinds of roles and no role in private time in the original MRM sheet, this study counted the number of samples that adopted other roles. According to the making up of the MRM sheets, no one adopted other roles. Namely, almost all roles surrounding university students could be covered by setting an academic, a club, a part-time job, a family, a human relationship roles and no role in private time. Hence, the hypothesis 1 was supported.

3.2 *Are the spillover, compensation, and segmentation confirmed?*

In order to examine the hypothesis 2 that every perspective in role interface (spillover, compensation, and segmentation) is supported, this study counted the number of samples with each perspective. The Table 1 was the spreadsheet of MRM data. According to these results,

Table 2. Some of the free descriptions related to self-understanding.

f1: Awareness of the existence of multiple roles
- I surprised that I had so many roles.
- I didn't know about my roles well.
- I have been spending a lot of times under unconsciousness about my roles.

f2: Awareness of the central role
- All my roles were influenced by the condition in the club role.
- I realized how much I prioritized the club role.
- I realized that playing a role of the part-time job affected other roles.

f3: Awareness of the positive and/or negative roles
- The private and family roles were necessary for me because these roles compensated my negative feeling in other roles.
- Playing an academic role may be a source of stress for me.
- Although playing club role produced may negative spillovers to other roles, I like it. So it was precious role for me.

f4: Awareness of the positive and/or negative spiral of performing roles
- I realized the best relationships between my roles.
- My fulfilling daily life could be achieved by not only playing club role but also playing academic role.
- If I can't spend refresh time in the private roles, many negative spillovers will cause in the other roles and I can't adjust the society.

f5: Awareness of the positive and/or negative personality features
- I tend to worry about negative spillovers directed to my family role in my home.
- I think I tend to bring my emotion from role to role and it was difficult for me to be pragmatic about everything.
- I realize I am an optimist.
- My personality features tend to switch when changing a role.

f6: Awareness of a suitable role
- I think I can act like oneself in playing a club role.
- The family role was important for me because I think playing a family role is suitable for me.
- I tend not to reveal myself easily in human relationship with my friends.

f7: Awareness of the isolated role
- My family role was separated from other roles.
- I concentrated on doing my club role and I couldn't think about other roles then.

f8: Awareness of the necessity of the role control
- I'd like to have a clearer image about positive spillover and it also better to rediscover my feelings when I face some negative situation in one role.
- I'd like to find the positive spillover more frequently in my daily life and to change my negative spillover to a positive one.
- Since I think a fulfilling school life is my first priority, I'd like to strike the proper balance between multiple roles.
- I think that I shouldn't drop out from my club activity now.

three perspectives were confirmed and hypothesis 2 was supported. Moreover, in our data, frequency of the *Academic to Part Time Job PSP* (19.0%), *Club to Academic NSP* (19.0%), *Club to Family NSP* (19.0%), *Part Time Job to No Role NSP* (19.0%), *Part Time Job to No Role Compensation* (19.0%), and segmentation between family role and human relationship (34.9%) were especially high.

3.3 *Does the MRM program enhance participants' self-understanding?*

In order to examine the hypothesis 3 that participation in the MRM program contributes to the enhancement of self-understanding, this study analyzed free descriptions about considerations toward the MRM sheet and awareness through the MRM program. In this study, collected free descriptions were analyzed by a psychiatrist and two psychologists belonging to the academic laboratory. As the results, free descriptions were categorized to eight factors of awareness relating to self-understanding. The names of eight factors and some descriptions in each factor were shown in the Table 2. Concretely, eight factors were named as "awareness of the existence of multiple roles (f1)," "awareness of the central role (f2)," "awareness of the positive and/or negative roles (f3)," "awareness of the positive and/or negative spiral of performing roles (f4)," "awareness of the positive and/or negative personality features (f5)," "awareness of the suitable role (f6)," "awareness of the isolated role (f7)," "awareness of the necessity of the role control (f8)." Hence, the hypothesis 3 was supported.

4 DISCUSSIONS

This study developed the MRM program for the enhancement of self-understanding among Japanese university students and examined its validity. As the results, university students had mainly five roles; academic role, club one, part-time job one, family one, and a human relationship one. Hence, adopting these roles in the original MRM sheet was appropriate for them. Next, we counted the number of samples with the spillover, compensation, and segmentation, every perspective in role interface was supported between these roles. Moreover, since some common features of them were emerged from the spreadsheet of MRM data, more detailed analyses were needed from now on. Furthermore, this study clarified that the MRM program led eight kinds of awareness relating to self-understanding.

Taking these results into considerations, it can be said that the validity of the MRM program was confirmed. Therefore, the MRM program is expected to offer an effective approach especially in the fields of stress management, career support, social skill education from now on. On the other hand, it is anticipated that a more sophisticated guideline and tool of the MRM in the near future.

5 CONCLUSIONS

Through this study, three conclusions were confirmed.

1. Japanese university students had mainly five kinds of roles and spillover, compensation, and segmentation were observed in the interfaces between these roles.
2. Participants of the MRM program reported mainly eight kinds of awareness about their multiple roles.
3. The validity of the MRM program as the interventional methods to enhance university athletes' self-understanding was confirmed.

REFERENCES

Adler, P.A. & Adler, P. 1991. *Backboards & blackboards: College athletics and role engulfment*. New York: Columbia University Press.

Edwards, J.R. & Rothbard, N.P. 2000. Mechanisms linking work and family: Clarifying the relationship between work and family constructs. *Academy of Management Review* 25: 178–199.

Hanson, G.C., Hammer, L.B. & Colton, C.L. 2006. Development and validation of a multidimensional scale of perceived work-family positive spillover. *Journal of Occupational Health Psychology* Jul 11(3): 249–265.

Hochi, Y., Yamada, Y. & Mizuno, M. 2009. Effects of Organization Development on the Psychological Aspects among University Students: Focused on the Transformation of the Understanding for their University. *News letter of Human Ergology* 91: 12.

Staines, G.L. 1980. Spillover versus compensation. A review of the literature on the relationship between work and non-work. *Human Resources* 33: 111–129.

Yamada, Y., Mizuno, M. & Hochi, Y. 2009. Does Organizational Development bring the Change of Self-Understanding among University Students? *25th Annual Meeting of the Japanese Association of Industrial and Organizatinal Psychology*: 91–94.

Yamada, Y., Oka, K., Kawata, Y., Mizuno, M. & Hirosawa, M. 2009. Relationship between negative spillover and depression among university athletes. *Juntendo Medical Journal* 55(4): 502–510.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Analysis of irradiation effects of font size, font thickness, and font type

Y.K. Kong, D.M. Kim, C.M. Lim & J.G. Han
I.E. Sungkyunkwan University, Suwon, Gyeonggi-do, South Korea

M.C. Jung
I.I.S.E. Ajou University, Suwon, Gyeonggi-do, South Korea

ABSTRACT: In this study, it was studied to know that contrast of luminance due to irradiation-effect affects legibility in various types, sizes and thickness of font. Thirty subjects were recruited from university students. Who have over 1.0 eyesight. The tested font type was Gothic, Ming-style, and Saemmul. And the font thickness was either normal or bold. We estimated the irradiation-effect of six combinations of font type and font thickness over font size 2~80 pt (1pt = 0.35146 mm). In result, font type ($p = 0.02$), font thickness ($p < 0.001$), font size ($p < 0.001$) were factors which have a statically meaningful effect in legibility for subjective assessment of font size in p-value 0.05. For all of the six combinations at font size 22~24 pt, we could observe that frequency of the white letters on black background (W/B) looked bigger than black letters on white background (B/W) was larger than frequency of the W/B looked similar with B/W and of the white letters on W/B looked smaller than B/W.

Keywords: irradiation, irradiation-effect, font size, font thickness, font type

1 INTRODUCTION

Our society has many written advertisements and written signs, and they deliver important information. That is, key function of letter is delivering information quickly and efficiently (Jung et al., 1993). Therefore, letter must be considered about legibility that precondition of information delivery (Choi et al., 1987).

Legibility means degree of how can we read the letter easily and various studies which were performed from early stage to VDT (Visual Display Terminal) of present have been existed.

The main factors which affect legibility is font size, type of font, thickness of font, space between fonts, contrast of colors, distance of sight, and illumination (Vartabedian, 1971; Hwang et al., 1997; Bernard et al., 2003). Considering Reading activity is practicable by comparing between letter color and background color, contrast of letter colors have a great effect on legibility (Legge, 1990; Knoblauch, 1991). Contrast of color is classified into two main factors; contrast of color tone and contrast of luminance (Legge et al., 1990). Contrast of color tone means because of mixed colors it looks different with actual color and contrast of luminance means because of difference of brightness it looks different with actual size. Through studies about contrast of color and legibility, many researchers report that contrast of luminance have a more effect on legibility than contrast of color tone (Bruce and Foster, 1982; Knoblauch et al., 1988; Legge et al., 1990; Shieh et al., 2000).

Knoblauch et al. (1988) reported that contrast of luminance have a great effect on legibility but contrast of color tone have a slight effect. Accordingly, Legge et al. (1990) analyzed a research of subjects whose eyesight was normal, background color was fixed white, divided luminance of font color into 8 stages then compare contrast of luminance with relation of objective reading rates, the result was that the higher contrast of luminance, the easier to read. So they report that black letters on white background (B/W) is the highest legibility. Also, Shieh and line (2000)'s study about legibility test at 12 combination of color support this Legge et al. (1990)'s research work. In other words, generally, most guidelines about luminance and legibility suggest linear relation that the more contrast of luminance between color of letter and background increase, the more legibility increase (Rivlen et al., 1990; Nielsen, 2000).

Although, many studies about contrast of luminance were performed, influence of irradiation-effect (Park, 1980), phenomenon white object or shape seems to be spread in black background, about subjective measurement variable-one of the measure of legibility—is highly insufficient.

The purpose of this research is study the influence of change of font size, font thickness, and font type, to suggest guideline that can be used to design notifications or signs in terms of optimizing legibility in condition of contrast of luminance (i.e. W/B, B/W).

2 METHODS

2.1 *Subjects*

In this research, fifteen males and fifteen females were recruited from university students in this study. The mean age of all subjects was 22.37 year (±1.96).

Subjects had no surgery or disease of eyes which affect on their eyesight and were not color-weak or color-blind and their eyesight was over 1.0.

Using eyesight-test-chart (3 m), measured the eyesight of subjects and if he or she wears glasses of contact lens, measure the corrected eyesight (Mclean, 1965).

2.2 *Equipment*

We built darkroom with rebar and dark fabric (Figure 1). Inside the darkroom, we made comfortable environment by using white fabric not to stimulate eye.

Experiment screen was suggested resolution of 1280 × 1024 pixel LCD monitor. Distance of eye and LCD monitor was maintained at regular distance by head fixing device with chin rest (Figure 1).

To maintain intensity of illumination, we placed LED lamp (KDT, Model No. R02TD) that can control brightness at ceiling of darkroom.

2.3 *Experimental design*

In condition of 6 combination of font type (Gothic, Ming-style, Saemmul) and thickness (normal, bold) at every font sizes (2~80 pt, 39ea), the experiment was carried out to measure the difference of W/B and B/W as a dependent variable (Table 1).

Figure 1. Equipment picture.
*left: outside of darkroom.
**right: head fixing device.

2.4 *Procedures*

First, we got subjects' agreement of experiment, and provide experiment information—methods, purpose of experiment, and procedure—using brochure. Before the experiment, we measured subjects' eyesight by using eyesight-test-chart, experiment host suggested 5 numbers and 5 Landholt ring, and if subjects make a match more than 3 times, we recorded subjects' eyesight (Mclean, 1965).

After the measurement of eyesight, to know factors which affect on irradiation-effect, we suggested white letter on black background (W/B) and black letter in white background (B/W) on monitor using MS Power Point 2007 as shown in Figure 2.

There were two types of fonts, rectangle shape, we determined that are Ming-style type and Gothic type. The Ming-style type is used general publications, newspaper and magazine. Otherwise, the Gothic type is used many kinds of signs, title and sentence of publications. Also, we determined that are Saemmul type which have non-rectangle shape widely used among young people. Font size was changed 2~80 pt (1 pt = 0.35146 mm), font thickness was either normal and bold. To carry

Table 1. Font thickness and font type.

	Gothic type	Ming-style type	Saemmul type
Normal type	민	민	민
Bold type	**민**	**민**	**민**

Figure 2. LCD serene used in the experiment.
*left: white letter on black background (W/B).
**right: black letter in white background (B/W).

out experiment in same condition, luminance and distance between object and eye were fixed at 600 ± 10 lux and 50 cm, respectively.

All subjects performed subjective assessment selecting bigger letter to between two letters shown LCD monitor (Figure 2).

All experiments proceeded randomly and after each task, to avoid cumulative fatigue. Subjects were provided about 2 minutes break time. After completing 3 tasks among 6 tasks, about 10 minutes break time was provided, continually performed 3 tasks.

2.5 *Analysis methods*

To evaluate the effect of font size, font thickness, and font type in condition of contrast of luminance (i.e. W/B, B/W) on legibility, we analyzed difference of subjective size assessment (i.e. dependent variable; difference of W/B and B/W) on independent variable (i.e. font size, font thickness, font type) by using SPSS 12.0 (statistic analysis program). Also, through analysis of frequency, examine font size-frequency of W/B looked larger than B/W is bigger than frequency of B/W looked larger than W/B or similar-then, performed cluster analysis. Finally, we performed ANOVA about thickness and font type.

3 RESULTS

3.1 *Analysis of variance*

As a result of ANOVA, font thickness ($p < 0.001$), font size ($p < 0.001$) and font type ($p = 0.020$) were factors which have a statically meaningful effect in irradiation-effect. There was no interaction between font size and font thickness ($p = 0.110$).

In result of paired comparison analysis about font thickness and font type, in font thickness, bold type was statically meaningful in subjective assessment of font size. In font type, except Gothic type, Ming-style type and Saemmul type had a statically meaningful difference ($p = 0.006$).

3.2 *Frequency analysis*

3.2.1 *Gothic type*

If font thickness is normal, font size which frequency of W/B looked larger than B/W (Figure 3, top, green line) is higher than frequency of W/B and B/W looked similar (Figure 3, top, red line) and B/W looked larger than W/B (Figure 3, top, blue line) is 22~24 pt. This result is similar to bold type. That is, at the approximately 22~24 pt, frequency of W/B looked larger than B/W heighten (Figure 3, bottom).

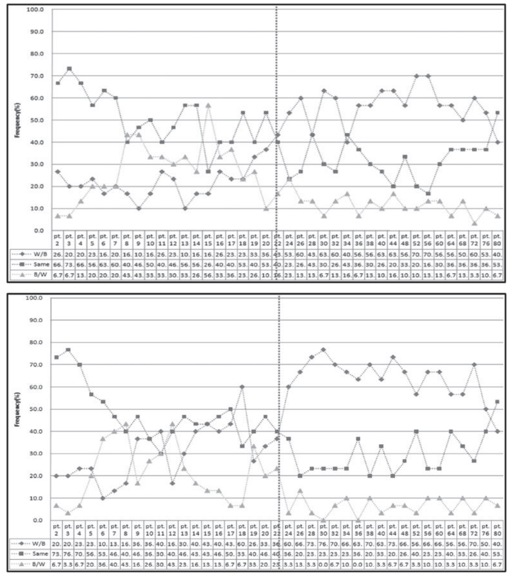


Figure 3. Frequency of subjective size assessment on font thickness.
*Gothic type, normal type (top); bold type (bottom).
**Blue line/diamond: frequency of assessment that B/W looks larger.
***Red/rectangle: frequency of assessment that B/W looks similar with W/B.
****Green/triangle: frequency of assessment that W/B looks larger.

3.2.2 *Ming-style type*

In Ming-style type like Gothic type—both normal type and bold type, font size is 22~24 pt that frequency which W/B looked larger than B/W (Figure 4, top/bottom, green line) is larger than both frequency which B/W looked larger than W/B (Figure 4, top/bottom, blue line) and frequency which W/B looked similar to B/W (Figure 4, top/bottom, red line).

3.2.3 *Saemmul type*

As before, in Saemmul type—both normal type and bold type, font size is 22~24 pt that frequency which W/B looked larger than B/W (Figure 5, top/bottom, green line) is larger than both frequency which B/W looked larger than W/B (Figure 5, top/bottom, blue line) and frequency which W/B looked similar to B/W (Figure 5, top/bottom, red line).

As before, in Saemmul type, in both normal type and bold type, font size appeared approximate 22~24 pt when frequency of W/B looked larger than B/W (Figure 5, top/bottom, green line) is bigger than frequency of B/W looked larger than W/B (Figure 5, top/bottom, blue line) and frequency of W/B looked similar to B/W (Figure 5, top/bottom, red line).

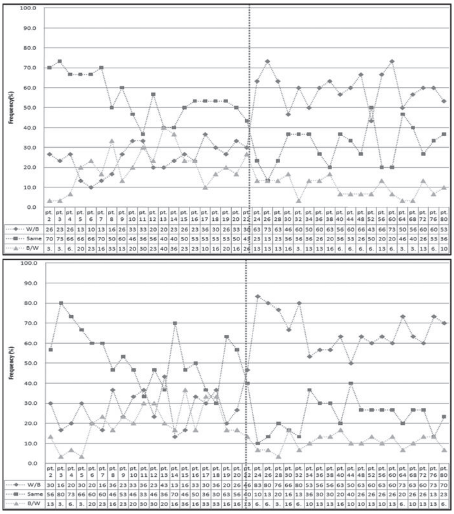


Figure 4. Frequency of subjective size assessment on font thickness.
*Ming-style type, normal type (top); bold type (bottom).
**Blue line/diamond: frequency of assessment that B/W looks larger.
***Red/rectangle: frequency of assessment that B/W looks similar with W/B.
****Green/triangle: frequency of assessment that W/B looks larger.

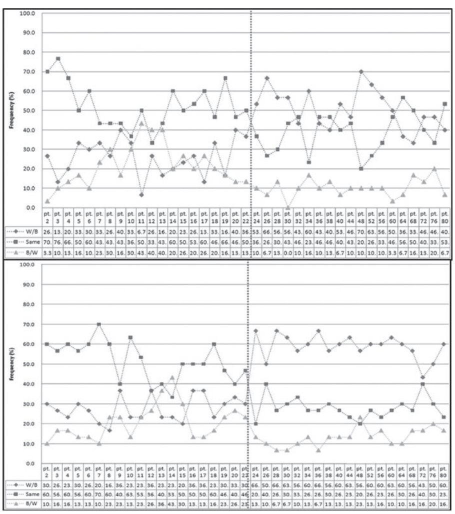


Figure 5. Frequency of subjective size assessment on font thickness.
*Saemmul type, normal type (top); bold type (bottom).
**Blue line/diamond: frequency of assessment that B/W looks larger.
***Red/rectangle: frequency of assessment that B/W looks similar with W/B.
****Green/triangle: frequency of assessment that W/B looks larger.

Table 2. Font thickness and font type.

	Gothic type	Ming-style type	Saemmul type
Normal type	24 pt	24 pt	22 pt
Bold type	24 pt	22 pt	24 pt

In conclusion, every font thickness and every font type, frequency of W/B looked larger than B/W is bigger than frequency of B/W looked larger than W/B and frequency of W/B looked similar to B/W at the font size 22~24 pt.

3.3 *Cluster analysis*

In conclusion of Cluster analysis about font thickness and font type, each W/B looked larger than B/W and B/W looked larger than W/B was grouped. And font size which divide formed groups were suggested, Table 2. The font sizes which divide the two groups were also 22 pt and 24 pt like an analysis of frequency (Table 2).

4 DISCUSSION

Recently, we have seen many words or characters for delivering some information to as. To deliver information, above all, we should read the letters easily. In this situation, as a medium of information delivery, letters about legibility have been studied in ergonomics field. According to Bruce et al. (1982), legibility was increasing in condition difference between brightness of text and background was increasing. That is to say, irradiation-effect generated by difference of brightness has a great effect on legibility.

By studying irradiation-effect which is a factor that had a influence on legibility, this research was to investigate readable font size, font thickness, and font type. Through Analysis of variance and cluster analysis, we can confirm in 22~24 pt font size, irradiation-effect is obvious. At the same time, in font type, we can confirm rectangle shape (Gothic type and Ming-style type) had a greater effect on subjective size assessment than non-rectangle shape (Saemmul type). Font size was not always have a greater effect on subjective size assessment, at some point, when the font size is increased, generally size have an effect obviously. Also, bold type has an effect on subjective size assessment greater than normal type. In font type, it is appeared that rectangle shape font had a greater effect on subjective size assessment than none rectangle shape font. In order of Gothic type, Ming-style type, Saemmul type, former one influenced the most, last one influenced the least. This finding might be explained by familiarity of font type. When the letter getting bigger, frequency of W/B looked larger than B/W was higher, but frequency of that in Gothic and Saemmul type dropped slightly before and after 80 pt because of distance factor (50 cm).

So far, research of irradiation-effect is insufficient. Through this study, we could know the factors that have effect on irradiation-effect, and from now on, we expect this study will be the important foundation of effect research of irradiation-effect in various conditions, and factors.

REFERENCES

Bruce, M. & Forster, J. 1982. The visibility of colored characters on colored backgrounds in view data displays. *Visible Language* 16: 382–390.

Bernard, M.L., Chaparro, B.S., Mills, M.M. & Halcomb, C.G. 2003. Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human Computer Studies* 59: 823–835.

Choi, D.C. & Park, Y.T. 1987. A study on stroke width-to-height ratio and reading distance of the Korean character. *Journal of the Ergonomics Society of Korea* 6(2): 9–18.

Hwang, W.S., Lee, D.C., Lee, S.D. & Lee, J.H. 1997. An experimental study on search speed and error rate according to Korean letter size and font on search task with VDT. *Journal of the Ergonomics Society of Korea* 16: 29–38.

Jung, W.H., Han, J.J. & Chung, C.S. 1993. The effects of hangul letter form on readability, Conference of Psychological Association of Korea, 491–500.

Knoblauch, K., Arditiand, A. & Szlyk, J. 1988. Reading rate depends on luminance contrast but not chromatic contrast, *Investigative ophthalmology and visual science* (suppl.), 29: 44.

Knoblauch, K., Arditiand, A. & Szlyk, J. 1991. Effects of chromatic and luminance contrast on reading, *Journal of the Optical Society of America A* 8: 428–439.

Legge, G., Parish, A. & Lee, H. 1990. Psychophysics of reading: Comparing color contrast and luminance contrast, *Journal of Optical Society of America A* 7(10): 2002–2010.

Mclean, M. 1965. Brightness contrast, color contrast, and legibility. *Human Factors* 7(6): 521–526.

Nielsen, J. 2000. *Designing web usability: The practice of simplicity.* IN: New Riders Publishing, Indianapolis.

Park, K.S. 1980. *Human factors*, Young-Ji publisher.

Rivlin, C., Lewisand, R. & Davies-Cooer, L.R. 2000. *Guidelines for screen design*, Blackwell Scientific Publications, Oxford.

Shieh, K. & Lin, C. 2000. Effects of screen type, ambient illumination, and color combination on VDT visual performance and subjective preference. *International Journal of Industrial Ergonomics* 26: 527–536.

Tinker, M.A. & Paterson, D.G. 1931. Studies of typographical factors influencing speed of reading: VII variations in color of print and background. *Journal of Applied Psychology* 15: 471–479.

Vartabedian, A.G. 1971. The effects of letter size, case, and generation method on CRT display search time. *Human Factors* 13(4): 363–368.

Part VI: Design & ergonomics

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

How exterior form of car influences drivers' judgment of distance between car and anterior as well as posterior objects

Min-Jho Young & Ming-Hsien Hsieh
Da-Yeh University, Changhua, Taiwan

ABSTRACT: Many drivers may feel it difficult to curb park a car quickly because of uncertainty in judging the gap between the car and surrounding obstacles. Such uncertainty arises from structurally induced blind spot around the car. This research aimed to find how exterior form of passenger car influences drivers' distance judgment between the car and anterior as well as posterior objects. Field tests were conducted using five passenger cars and twenty subjects participated in the experiment. The results show that the car shapes do influence driver's judgment of safe distance. Appropriate design of car shape can provide reference to assist drivers in gap judgment.

Keywords: car, shape, distance, judgment

1 INTRODUCTION

Many drivers may feel it difficult to curb park a car quickly because of uncertainty in judging the gap between the car and surrounding obstacles. Such uncertainty arises from structurally induced blind spot around the car. Although nowadays CCD camera and radar have been widely investigated (Aguilar et al. 1996, Burie et al. 1995, Chavand et al. 1997, Handmann et al. 2000, Hicks et al. 1999, Kunert 1999, Kuno & Sugiura 1999, Miyahara 2005, Miyahara 2007, Miyahara et al. 2009, Oizumi et al. 2003, Takahashi et al. 2002) and used to warn drivers the proximity of the car and the obstacles, it would be more beneficial if the driver can directly judge the gap between the car and the obstacles. This is possible if the exterior form of a car is well designed to provide some references to drivers. People normally cannot have correct distance judgment between two objects in open space. Sinai's research (1998) showed that without ground as reference people tended to overestimate the distance between two objects. Even with CCD in car, Young & Chiu (2007) shows that only with appropriate installation to have ground as background can it provide useful reference for judging gap between car and posterior object.

This research hopes to know how exterior forms of passenger cars influence driver's judgment of gap between car and anterior as well as posterior objects. Five passenger cars with different form were tested and safe distances were compared. How the form of front end and rear end influence gap judgment was also investigated.

2 METHOD

2.1 *Subject*

Twenty subjects participated in the experiment, ten males and ten females. All subjects were licensed drivers with age ranging from 20 to 30.

2.2 *Apparatus*

Five passenger cars with different styles were used, including Mitsubishi Lancer, Honda Civic, Nissan March, Nissan Sentra, Renault Twingo, as shown in Figure 1 and Table 1. All five cars were in good condition when experiments were conducted.

As for the anterior and the posterior objects, to avoid the possibility of damaging testing cars in case of subjects' inappropriate maneuver, a plastic column with 20 cm of diameter and 200 cm high, as well as a wood wall of 6 cm thick, 210 cm high and 250 cm wide, were used. Both column and wall were painted in gray.

2.3 *Test site*

To avoid that any surrounding object may provide reference for gap judgment, the experiments were conducted in an open asphalt-paved flat ground, about 50 meters long and 18 meters wide.

2.4 *Procedure*

The experiments were conducted in the daytime between 9 am and 5 pm. Simulated obstacles, column and wall, were put in the center of the testing

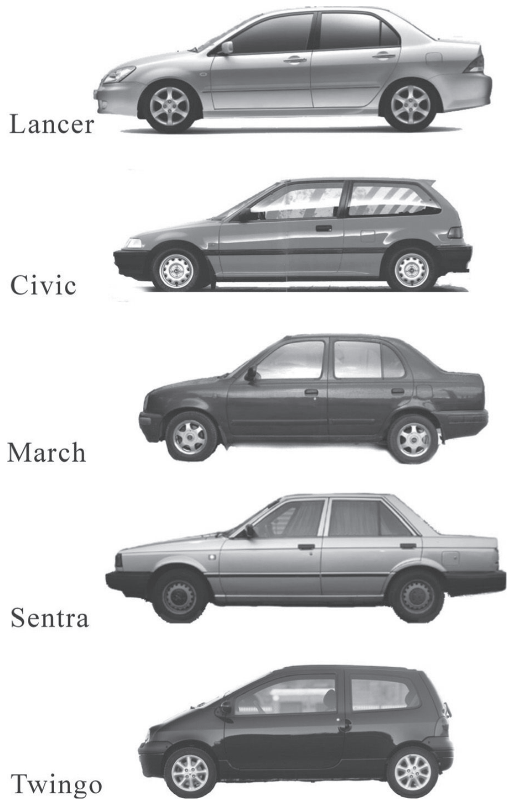


Figure 1. Five testing cars.

Table 1. Information of testing cars.

Model (cm/cm/cm)	Length/Width/ Height (cm)	Wheelbase	Remark
Mitsubishi Global Lancer (2003)	4500/1695/1430	2600	Sedan
Honda Civic (1991)	3995/1680/1330	2500	Hatchback
Nissan March (1995)	4060/1585/1430	2360	Sedan
Nissan Sentra (1991)	4175/1640/1400	2430	Sedan
Renault Twingo (1997)	3425/1630/1410	2345	Hatchback

site alternately. The car was initially positioned at 7 meters away from the simulated object.

Before the experiment, a short briefing regarding testing objectives and procedure was given to the subjects. The subjects were also requested to take a detailed look at the car to have a clear understanding of the forms of front and rear ends. After sitting in the car the subject might adjust the seat to his or her most comfort position.

It is hypothesized that the more the subject can see the far end of hood the better the subject can perceive the gap. A virtual line can be drawn from subject's eyes tangent to the hood, as shown in Figure 2. The magnitude of "a" is considered as an important factor influencing the gap perception. To measure "a", an experimenter, standing alongside the front end, would move a stick from rear end of hood slowly in subject's eye-view direction to the car front until subject in driving position could not see the portion of hood beyond that point. Horizontal distance between that point and the forefront of the car was measured, as shown in Figure 2.

2.5 *Experiment A: Comparison of safe distance of five testing cars*

After getting familiar with the car operation the subject was requested to drive the testing car forward or backward at his or her preferred speed and brought the car to a completely stop when he or she felt the car was at appropriate safe distance away from the objects. The gap between the car and the objects was measured. The subject then moved the car back to its original position and repeated the process. For each experiment every subject had nine trials. From the fourth till the sixth trials the subject would be informed of the testing result of previous trial.

The procedure for moving backward test was similar. The subject must turn his or her head to observe the rear scene without using the rearview mirror.

2.6 *Experiment B: How form of car influences gap judgment*

To investigate how form of front end and rear end influences gap judgment, a special experiment was conducted using Mitsubishi Lancer. The testing procedure was similar to the previous ones except, during the test, a cardboard was put at two locations, the lower side of windshield

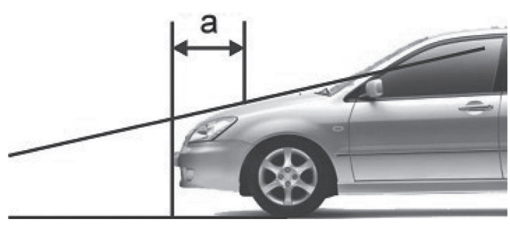


Figure 2. Virtual line from driver's eyes and tangent to hood.

and the forefront of the car, respectively. While at the lower side of windshield, the height of the cardboard was adjusted until the subject just could not see the hood. While at the forefront of the car, the height of the cardboard was adjusted until its top coincided with a virtual line drawn from subject's eyes and tangent to the hood, as shown in Figure 3. Similar procedure was applied for moving backward test.

3 RESULTS

3.1 *Learning effect*

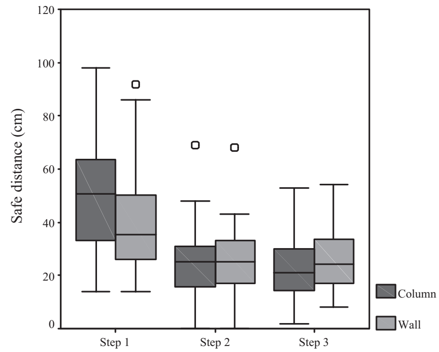
For each testing car each subject had nine trials. To see if practice made difference, the first three, the middle three, and the final three trials were averaged and analyzed using SPSS. Results are shown as box plots in Figure 4. For each box plot, the top of the box represents the 75th percentile, the bottom of the box represents the 25th percentile, and the line in the middle represents the 50th percentile. The whiskers, lines that extend out the top and bottom of the box, represent the highest and lowest values that are not outliers or extreme values. Circles beyond the whiskers represent outliers, values that are between 1.5 and 3 times the interquartile range, and extreme values, values that are more than 3 times the interquartile range. It indicates that the safe distance associated with the first group is larger than the other two. The results of second group and the third group are nearly the same. It implies that after three time of practicing the subject would have getting used to the car. Consequently the average results of the third group, the seventh to ninth trials, were used for further analysis.

3.2 *Experiment A: Comparison of safe distances of five testing cars*

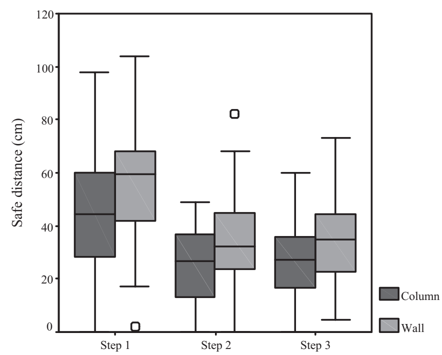
Figure 5 shows the box plot of safe distances of five testing cars. The safe distance associated with Mitsubishi Lancer and Twingo were larger than the other three. Table 2 shows the average of parameter "a" as defined in Figure 2. The relation between safe distances and "a" can be shown as scatter plot in Figure 6. A correlation analysis shows that Pearson correlation coefficient = 0.445 (P = 0.01). It shows that the smaller the parameter "a" the smaller the safe distance. Although Twingo looks smaller than the other four cars, it has steeper hood so that subjects cannot well perceive the gap between the car and the anterior objects.

Safe distance for moving backward tests is shown in Figure 7. Safe distances of Mitsubishi

Figure 3. Position of cardboard.



(a) Moving forward test



(b) Moving backward test

Figure 4. Can practice make difference.

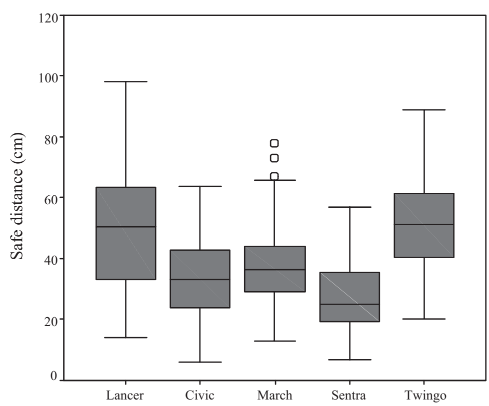


Figure 5. Distribution of safe distance of five testing cars in moving forward tests.

Table 2. Average of parameter "a" for five testing cars.

	Lancer (cm)	Civic (cm)	March (cm)	Sentra (cm)	Twingo (cm)
Parameter "a"	61	48	30	25	55

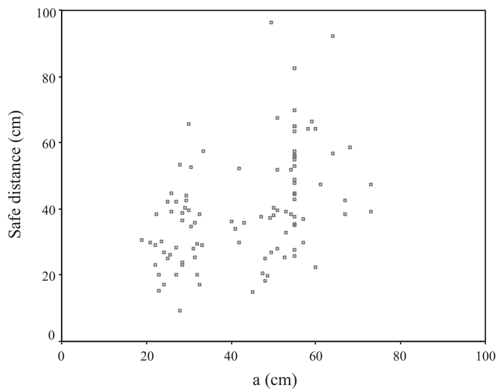


Figure 6. Relation of safe distance to parameter "a" (Correlation coefficient = 0.445, P = 0.01).

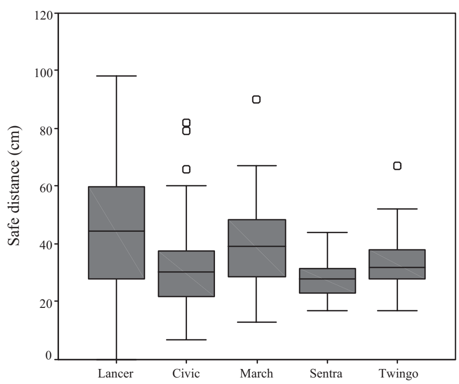
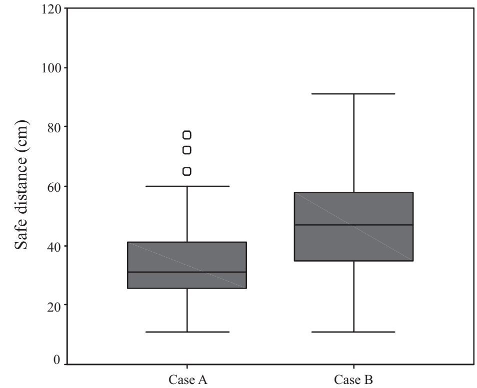


Figure 7. Distribution of safe distance of five testing cars in moving backward tests.

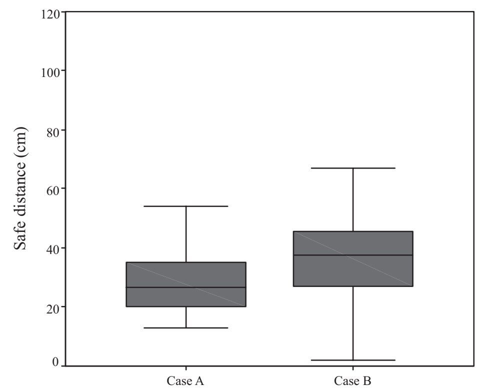
Global Lancer and Nissan March are larger than that of the other three. Although both cars have trunks, drivers cannot see the trunk through back-window. It indicates that the relative height of trunk and back-window may bring uncertainty to the drivers.

3.3 *Experiment B: How form of car influences gap judgment*

As shown in Figure 8 the results in moving forward test indicate that the safe distance associated with



(a) Global Lancer



(b) Honda Civic

Figure 8. Comparison of safe distances in moving forward tests.
Case A: Cardboard at the forefront of the car.
Case B: Cardboard at the lower end of the windshield.

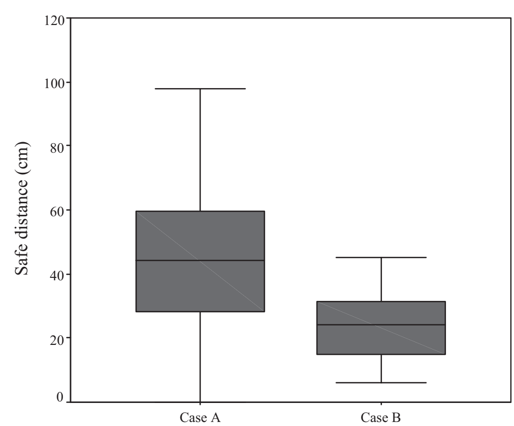


Figure 9. Comparison of safe distances in moving backward tests (Global Lancer).
Case A: Cardboard at the lower end of the back-window.
Case B: Cardboard at the rear-most of the car.

the cardboard at the forefront of the car is smaller than that associated with the cardboard at the lower end of the windscreen, similar for moving backward test, as shown in Figure 9.

4 CONCLUSION

By testing five passenger cars in approaching to anterior and posterior objects to see how the form influences driver's judgment of safe distance between car and objects, conclusion can be summarized as follows:

1. The safe distance between the car and the object varied with the exterior form of the car. A car with shorter front end or a car with trunk lid doesn't assure shorter safe distance.
2. Safe distance depends on if the shape of the hood and truck lid can provide as reference for drivers' perception of the gap. The more the driver can see the fore-most portion of the hood, or the rear-most portion of the trunk lid, the smaller the safe distance.
3. It is anticipated that, with appropriate design, the forms of hood and trunk lid can be good references for drivers to directly well perceive the gap between the car and the anterior or posterior objects. It will be helpful for parallel parking along the roadside.
4. By avoiding the possibility of damaging the testing cars a plastic column and a wall made of wood-board were used as anterior and posterior objects in this experiment, which may not resemble the real situation. It is worth further investigation if a real car is used as the anterior or posterior objects.

REFERENCES

Aguilar, J.J., Torres, F. & Lope, M.A. 1996. Stereo vision for 3D measurement: accuracy analysis, calibration and industrial application. *Measurement* 18(4):193–200.

Burie, J.C., Bruyelle, J.L. & Postaire, J.-G. 1995. Detecting and localizing obstacles in front of a moving vehicle using linear stereo vision. *Mathematical and Computer Modeling* 22(4–7): 235–246.

Chavand, F., Colle, E., Chekhar, Y. & Ni, E.C. 1997. 3-D Measurements using a video camera and a range finder. *IEEE Transactions on Instrumentation and Measurement* 46(6): 1229–1235.

Handmann, U., Kalinke, T., Tzomakas, C., Werner, M. & Seelen, W.V. 2000. An image processing system for driver assistance. *Image and Vision Computing* 18: 367–376.

Hicks, R., Schofield, K., Tarnow, P. & Veiseh, M. 1999. Panoramic electronic rear vision for automotive applications. *SAE* 1999-01-0655.

Kunert, M. 1999. Radar-based near distance sensing device for automotive application. *SAE*: 1999-01-1239.

Kuno, T. & Sugiura, H. 1999. Detection of road condition with CCD Cameras mounted on a vehicle. *System and Computers in Japan* 30(14): 2301–2310.

Miyahara, S. 2005. New Algorithm for the Range Estimation By a Single Frame of a Single Camera. *SAE* 2005-01-1475.

Miyahara, S. 2007. Target Tracking by a Single Camera Based on Range-Window Algorithm and Pattern Matching: Real Time Operation. *SAE* 2007-01-1320.

Miyahara, S., Freeman, K.A., Koulinitch, A. & Tiedje, K. 2009. Blind spot monitoring by a single camera. *SAE* 2009-01-1291.

Oizumi, K., Yamamoto, Y., Sakata. M. & Inoue, M. 2003. Development of all-around view system. *SAE* 2003-01-0017.

Sinai, M.J., Ooi, T.L. & He, Z.J. 1998. Terrain influences the accurate judgments of distance, Letter to *Nature* 395: 497–500.

Takahashi, A., Ninomiya, Y., Ohta, M., Nishida, M. & Takayama, M. 2002. Rear view lane detection by wide angle camera. *Intelligent Vehicle Symposium*, IEEE, Versailles, France.

Young, M. & Chiu, F. 2007. The influence of CCD position on drivers' perception of the gap between passenger car and posterior obstacles. *Journal of Science and Engineering Technology* 3(2): 27–36.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Research development of a sports specific back bag with breathable comfort moisture ergonomic features

Chih-Fu Wu & Neng-Chiao Weng
The Graduate Institute of Design Science, Tatung University, Taipei, Taiwan

Hsi-Sheng Sun & Wei-Chun Hsueh
Footwear & Recreation Technology Research Institute, True East Enterprise Co., Ltd., Taichung, Taiwan

ABSTRACT: The structure of the sport and leisure use back bag is designed through research into general ergonomic requirements, by the use of polyurethane foam sheet (Polyurethane, PU) and polyethylene foam sheets (Expandable Polyurethane, EPE) as an internal filling material, in order to achieve comfort for the human back through lightweight, shock absorption, and sweat absorption capabilities. Footwear & recreation technology research institute (FRT), industry, academic et al. In the material properties of this polyurethane foam include anti-seismic, thermal insulation, absorption of moisture, antistatic, good flexibility, abrasion resistance, light weight and other characteristics but are not easily deformed. Up to now, there are no studies about polyurethane foam material and polyethylene foam material used in sports, leisure and back bags, which address the consumer use, comfort, breathability, moisture absorption and heat and so on in Taiwan. In this study, in addition to general commercial materials, a new research and development is applied called AIR permeability study (materials provided by the True East Enterprise Co., Ltd.). The purpose is to understand whether or not the material properties and structural design will affect the temperature and humidity of the back of the human body. The case experiment, with five healthy male volunteers with appropriate exercise habits for the study, were tested in the campaign. The results show that the 30-minute experiments, with back bags filled with polyurethane foam material, the skin temperature rose by an average of 2.42 degrees. The AIR permeability test results, showed skin temperature increased by an average of 1.25 degrees, and the use of polyethylene foam material enhanced the average temperature by 2.87 degrees. The results showed that the structure of the sports leisure back bag after use of the AIR breathable air cushion provided greater comfort, dehumidification and ventilation functions, whilst the polyurethane foam material and polyethylene foam material is highly insulating, and has moisture absorption function.

Keywords: back bag, comfort, moisture, breathable, air cushion

1 INTRODUCTION

The purpose of this study was to determine the most effective filling for the back bag in relation to the physical properties and structural design, and the effect upon the human back in motion, including comfort, moisture absorption, ventilation and other factors.

The researchers examined the use of a commercially available back bag after exercise utilizing polyurethane foam material, polyethylene foam material, AIR breathable air cushion filler and compared the three experiments.

2 METHODS

2.1 *Participants and materials*

Five healthy male volunteers were selected to do the study. In this study, the average age was 31 years (range 30–33), the average weight was 67 kg (range 60–75), average height was 169 cm (range 167–172). The participants each used commercially available back bag products, adopted the same style, with three different filling materials respectively, at each stage of the experiment.

The change of filler, filling the physical properties and dimensions based bags, The industry standard implementation process was used for changing the filler, and meeting the standardized physical properties and dimensions of the basic bags. The experimental materials included sports and leisure bag (Figure 1), polyurethane foam material foam density 25 times, the size was 38x long and 25x wide and 1 cm thick. The polyethylene foam material foam density 25 times, the length was 25x, width 38x, 1.2 cm thick.

The size of AIR breathable air filled of the TPR cushion material was 38x long 20x wide and 1.5 cm thick. After exercise, the back bag structure was 45x long, 27x wide and 1.6 cm thick, and the area fitted to the back of the human body 38x long, 25 cm wide (Figure 2).

In addition, to obtain the standard with the test results of the participants body back skin temperature, five male volunteers were wearing a uniform heat absorption, humidity and ventilation of 20 cotton method, the 300 yards thickness of 100% cotton white T shirt for testing.

2.2 *Apparatus*

The apparatus supplied by Taishi electronics industrial co., ltd., comprised a TES-1307 K/J thermometer and memory, combined with a K-type thermocouple (Figure 3) in measurement of the human dorsal skin surface temperature of the participants. The measurement scope of the K-type was from (−190°C to 1333°C) (−310°F to 2431°F), accuracy (0°C~1000°C) (+/−0.1% rdg + 0.5°C), and the maximum input voltage overload protection 60VDC/24V rms.

2.3 *Procedures*

Back bags were fitted to the back of the five male volunteers and the A.B.C. three regions use of 9 K-type joints were lodged evenly in the back skin of the human body (Figure 4), with a TES-1307 K/J thermometer to memorize records of the skin temperature. In this study, each participant was carrying 3.5 kilograms of back bag within a constant laboratory temperature of 26 degrees (+/−1), for a 30 minute walking test at a speed of 4 km per hour. Each participant was subjected to a total of three tests in order to ensure the accuracy of the experimental data, between each test and experiment the researchers waited until the back skin temperature of the subjects to normal before implementing the next experiment. Thus the experiment aimed to understand the experimental results of the different control groups.

Figure 1. The back bag for sports and leisure.

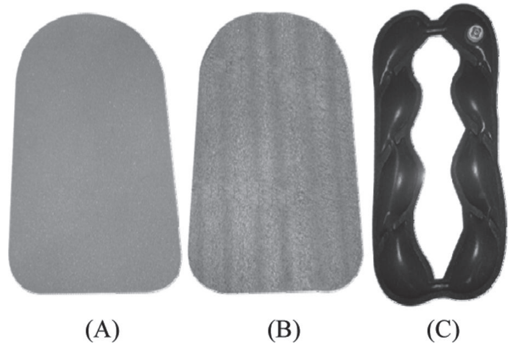


Figure 2. This study used three different properties of the foam material, from left to right order of: A. Polyurethane (PU) B. Expandable Polyurethane (EPE) C. Air permeability.

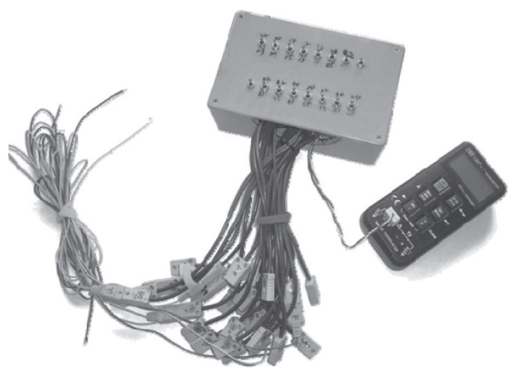


Figure 3. TES-1307 K/J thermometer and memory, combined with K-type thermocouple.

2.4 *Data analysis*

Each participant was tested using polyurethane foam materials, polyethylene foam material, AIR breathable air cushion with the three bags filled

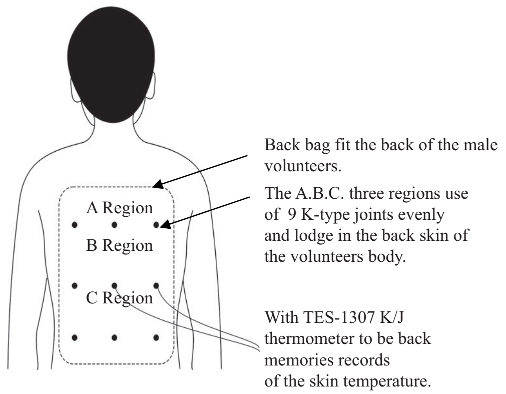


Figure 4. The back bag fit the five (participants) male volunteers back, use of 9 K-type joints evenly and lodge in the back skin of the human body, with TES-1307 K/J thermometer to be back memories records of the skin temperature.

with different materials. The participants carrying back bags continued to walk on a treadmill for the 30 minutes of the experiment, and after stopping the exercise the participants back skin temperature was recorded within 10 minutes the cessation. A total of four stages of the recording, with a record every 10 minutes through the 9 K-type joints of the phenomenon of human skin temperature, every stage recorded the temperature change twice: the data of the temperature change to take a single point K-type the maximum junction temperature and minimum temperature record for the stage, and the average obtained with the single point of data analysis. Participants results were obtained by the back nine measuring points of a single point average, and the overall average value of back skin temperature, and at the same time compared to the average value, as a follow-up analysis.

3 RESULTS

Before the test utilizing polyurethane foam material (Polyurethane, PU) within the filling of bags, the average temperature of the back of the human body skin of the five participants was 32.86 degrees; after exercise of 10 minutes to enhance skin temperature the average temperature increased 1.75 degrees, after 20 minutes the average temperature had increased 2.42 degrees, after 30 minutes the average temperature had increased 2.61 degrees, after 30 minutes exercise, and then stopped for 10 minutes, it mean after 40 minutes, the average temperature has increased 2.42 degrees (Figure 5).

Before the test utilizing polyethylene foam sheets (Expandable Polyurethane, EPE) within the filling of bags, the average temperature of the back of the human body skin of the five participants was 33.15 degrees, after exercise of 10 minutes to enhance skin temperature the average temperature increased 1.49 degrees, after 20 minutes the average temperature had increased 2.58 degrees, after 30 minutes the average temperature had increased 2.88 degrees, after had exercise stopped for 10 minutes the average temperature had increased 2.87 degrees (Figure 6).

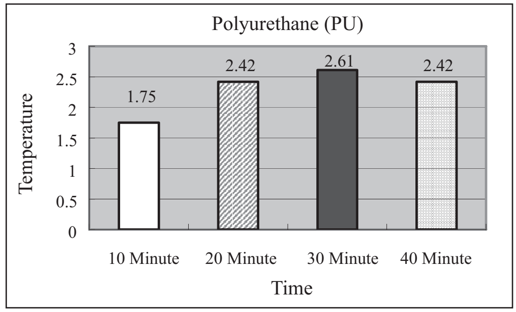


Figure 5. Experimental results of the Polyurethane (PU) in 40 minute.

Before the utilizing the TPR material of the AIR breathable air cushion within the filling of bags, the average temperature of the back of the human body skin of the five participants was 32.83 degrees, after exercise of 10 minutes to enhance skin temperature, the average temperature increased 1.17 degrees, after 20 minutes the average temperature had increased by 1.8 degrees, after 30 minutes the average temperature had increased 1.87 degrees, after exercise stopped for 10 minutes the average temperature had increased 1.25 degrees (Figure 7).

The five male volunteers experiments after exercise 10 minutes, The body back and knapsack fit area, the skin of the average temperature changes. A region on the up back area, five volunteers were tested three kind materials and the average temperature upgrade 0.69 degrees, B region of the middle back area, the average temperature 0.96 degrees to enhance, C region under the back area, the average temperature increases 1 degree. The test results raise the higher temperature area falls in the middle and under of the back (Table 1).

The five male volunteers experiments after exercise 30 minutes, The body back and knapsack fit area, the skin of the average temperature changes. A region on the up back area, five volunteers were tested three kind materials and the average temperature upgrade 1.17 degrees, B region of the middle back area, the average temperature 1.61 degrees to enhance, C region under the back area, the average

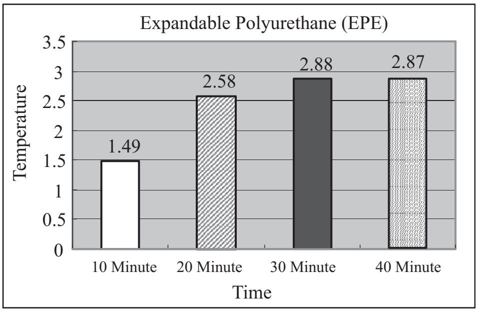


Figure 6. Experimental results of the Expandable Polyurethane, (EPE) in 40 minute.

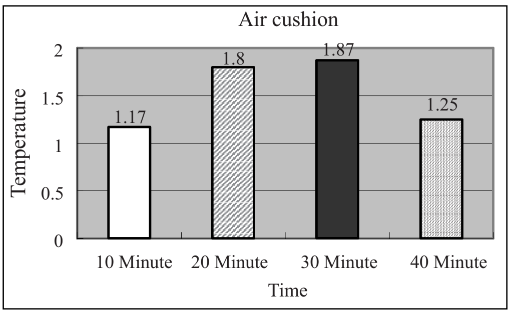


Figure 7. Experimental results of the air cushion in 40 minute.

temperature increases 1.67 degree. The test results raise the higher temperature area falls in the middle and under of the back, similar as the experiments after 10 minutes (Table 2).

The five male volunteers experiments after stops exercise10 minutes, The body back and knapsack fit area, the skin of the average temperature changes. A region on the up back area, five volunteers were tested three kind materials and the average temperature upgrade 0.99 degrees, B region of the middle back area, the average temperature 1.5 degrees to enhance, C region under the back area, the average temperature increases 1.42 degree. The test results raise the higher temperature area falls in the middle and under of the back, similar as the experiments after 10 minutes.

The other, 40-minute experiment with three kind materials, PU material and AIR breathable air cushion, the temperature of the body back skin, C region near under the body back area, the average temperatures were the highest, but the experimental results, use EPE material, the middle of the body back B region, the temperature high than near under body back (Table 3).

Table 1. Three kinds of material under test after exercise 10 minutes, the average temperature of the three regions.

Materials	A region	B region	C region
Polyurethane (PU)	0.71	1.16	1.27
Expandable (EPE)	0.7	1.08	0.94
Air Cushion	0.65	0.64	0.81
Three kinds of material	0.69	0.96	1

Table 2. Three kinds of material under test after exercise 30 minutes, the average temperature of the three regions.

Materials	A region	B region	C region
Polyurethane (PU)	1.11	1.67	1.9
Expandable (EPE)	1.31	2.1	1.75
Air Cushion	1.08	1.07	1.35
Three kinds of material	1.17	1.61	1.67

Table 3. Three kinds of material under test after exercise 30 minutes, and then stop 10 minutes, the average temperature of the three regions in 40 minutes.

Materials	A region	B region	C region
Polyurethane (PU)	0.87	1.65	1.83
Expandable (EPE)	1.49	2.06	1.62
Air Cushion	0.62	0.8	0.8
Three kinds of material	0.99	1.5	1.42

4 DISCUSSION

The use of PU, EPE, Air Cushion as the filling within the three bags showed different experimental results on the human body as follows: after 10 minutes the experiment using PU material to upgrade the body back skin temperature showed increase temperature averaging 1.75 degrees, after the experiment following 20 minutes exercise, EPE material increased the maximum temperature of the body back skin 2.58 degrees, after the experiment had run for 30 minutes, the EPE materials used still showed temperature up 2.88 degrees, after exercise had stopped for 10 minutes, the use of EPE materials of human back skin temperature remained at 2.87 degrees. The results indicated that the insulation material results from the three samples tested demonstrated that the EPE was the best.

Table 4. Experimental results of the comparison of three materials in 40 minute.

Materials	10 minute	20 minute	30 minute	40 minute
Polyurethane	1.75	2.42	2.61	2.42
Air Cushion	1.17	1.8	1.87	1.25
Expandable Polyurethane	1.49	2.58	2.88	2.87

However, for the average participant in the physical exercise, the body back showed significant sweating and temperature upgrade, using the EPE material filling for the back bag. Thus the body back got more sultry through the temperature and humidity increase (Table 4).

4.1 *Limitations of this study*

Variations in the state of motion, time and environment, such as in the actual sports environment experiment, can influence the validity. Also good observation can reflect the actual impact of the movement within different environments.

5 CONCLUSION

PU, EPE, Air Cushion use within three back bags show different experimental results on the back skin temperature of the human body as follows: 10 minutes after the experiment using the PU material the back of the body skin average temperature was 0.58 degrees higher than the Air Cushion, 20 minutes after the test material, the back of the body skin using the EPE material, the average temperature was 0.78 degrees higher than the Air Cushion, 30 minutes after the test material on the back of the body skin, the EPE average temperature was 1.01 degrees higher than the Air Cushion, after stopping exercising for 10 minutes the experiment using the EPE material showed the back of the body skin of average temperature was 1.62 degrees higher than the Air Cushion.

In the 30-minute experiment, back bags filled with polyurethane foam material, the skin of the body temperature average of 2.42 degrees increase. For the AIR permeability test results, the human back skin temperature had increased 1.25 degrees on average. Using the polyethylene foam, the average temperature of the body back skin upgraded 2.87 degrees.

After the actual human body test results, the AIR breathable air cushion provided greater comfort, dehumidification and ventilation functions for use in ventilation products, and, the polyurethane foam material, polyethylene foam material with a high degree of thermal insulation, moisture absorption features are suitable for thermal insulation functionality.

The box and the bags industry, air breathable air cushion suitable for use in direct contact with the skin back style sports-specific bags, leisure bags, computer bags and other products, can achieve effective ventilation and heat effect. The function of a thermal insulation material polyurethane foam, polyethylene foam material is used for insulation features hard and soft style fishing refrigerator, lunch bags, and with heat insulation features shopping cart.

At present, the world's bags box of manufacturing, In general, the comfort features computer bag and back bag, commonly used polyurethane foam material, polyethylene foam material. In this study after the actual test and analysis, The experimental results can be used as a reference for research study and can be used as reference for bag manufacturers design development.

ACKNOWLEDGEMENTS

Thanks to Footwear & Recreation Technology research institute (FRT) testing group for providing laboratory equipment, as well as the proposed experimental methods, and also the true east enterprise co., ltd., who provided sample bags, PU, EPE, Air Cushion and other materials for this experiment. And also grateful to FRT colleagues Liu Jinzhong, Sunxi sheng, Yang same, Wang xin he, and Weng ming chian, for their volunteer work For the testing of this research.

REFERENCES

Chen, W.C. & Lee, T.H. 2008. *Evaluation of low back mechanics and risk for maximum isometric lifting exertion*.

FRTRI 2003. *Physical and chemical testing for footwear and related material on contract basis*. Footwear & Recreation Technology Research Institute, Taiwan.

Hsiung, S.K. & Yu, C.H. 2004. *Study of the humidity sensor based on the polymeric compounds technology*.

Wang, C.F. & Chung, H.C. 2008. *The effects of marching load, duration, and speed on heart rate, breath rate, and rating of perceived exertion*.

Wu, D.H. & Tu, C.H. 2007. *Using multi-walled carbon nanotube/PU foam composite as vibration isolator*.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The design model study of bicycle's seat

Chih-Fu Wu, Li-Lun Huang & Shih-Chang Chen
The Graduate Institute in Design, College of Design, Tatung University, Taipei, Taiwan

ABSTRACT: A After a long bike product development, riding on the body size and frame size relative to the cushion pressure which is generated by the rider. The bicycle seat can be physical size and frame size in relation to the different models in (cf. Chen Shih-Chang & Wu Chih-Fu, 2001). However, in product design, if taking into account all the variables, in addition to increased complexity, the cost would be inconsistent with the actual product development. Therefore, it is necessary to explore the design some suitable seat design models. In this study, we use canonical correlation to find the correlation among the size of variables and pressure variables. Findings, the "weight" and "big rotor wall" are the most important factors of physical size and the secondary is the "big rotor height". The "Five-point cushion," is he most important element factor of frame size, secondary to "cushion handle relatively high" and "seat handle relatively long". Next, the principal component analysis is used to reduce the variables for the cushion pressure and for the frame size and physical size. The most important factor of Pressure is "the P'max value of hip bone area". Therefore, the size of the variables on the nine variables and pressure of the five variables of the 14 variables, we simplified into two dimensions variable with one variable pressure on the three variables. In practice, the product design should take into account the important variables, this study is that a bike seat for product design, providing there could be a common design model as a reference for other commodities.

Keywords: pressure of seat, principal component analysis, canonical correlation, regression analysis

1 INTRODUCTION

The development of Taiwan's bicycle is more than 40 years. It was the largest exporter of bicycles in the world. For the continued decline in cost competitiveness, bicycle factory began to move to mainland China. Due to the low-cost competition from China, the export value showed negative growth during 2001~2002 consequently, the industry has developed high-end bikes gradually since 2003, The average unit price of bicycle exports began to soar even higher. Till to 2005, the average export unit price of bike has been upgraded to $200.

Three major domestic bicycle factories are 'Giant', 'Merida', 'Ideal'. Most of them export to Europe and the United States market. The business behavior, the 'Ideal' emphasized on OEM-based, while the 'Giant' and 'Merida' has also been actively launching its own brand, now owned accounts for a substantial increase in brand sales, 'Merida' up to 90%, 'Giant' is 70%. 2008, domestic sales of 110 million units, the output value of 800 billion dollars to export about 5.4 million units, exports amounted to NT $49 billion, with an average unit cost of $9,000. The Production in the world's are third, second respectively in the world export value, about 810 factories, There are about 15000 employee (TBEA, 2009).

Taiwan's bicycle industry is already a mature industry, the commercial quality of the design also needed to further explore for manufacturers to provide consumers more satisfied product. The bicycle saddle of the commodity for the riding comfortable is very important. Therefore, this study develops some methods in product design for establishing a standard design model.

When riding a bike, human's body and the bike's frame will affect the comfort of sitting. Human's size and bike's frame size are very important reference for consumer to select a more comfortable bike. By designing experiments to explore, riding different bike due to the different human's size and frame's size, the resulting pressure on the relationship between size and saddle's pressure. In the study of "A Study on Saddle Design for Various Types of Bicycles Based on the Seat Pressure of People with Different Anthropometric Characteristics" (Chih-Fu Wu & Shih-Chang Chen), preliminary estimates the initial predicted the relationships between the pressure cushion and frame size and body size. However, the production needs to be more work on the relationship between variables for further integration of the product design. It should take into account the many variables. The production in the bicycle seat design on the rider for the design of different models available seat, the Institute produced a more general design model, but also other commodities as a basis for the design of the design model.

We propose some hypotheses as follows:

a. Hypothesis 1: Were due to differences in human size and saddle pressure.
b. Hypothesis 2: Were due to differences in frame size and saddle pressure.
c. Hypothesis 3: Were due to differences in saddle pressure and human size and frame size.

2 RESEARCH METHODS

In this study, the object of sampling is Tatung University, Department of Industrial Design students. Range in age from 19 years to 23 years of experience with riding a bike scores on the male and female students, male subjects were taking height range from 169 to 180 cm, female subjects were taking height range from 158 to 175 cm. "A Study on Saddle Design for Various Types of Bicycles Based on the Seat Pressure of People with Different Anthropometric Characteristics" (Chih-Fu Wu & Shih-Chang Chen).

Use Cannonical Analysis to find the correlation of the variables between the intensity and related analysis is the most important procedure. Next, we use principal component analysis to do the variables reduction of the pressure cushion and the frame size, to understand the important features of the saddle's pressure and the physical size and the frame size. Finally, we use regression analysis to construct the model that can predict the saddle's pressure by size variables. Product design should take into account many variables, of course, but through the model, the products in the bicycle seat design can provide different models designed to cushion the rider who can produce a more universal design model, and as other product design basis for the design of models to study the procedure shown on Figure 1.

The structure of this study is shown on Figure 2.

3 THE RESULTS OF DATA ANALYSIS

In a previous study, the definition of the original variables of size, including six variables of the human' size and three variables of the frame size, are nine original variables (see Figures 3 & 4). The original variables of pressure are including five

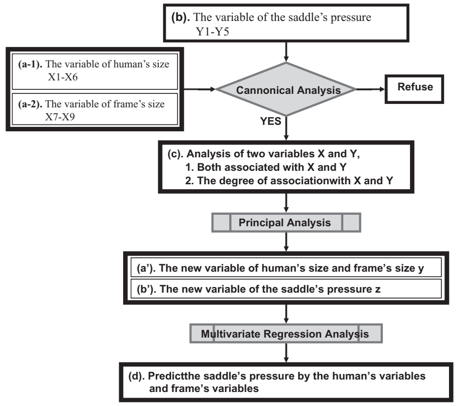


Figure 1. The procedure of study.

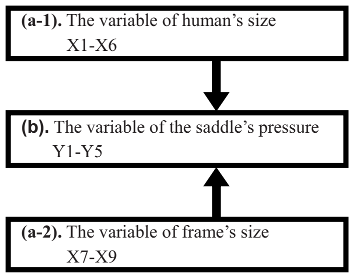


Figure 2. The structure of study.

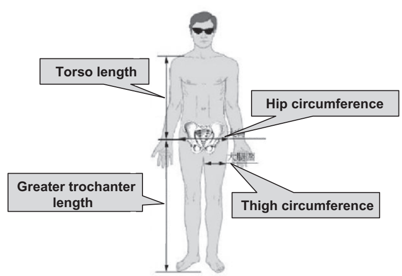


Figure 3. The variables of human's size.

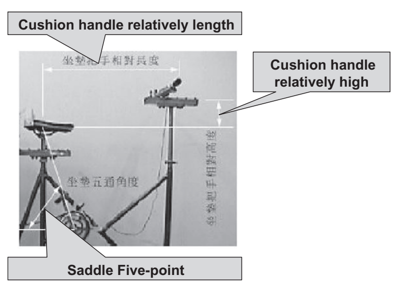


Figure 4. The variables of bike's size.

variables (see Figure 5). Based on the structure and procedure in this study, according to the gender of their Canonical analysis, Principal component analysis and Multiple regression analysis. The original variables of pressure, Pressure in the original variables, "P'max value" for the X-axis value and Y-axis value of the square root, are not independent variables, so "the P'max value of hipbone area" and "the P'max value of ischium area", the two factors in this study should be excluded, so the pressure variables in the analysis, the original pressure variables selected "the Y-axis value of hipbone area", "the X-axis value of ischium area" and "the Y-axis value of the ischium area" to carry out the analysis of the three factors.

Based on the framework of this study, to determine the size variables and pressure variables between this two groups is relevant, shows the correlation coefficient between two variables. Understand the correlation of two structures, the subsequent steps of analysis makes sense.

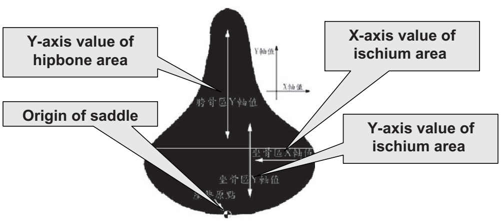


Figure 5. The variables of saddle's pressure.

3.1 *Canonical analysis*

3.1.1 *Canonical analysis of female's data*

The standardization coefficient of pressure's.

Standardization	1	3	4
y1	0.9194	–0.2772	0.0013
y3	–0.3644	0.9608	–0.8673
y4	0.1480	–0.0073	0.4977

The standardization coefficient of size's.

Standardization	1	3	4
x1	0.0916	–0.1397	–0.2470
x2	0.0217	0.0236	–0.0213
x3	–0.0179	–0.1808	–0.1207
x4	–0.0331	0.0159	0.0555
x5	0.0595	–0.1803	–0.0768
x6	–0.0142	0.5326	0.3752
x7	–0.1511	0.0226	0.0486
x8	0.0000	0.0000	0.0000
x9	–0.9814	0.7940	0.8785

Eigenvalues and canonical correlations.

Root no.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	25.907	93.059	93.059	0.981	0.963
2	1.716	6.164	99.223	0.795	0.632
3	0.216	0.777	100	0.422	0.178

The 1st correlation coefficient		The 2nd correlation coefficient		The 3rd correlation coefficient	
w1	v1	w2	v2	w3	v3
0.0916	0.9194	–0.1397	–0.2772	–0.2470	0.0013
0.0217	–0.3644	0.0236	0.9608	–0.0213	–0.8673
–0.0179	0.1480	–0.1808	–0.0073	–0.1207	0.4977
–0.0331		0.0159		0.0555	
0.0595		–0.1803		–0.0768	
–0.0142		0.5326		0.3752	
–0.1511		0.0226		0.0486	
0.0000		0.0000		0.0000	
–0.9814		0.7940		0.8785	
$\hat{CR}_1 = 0.981$		$\hat{CR}_2 = 0.795$		$\hat{CR}_3 = 0.422$	
$\hat{CR}_1^2 = 0.963$		$\hat{CR}_2^2 = 0.632$		$\hat{CR}_3^2 = 0.178$	

Canonical correlation analysis results, the first of the typical correlation coefficient 0.981, indicating a typical variables w1 and v1 is very relevant. By the square of the first canonical correlation coefficient (= 0.963) explained the representative of w1, v1, 96.3% of the errors can be reduced, very high values. w1 explain the size variation of the percentage of variable 14%, v1 take pressure variables to explain the percentage of variation was 4.1%. And from the first of the typical weights, we see the relevant groups, mainly: "Five-point cushion" (weights –0.9239) on: "the Y-axis value of hipbone" (weight is 0.9194) effects. The second variable in the typical, w2 size variables to explain the percentage of variation of 47.4%, v2 take pressure variables to explain the percentage variation of 30.7%. Value of 0.795 indicated that the two groups, and the correlation is also quite strong and from = 0.632 to explain that in order to take pressure of variable size variables, 63.2% of the errors can be reduced. The two sets of related variables, mainly: "Five-point cushion" (weight 0.7940) on: "the X-axis value of ischium area" (the right number is 0.9608) effects. The third of the typical variables, w3 size variables to explain the percentage of variation of

58.4%, v3 take pressure variables to explain the percentage of variation of 18.3%. = 0.422 means that only by the 17.8% error reduction compared with the first canonical correlation strength of the second canonical correlation was weak and mainly represents: "Five-point cushion" (weights 0.8785) on: "the X-axis value of ischium area" (the right number is –0.8673) effects.

The first of the typical variables:

w1 = "greater trochanter height"* (0.0916) + "torso length" * (0.0217) + "greater trochanter circumference"* (–0.0179) + "Thigh circumference" * (–0.0331) + "weight"* 0.0595 + "body fat ratio"* (–0.0142) + "cushion grip relative High"* (–0.1511) + "cushion handle relatively long"* (0.0000) + "saddle Five-point"* (–0.9814).

v1 = "the Y-axis value of hipbone"* (0.9194) + "the X-axis value of ischium area"* (–0.3644) + "the Y-axis value of ischium area"* 0.1480.

The second of the typical variables:

w2 = "greater trochanter height"* (–0.1397) + "torso length"* (0.0236) + "greater trochanter circumference"* (–0.1808) + "Thigh circumference"* (0.0159) + "weight"* (–0.1803) + "body fat ratio"* (0.5326) + "cushion handle relatively high"* (0.0226) + "cushion handle relatively long"* 0.000 + "saddle Five-point"* (0.7940).

v2 = "the Y-axis value of hipbone"* (–0.2772) + "the X-axis value of ischium area"* (0.9608) + "the Y-axis value of ischium area"* (–0.0073).

The third of the typical variables:

w3 = "greater trochanter height"* (–0.2470) + "torso length"* (–0.0213) + "greater trochanter circumference"* (–0.1207) + "Thigh circumference"* (0.0555) + "weight"* (–0.0768) + "body fat ratio"* 0.3752 + "cushion grip relatively high"* (0.0486) + "cushion handle relatively long"* 0.000 + "saddle Five-point"* (0.8785).

v3 = 'the Y-axis value of hipbone"* (–0.6760) + "the X-axis value of ischium area"* 0.3791 + "the Y-axis value of ischium area"* 0.6319.

By the canonical correlation analysis results, identify the variables female size and pressure variables are related, and it is quite high.

The first corresponds to the typical variables of the canonical correlation coefficient of 0.940, by

3.1.2 *The data analysis of male's canonical analysis is with the same way as females'*

The standardization coefficient of pressure's.

Standardization	1	3	4
y1	0.3337	0.2462	–0.6760
y3	0.8470	–0.8526	0.3791
y4	0.4139	0.4610	0.6319

The standardization coefficient of size's.

Standardization	1	3	4
x1	0.2579	–0.2112	0.1500
x2	0.1686	–0.0292	0.1429
x3	–0.0255	0.1073	–0.1150
x4	0.0298	0.0073	0.1056
x5	0.0879	–0.1170	–0.0107
x6	–0.0538	0.1363	0.0955
x7	–0.1984	0.0363	–0.0393
x8	0.0000	0.0000	0.0000
x9	–0.9239	0.9536	–0.9602

Eigenvalues and canonical correlations.

Root no.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	7.5700	81.1730	81.1730	0.9400	0.8830
2	1.5870	17.0150	98.1880	0.7830	0.6130
3	0.1690	1.8120	100.0000	0.3800	0.1450

The 1st correlation coefficient		The 2nd correlation coefficient		The 3rd correlation coefficient	
w1	v1	w2	v2	w3	v3
0.2579	0.3337	–0.2112	0.2462	0.1500	–0.6760
0.1686	0.8470	–0.0292	–0.8526	0.1429	0.3791
–0.0255	0.4139	0.1073	0.4610	–0.1150	0.6319
0.0298		0.0073		0.1056	
0.0879		–0.1170		–0.0107	
–0.0538		0.1363		0.0955	
–0.1984		0.0363		–0.0393	
0.0000		0.0000		0.0000	
–0.9239		0.9536		–0.9602	
$\hat{CR}_1 = 0.940$		$\hat{CR}_2 = 0.783$		$\hat{CR}_3 = 0.380$	
$\hat{CR}_1^2 = 0.883$		$\hat{CR}_2^2 = 0.613$		$\hat{CR}_3^2 = 0.145$	

the square of the first canonical correlation coefficient (=0.883) can be understood by the body frame size (w1) Block explained pressure (v1) errors can be reduced 88.3%. And from the first of the typical weights, we see the relevant groups, mainly: "Five-point cushion" (weights –0.9239) on: "the X-axis value of ischium area" (weight is 0.8470) effects. The second corresponds to the typical variables on the two groups, and the correlation is also quite strong, and 0.783, indicating 61.3% of the errors can be reduced, and this relationship is the representative: "saddle Five-point" (weights 0.9536) on: "the X-axis values of ischium area" (weights –0.8526) effects. In the third variable, the relationship between the typical weak, and 0.380, only 14.5% of error reduction, mainly on behalf of their relations: "saddle Five-point" (weights 0.9536) on: "the Y-axis value of hipbone" (weights –0.8526) effects. The third variable related to the typical weak, the third on the second than the first pair of variables and in the typical variation is much lower.

The first of the typical variables:

w1 = "greater trochanter height"* (0.2579) + "torso length"* (0.1686) + "greater trochanter circumference"* (–0.0255) + "Thigh circumference"* (0.0298) + "weight"* 0.0879 + "body fat ratio"* (–0.0538) + "cushion handle relatively high"* (–0.1984) + "cushion handle relatively long"* (0.0000) + "saddle Five-point"* (0.9239).

v1 = "the Y-axis value of hipbone"* (0.3337) + "the X-axis value of ischium area"* 0.8470 + 'the Y-axis value of ischium area"* 0.4139.

The second of the typical variables:

w2 = "greater trochanter height" * (–0.2112) + "torso length"* (–0.0292) + "greater trochanter circumference"* (0.1073) + "Thigh circumference"* (0.0073) + "weight"* (–0.1170) + "body fat ratio"* (0.1363) + "cushion handle relatively high"* (0.0363) + 'cushion handle relatively long" * 0.000 + "saddle Five-point"* (0.9536).

v2 = "the Y-axis value of hipbone"* 0.2462 + "the X-axis value of ischium"* (–0.8526) + "the Y-axis value of ischium"* 0.4610.

The third of the typical variables:

w3 = "greater trochanter height"* 0.1500 + "torso length"* 0.1429 + "greater trochanter circumference* (–0.1150) + "Thigh circumference"* (–0.1056) + "weight"* (–0.0107) + "body fat ratio"* 0.0955 + "cushion handle relatively high"* (–0.0393) + "cushion handle relatively long"* 0.000 + "saddle Five-point" * (–0.9602).

v3 = "the Y-axis value of hipbone"* (–0.6760) + "the X-axis value of ischium" * 0.3791 + "the Y-axis value of ischium" * 0.6319.

Individual sex canonical correlation analysis, determine the size of variables and stress variables are relevant and very high. People because of size and frame size can be the principal component analysis.

3.2 *Principal component analysis*

The size of variables and pressure variables, two components of each call the shots:

3.2.1 *Principal component analysis of size's variable*

Principal component	Eigen-value	Differ-ence	Cumula-tive	Pro-por-tion	Squre of Eigen-value
1	3.39	37.62	37.62	3.39	1.84
2	2.38	26.41	64.03	2.38	1.54
3	2.06	22.93	86.96	2.06	1.44
4	0.43	4.83	91.78	0.43	
5	0.39	4.31	96.10	0.39	
6	0.18	2.03	98.13	0.18	
7	0.14	1.58	99.71	0.14	
8	0.03	0.29	100.00	0.03	
9	0.00	0.00	100.00	0.00	

Standardization of Principal component matrix	1	2	3
Greater trochanter height	0.469	0.058	0.160
Torso length	0.471	0.084	0.192
Greater trochanter circumference	–0.025	0.587	0.056
Thigh circumference	–0.249	0.455	–0.047
Weight	0.328	0.441	0.161
Body fat ratio	–0.271	0.482	–0.052
Cushion handle relati vely high	–0.142	–0.058	0.668
Cushion handle relati vely length	0.142	0.058	–0.668
Saddle Five-point	0.524	0.059	–0.101

Explain the variance in scale, it shows that the fourth principal component eigenvalue <1, so they chose the three main components, re-named as the new variables y1, y2 and y3, easy to distinguish from the research steps.

Equation with a score of principal component:

y1 = "greater trochanter height" * 0.469 + "torso length" * 0.471 + "greater trochanter circumference" (–0.025) + "Thigh circumference" * (–0.249) + "weight" 0.328 + "body fat ratio" * (–0.271) + "cushion handle relatively high" * (– 0.142) + "cushion grip cushion relatively long" * 0.142 + "saddle five-point" * 0.524.

y2 = "greater trochanter height" * 0.058 + "torso length" * 0.084 + "greater trochanter circumference" * 0.587 + "Thigh circumference" * 0.455 + "weight" * 0.441 + "body fat ratio" * 0.482 + "cushion handle relatively high" * (–0.058) + "cushion handle relatively long" * 0.058 + "saddle Five-angle" * 0.059.

y3 = "greater trochanter height" * 0.160+ "torso length" * 0.192 + "greater trochanter circumference" *

0.056 + "Thigh circumference" * (–0.047) + "weight" * 0.161 + "body fat ratio" * (–0.052) + "cushion handle relatively high" * 0.668 + "cushion handle relatively long" * (–0.668) + "saddle Five-point" * (–0.101).

As the results of size's Principal component analysis, we can find two points as following:

i. The first principal component explained about 37.62% for the second principal component of about 26.41 percent, the second principal component of about 22.93%, 86.96% total has been about.
ii. Showing a size of the first principal component variables by the "saddle five-point" (0.524) the greatest impact, the "greater trochanter height" (0.469) and "torso length" (0.471) the second place, are subject to height and frame size factor and the impact of; second principal component is the "greater trochanter circumference" (0.587) most affected, followed by "thigh circumference" (0.455), "weight" (0.441) and "body fat ratio" (0.482), belong to the people because of size, the changes related with body fat or thin items, as degree of obesity (fat percentage) the maximum impact strength; the third principal component, was "cushion handle relatively high" (0.668) and the "cushion handle relatively long" (–0.668) of the two variables affecting the hand the position, as belonging to the highest intensity of vehicle impact.

3.2.2 *Principal component analysis of pressure variables*

Principal-component	Eigen-value	Differ-ence	Cumu-lative	Pro-portion	Squre of eige-value
1	1.962	65.408	65.408	1.962	1.401
2	0.665	22.171	87.579	0.665	
3	0.373	12.421	100.000	0.373	

Standardization of principal component matrix	Composition
The Y-axis value of hipbone	0.588
The X-axis value of ischium	0.622
The Y-axis value of ischium	0.517

Explain the variance in scale, it shows that only the first principal component eigenvalue >1, so they chose a main component, re-named as the new variables z1, easy to distinguish from the research steps.

Equation with a score of principal component:

z1 = "the Y-axis value of hipbone" * 0.588 + "the X-axis value of ischium" * 0.622 + "the Y-axis value of ischium" * 0.517.

Presented by the first principal component "the X-axis value of ischium area" the greatest impact, followed by close to strength, "the Y-axis value of hipbone area" and "the Y-axis value of ischium area."

As the results of pressure's Principal component analysis, we can find two points as following:

i. The first principal component to explain about 37.62 percent, the second principal component of about 26.41 percent, the second principal component of about 22.93%, 86.96% total has been about.
ii. Showing a size of the first principal component variables by the "saddle Five-point" (0.524) the greatest impact, the "greater trochanter height" (0.469) and "torso length" (0.471) is second with frame size factor are affected by the impact; second the major components of the "greater trochanter circumference" (0.587) most affected, followed by "thigh circumference" (0.455), "weight" (0.441) and "body fat ratio" (0.482), belong to the people because of size, the fat or thin body with the relevant variables, as for the degree of obesity (fat percentage) the maximum impact strength; the third principal component, was "cushion handle relatively high" (0.668) and the "cushion handle relatively length" (–0.668). The two variables that affect the hand position, as belonging to the highest intensity of vehicle impact.

3.3 *Regression analysis*

Mode	R	R' square	Modi-fied R' square	Standard error of estimate	Durbin-watson coefficient
1	0.844	0.712	0.691	0.556	1.999

i. Forecast variables: (constant), size vehicles with REGR factor score 3 for analysis 4, size vehicles with REGR factor score 2 for analysis 4, size vehicles with REGR factor score 1 for analysis 4, M: 1; F: 0
ii. Dependent variables: the first principal component of pressure is REGR factor score 1 for analysis 3.

Mode	Beta estimate value	Deviation	Beta distribution	P value
(constant)	–0.019	0.156		0.904
Male:1; female:0	0.038	0.277	0.019	0.892
The 1'st component	0.188	0.135	0.188	0.172
The 2'nd component	0.154	0.072	0.154	0.038
The 3'rd component	–0.809	0.079	–0.809	0.000

Residuals Satistics

	Min.	Max.	Average	Deviation	Number
Predictive value	-1.172	1.232	0.000	0.844	60
Residuals	-1.145	1.823	0.000	0.536	60
Standardization of Predictive value	-1.388	1.459	0.000	1.000	60
Standardization of Residuals	-2.062	3.281	0.000	0.966	60

$\hat{CR}_1^2$ = MSE = 0.712, Durbin-Watson Test = 1.999. Regression analysis predicted that effective and good, assuming a simple binary model becomes a linear equation. However, only the third principal component in the prediction was significant pressure on. In the regression analysis to establish the prediction model, but the third principal component can predict the pressure of people because of the height variables are the second principal component and the variables that are part of human obesity due to the third degree of the main ingredients, but can not predict the pressure should be as close to the sample data, the nature of the relationship.

4 CONCLUSION AND RECOMMENDATIONS

Size and pressure groups in the canonical correlation analysis of variables in the first two groups of men showed typical correlation coefficients are: "saddle Five-point" for: "the X-axis value of ischium area" strongest influence. However, women are the first group of canonical correlation is: "saddle five-point" for: "the Y-axis value of hipbone area" the strongest impact; second canonical correlation is the same with men is: "saddle five-point" for: "the X-axis value of ischium area", the strongest effect. Show that the impact of gender differences in stress variables in the size variable is some different.

Standardization of principal component matrix	Composition	Ratio
The Y-axis value of hipbone	0.588	9
The X-axis value of ischium	0.622	10
The Y-axis value of ischium	0.517	8

Pressure variables in the main components, the standardization of weights for the three main components, the access ratio for the "the X-axis value of ischium area": "the Y-axis value of ischium area": "the Y-axis value of hipbone area", ≒10:8:9 may be seen in three dimensions cushion ratio, which means "the X axis on ischium area": "the Y axis on ischium area": "the Y axis on hipbone area": ≒ 10:8:9. This ratio allows manufacturers to obtain more moderate shape ratio.

In the regression analysis of the prediction model, only bike's model could predict the saddle pressure, that sampling a smaller range. In this study, the experimental sample in the north metropolitan area, range in age from 19 to 23-year-old boys and girls, young and vigorous activity with the students, certainly more difficult in the degree of obesity are different, with fewer samples, but also will make the size of the principal component variables is difficult to predict the pressure cushion. In another perspective, that to meet this as the space of consumers, is the diversification of products for the bicycle seat cushion grip handle relatively high and relatively long model changes, the same models for the effects of other variables can not see. Since this study is to establish the basic forecasting model based testing, the sample used to quickly obtain a lesser extent, data, forecast model is consistent with the results of sampling.

REFERENCES

Huang, D.Y. 2001. *Applied regression analysis*. Hwa Tai Publishing Co. Ltd., Taipei.

Lee, R.F. & Huang, D.Y. 2001. *Fu Jen management review*. College of Management, Fu Jen Catholic University. (Thesis).

SPSS 1994. Advanced Statistics 6.1, SPSS Inc.

SPSS 2000. Advanced Statistics 12.0.1C. SPSS Inc.

Wu, C.F. & Chen, S.C. 2001. *A study on saddle design for various types of bicycles based on the seat pressure of people with different anthropometric characteristics*. College of Design, Tatung University, Taipei, Taiwan. (Thesis).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Human factors integration in railway vehicle design

Karen Priestley & Plato Chan
Lloyd's Register Rail (Asia) Limited, Hong-Kong, China

ABSTRACT: This paper discusses the approach to integrate Human Factors (HF) in a railway vehicle design project in China. It gives an overview of human factors integration in railway projects and shares the experience of applying human factors integration techniques to a rolling stock project in China. The use of Human Factors Issue Register (HFIR) was established on the project and proved invaluable as a communication and tracking tool between elements of the project team separated by geography and language. Three case studies are presented from the project to demonstrate how some of the issues raised on the HFIR were addressed and providing evidence of the value of human factors in (1) Uncovering the user requirements; (2) Improving the existing design through HF analyses and (3) Evaluating design concepts of a new system feature. This HF integration experience highlights the importance to consider the requirement for human factors early in the project so the scope and focus of activities can be determined to add value to the project goals.

Keywords: railway vehicles, human factors integration, systems design

1 INTRODUCTION

Human Factors is an interdisciplinary field of study that "discovers and applies information about human behavior, abilities, limitations and other characteristics to the design of tools, machines, systems, tasks, jobs and environments for productive, safe, comfortable and effective human use" (Sanders & McCormick, 1992). Within the rail industry, the purpose of applying human factors in projects is to minimize opportunities for operators, maintainers or passengers to make errors, improve efficiency and reduce training requirements for the resulting system design. Through the application of these techniques there can be benefits to the design process itself through derivation of more informed requirements, avoidance of redesign and improved acceptance of the system by the end users.

Many high profile rail accidents have raised awareness of the need to better understand and evaluate the role of people in the design of systems and the achievement of the desired system performance. Incidents such as the train derailment at Ladbroke Grove, UK, highlight problems with signaling systems; Waterfall in Australia revealed issues with the design of safety systems in the train cab; Daegu in Korea and Kings Cross in the UK were examples where the station design did not facilitate the effective evacuation of passengers.

Furthermore, there are examples of expensive systems introduced to try and prevent human error that were not as effective as expected or introduced additional errors because they failed to consider human limitations and behaviour in the design. For example, the Automatic Train Warning System in the UK gives an audible alarm that the operator must acknowledge when approaching a yellow signal. However, various problems have been identified with this solution, including the automatic response to the alarm by the operators who then fail to slow down the train accordingly. These types of incidents and disappointing results from engineered solutions have led to the recognition of the importance of human factors in Europe and Australia and the requirements for human factors to be given equal consideration in the design to any other engineering discipline. Human factors processes and techniques are integrated into the design process and considered from the earliest opportunity.

While the need for human factors is now recognised on European and Australian rolling stock projects, it is not frequently found in Asian projects. Consequently the concepts, techniques and potential value are not widely known to rolling stock manufacturers in the region.

The MTR Corporation Limited (MTR) from Hong Kong is one Asian client that does specify requirements for human factors issues to be addressed in their rolling stock design and build contracts. Lloyds' Register Rail (LR Rail) is currently supporting CNR Changchun Railway

Vehicles Co. Ltd. (CNR CRC) to meet the MTR human factors requirements to provide a fleet of 10 trains for its existing operation and future line extension.

The design of these trains is largely based on existing rolling stock within the MTR fleet. Thus many issues associated with the physical design and control systems were already established. The key human factors input to this project was to evaluate potential enhancements to the existing design and the introduction of new systems and assess the impact of differences to the performance of the operators, maintainers and passengers.

2 HUMAN FACTORS INTEGRATION

2.1 *Standards*

A generic framework for human factors integration is available in the document ISO 13407:1999—Human-centred design processes for interactive systems. The human factors integration process described in this document includes an understanding of the end users of the system, the tasks to be performed and the context in which they are undertaken. This is achieved through involvement of stakeholders and the application of available data on human requirements and performance. This forms the basis of the development of user requirements to pass on to the design team. The standard requires that an iterative design process be adopted such that the user requirements are evaluated throughout to ensure that either they can be met or that the requirements are amended in the light of further understanding about the design and context of use. The activities required to integrate this process with the design process should be outlined in a Human Factors Integration Plan. An overview of the human factors integration process is shown in Figure 1.

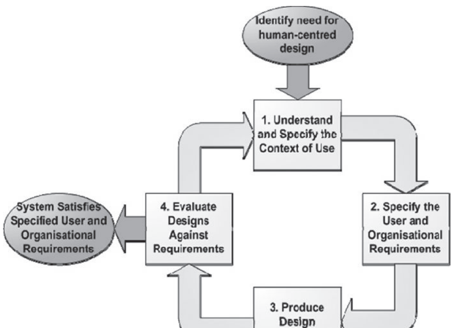


Figure 1. The process of human factors integration described in ISO 13407.

In railway engineering safety management, standards such as EN 50126:1999 requires that 'an analysis of human factors, with respect to their effect on system RAMS' (Reliability, Availability, Maintainability and Safety). The guidance document 'Engineering Safety Management' (the Yellow Book, Railway Safety and Standards Board, 2007) has integrated requirements for human factors activities through the lifecycle phases from concept and feasibility to decommissioning and disposal, into its latest edition.

In the UK, the requirements from all these standards have been adopted into internal standards for the integration of human factors by organizations such as Network Rail, London Underground and some of the larger contractors. To provide further assistance to the rail industry, the Rail Safety and Standards Board (2008) released a handbook of human factors techniques that could be applied to rail projects.

2.2 *Human factors requirements on the project*

The requirements for human factors activities in this rolling stock project were set out in the particular specification. This called for specific activities like the development of an integration plan, task analysis, the conduct of user trials to evaluate conformance with specific requirements e.g. human physical dimensions and sightlines and the identification of training requirements to be undertaken etc.

Not all the features of the integration standards appeared in the client's specification. Through the integration plan, LR Rail introduced the concept of the Human Factors Issue Register (HFIR), a fundamental technique for integrating human factors within a project described in the 'Yellow Book'.

2.3 *The HFIR*

The HFIR is an important tool in the management of Human Factors Issues on this project. It is essentially a spreadsheet that is used to hold a central record of the issues identified and their status on the project. A procedure was developed to describe how human factors issues are identified, logged, addressed, monitored and resolved. It is used to co-ordinate and communicate information on activities undertaken by three main parties; namely CNR CRC the contractor, MTR the client and LR Rail the human factors specialist. Clearly, one of the advantages of this approach is to enable efficient communication and co-ordination between three parties that speak three different languages, based in two different geographical locations, ensuring that issues are tracked through to closure and not lost.

The HFIR served as a record of the issue, the issue owner, the progress to resolution, the mechanism for evaluation (design drawings, demonstration or testing through trials) and the status. As the design progressed further trials were conducted on prototypes and mock-ups to test the operability and maintainability performance of the design according to requirements laid down in the particular specification and requirements derived from analysis of the issues on the HFIR. This provided demonstration that issues had been adequately addressed by the design, but in some instances revealed additional issues to be recorded and mitigated.

Not all issues can be addressed through design alone. Where additional measures are required to address an issue through operational means, e.g. training and procedures, these issues have been highlighted for transfer to the respective parties developing these elements.

2.4 *Identifying human factors issues to go on the HFIR*

As this project was essentially to improve upon an existing rolling stock design it was not necessary to start from a blank page. Many design parameters were already established from existing standards within MTR and provided in the client's specification along with requirements for testing. The key issue for this project was therefore to identify appropriate elements for improvement from the existing design and the implications of new systems and functionality being introduced to the train.

Initially, issues were identified through consultation with train operators regarding aspects of the existing rolling stock that they would like to see improved. Subsequently, as the design progressed, techniques such as task analysis, error analysis and difference analysis were used to identify potential problems with the design of specific elements of the rolling stock system. These issues were recorded on the HFIR. Designers and clients were also free to raise concerns onto the HFIR for evaluation by the human factors team.

3 CASE STUDIES

The MTR Corporation has an enviable safety and reliability record and the performance of the rolling stock plays a significant part in this result. The use of existing rolling stock as the basis for the specification of the design should enable the existing record to be matched if not exceeded. With regard to safety, many of the standards required and features of the existing design have been proven in operation. What remains for improvement is the enhancement to performance and the reduction of response times to incidents that do occur. Compared to other railways in the world, the issues described in the following case studies may appear relatively minor and have little impact on safety performance. However, in the context of MTR where delays over two minutes are investigated and those over eight minutes are reported to Government bodies, the performance improvements achieved have real value.

3.1 *Case study 1—distinguishing 'needs' from 'nice to haves'*

In order to identify potential improvements for existing designs and to elicit user requirements it is necessary to talk to the operators of existing systems. This is something that is required by human factors integration standards and should be embraced rather than feared. There can be a lot of resistance to speaking with operators or end customers from project teams. This is often rooted in concerns that a 'wish list' will be generated that the project team will be unable to meet. This sometimes arises because operators jump to conclusions regarding the solution to the problem that they are experiencing rather than describing the problem. The role of human factors is to identify the issues that impact upon performance and therefore to undertake systematic analysis and evaluation of user input to define performance problems and user requirements to be addressed by the project team.

The first case study relates to this kind of situation. In this project, the initial list of issues identified for enhancement by the operators included a request for a larger screen for the Train Management System (a computer control system). This led to concerns from the project team with regard to cost, physical space to accommodate a larger screen and impact on sightlines to signals through the windscreen from placing a larger TMS screen on the console.

However, after analysing the task and discussing with the operators we identified that the root problem was that one essential piece of information was missing from the display of one out of the eight sub-systems managed through the TMS. Finding a way to present this information clearly to the operator could lead to a solution that did not involve increasing the dimensions of the screen.

The problem for the operators was that in the event of a door failure they are under intense time pressure to meet recovery targets. It is imperative to convey the information on the faulty door to the station staff. Any errors made in identifying the location of the faulty door would lead to extended delays, which is a key operational concern. On the existing system the door numbers were not shown and so it was easy to make a mistake, particularly as the numbering changes direction depending upon the configuration of the cars. It is necessary for the operator to remember the numbering sequence and start to count the boxes on a small screen. There are many opportunities to go wrong. To try and help prevent errors, operators taped their own door numbering aids to the edge of the monitor. This had led them to the assumption that a larger screen was required in order to incorporate the door numbers into the interface.

The human factors team asked the designers how they could present this information within the existing screen dimensions. This enabled the design team to come up with a solution within the physical limits of the existing design, thus saving considerable cost, time, design effort and ongoing maintenance issues. The solution was acceptable to the operators and will reduce the occurrence of errors and delays in responding to door faults due to misdirection of station staff to the wrong door. The resulting solution is presented in Figure 2 below.

Although the solution still requires the drivers to count, it is greatly simplified by the new design and there is no longer a requirement on the operator to remember the numbering sequence. By fully understanding the task and the pressures that the operator is under when he needs this information it was possible to boil down their need to a particular piece of information that was missing from one display.

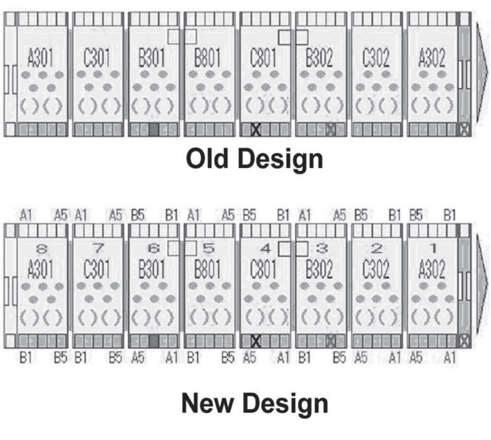


Figure 2. The train door display in the previous train (top) and the new train (bottom), with car and door number added.

3.2 *Case study 2—enhancing usability of existing designs*

The second case study demonstrates how the existing design can be further improved by human factors analyses. A number of issues were raised by the operators regarding the Digital Voice Announcement System (DVAS), particularly in response to passengers pulling the alarm handle in the saloon. This required a long sequence of buttons to be pressed before the operator could hear the passenger. Often by the time the operator had finished the sequence the passenger had finished talking and may have walked away, leaving the operator no wiser regarding the problem. Again this situation could increase the time to recover from the delay.

Again by obtaining a full understanding of the task and the context in which the task is being performed it is easy to see that using the DVAS system is stressful for operators in response to a passenger alarm. The operator not only has to use the DVAS system, but will need to check any associated failure indications on the TMS before relaying information to the control centre on the radio system. There are many potential distractions and the opportunity for errors and associated delays in resolving the situation are increased by the existing design of the DVAS system.

As each button had to be held down for 0.5 s, it was not uncommon for operators rushing under time pressure and stress to fail to hold the buttons down long enough before pressing the next button in the sequence. Thus they could get to the end and find they had made a mistake and be unclear at what point the mistake was made, leading to longer delays as they try to figure out where they went wrong.

A couple of user requirements were self-evident from this description. Firstly, that there should be no requirement to hold the buttons down to operate the system. Secondly, that an operator should be able to hear the passenger immediately the alarm handle is operated in the saloon. These were recorded on the HFIR and translated into design criteria for the design team to develop solutions.

Closer examination of the procedures for using the DVAS system and inspection of the existing interface revealed that further improvements could be made to conform to basic human factors principles for interface design, namely:

- Frequency and Criticality—those controls that are used most frequently or that are critical should be within easiest reach of the operator
- Sequence—layout of controls should follow the task sequence if there is one
- Functionality—controls for similar functions should be grouped together.

It was observed from the procedures that for any of the modes of operation a similar sequence was required as follows:

1. Select the appropriate mode
2. Use the scroll keys to identify the right message
3. Broadcast the selected message
4. Reset the system when message no longer needed.

Within the constraints of the space limitations and the desire to keep the buttons located around the display screen, an alternative proposal for the layout of the buttons was proposed that was compliant with the human factors design principles. The before and after layouts are shown in Figure 3. The 'link analysis' technique is used in this diagram to evaluate the sequence of control operation for each layout.

From Figure 3 it can be seen that the revised design follows the task sequence from top to bottom. The most frequently used buttons are on the right hand side which is closest to the operator. Thus, the revised layout is more intuitive and efficient than the first. Less time is spent visually searching for the right button and it will be less likely that mistakes are made in the actual selection of the buttons because the sequence moves in the same direction rather then going back and forth repeatedly.

The human factors input to this task has not only had a positive impact on the ability of the operator to use the interface to respond efficiently to passenger alarms, it also facilitated more effective communication between customer, contractor and sub-contractor in developing the solutions. This in turn increases the chance of user acceptance.

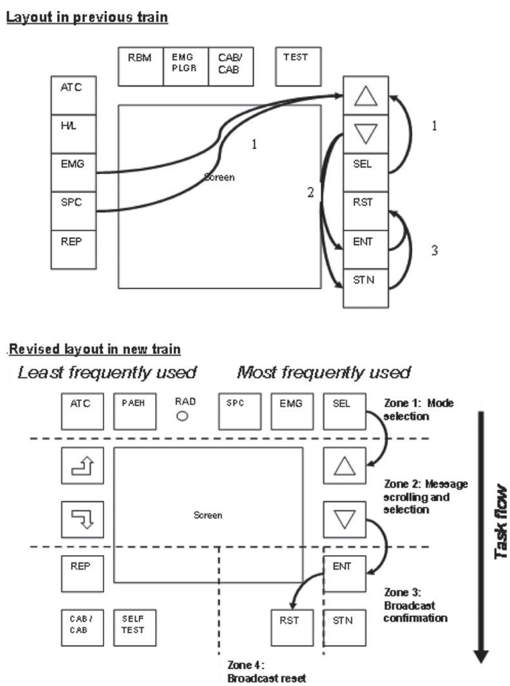


Figure 3. DVAS control panel layout in the previous train (top) and new train (bottom). The layout of the buttons was rearranged according to the task flow to improve usability and efficiency.

3.3 *Case study 3—evaluating concepts of new proposed systems*

The third case study involves the introduction of a new in-saloon CCTV system to be controlled by the operator in the train cab. It illustrates how human factors inputs can be utilized to evaluate concepts and help to define user requirements very early in the design. This new system will allow the operator to select the appropriate camera to see inside the saloon of any car through the TMS. It was necessary to consider in advance what the issues might be with the operation of the system in the context of the operators' overall tasks.

One of the issues arising from the task and error analysis was whether the operator could easily identify the correct camera to display the correct section of the saloon that they wanted. Similar to the previous case study on DVAS, communication between the client, contractor and HF consultant was necessary. Also, the subcontractors for the TMS and the CCTV system would need to be co-ordinated.

For the control interface to be intuitive for the operators it must reflect operators' concept of the system. Where the concept of the system designer and the user do not match it is generally found that the system encourages error. Simple examples are number pad layouts, see Figure 4. The one on the left is the intuitive design for people who are

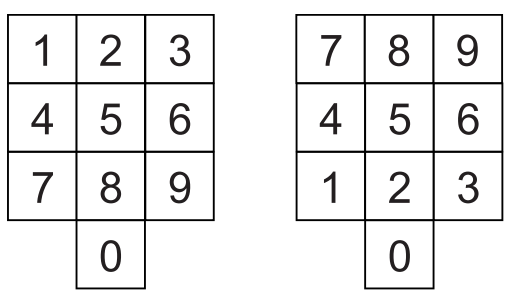


Figure 4. Numeric keypad with digits running to top-left to bottom-right (left) were more intuitive than the one running from bottom-left to top-right (right), with less errors committed by the users.

trained to read from left to right and from top to bottom (e.g. in English) and results in fewer errors than that on the right. This is because it matches their habitual pattern for searching information with the sequence of numbers.

The development of a new system presents an opportunity to study and integrate users' conceptual model into the design at early stage of the project.

In this case study, to select an in-saloon camera for display on the CCTV monitor the operator must:

1. Think of the car for the desired display image
2. Think of the section of the car for the desired display image
3. Review TMS interface to select the appropriate camera
4. Check the information shown on the CCTV monitor to verify that the correct image is displayed
5. Review the scene as needed.

The human factors team had concerns that the original proposals for the control interface design may not reflect the operators' internal concept of the system. This is because the controls in preliminary designs were laid out in a quadrant that represented the way the images were displayed on the CCTV monitor rather than how the images were captured in the saloon. Effectively the images in the saloon were captured in a linear fashion starting one end of the cab in four sections to the other end, see Figure 5.

To ensure that the operator does not make any mistakes in selection it is better to represent this linear arrangement than require them to mentally translate the linear arrangement into a quadrant.

In order to evaluate these and other human factors concerns regarding the interface design a series of paper-based tests were conducted to obtain insight into the operators' concepts. This involved setting a number of scenarios and presenting the operator with blank grids and for each scenario

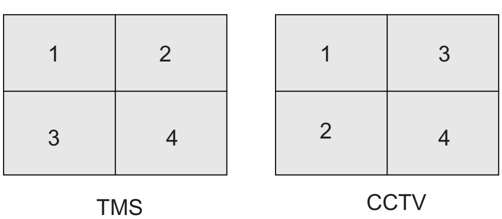


Figure 5. The original layouts of in-saloon CCTV selection panel in TMS (left) as compared to the CCTV display screen (right). The numbers in the quadrants indicate the camera numbers. The cameras were not installed in sequence following the numbers.

Figure 6. The linear layout of camera images on CCTV display (left) and matches with the selection panel on TMS (right). The numbers represent sections of the car instead of camera numbers. The camera numbers are transparent to the operators and a translation matrix was developed for the engineers based upon what the operators wanted to see on the interface.

asking which control they would expect to use and where they would expect the image to be presented. The results from this test showed that the linear layout in camera selection screen together with the use of section number instead of camera number were preferred by the operators. Furthermore, the operators pictured the cab at the front of the train, so that the section numbers counted away from the cab whichever journey was being undertaken. It was necessary to invert the presentation of the numbering on the interface when the train changed ends. This finding was turned into a requirement and an algorithm developed for the design team.

For other features the new in-saloon CCTV control system, participants were asked to perform a series of tasks using the paper-based mock-up. They indicated by pointing to the picture which control they would select and explained why. The tester then changed the image based upon the participants' selection. Time to complete the task, the number of errors made and feedback from the participant on the interface were captured. A comparison was made between the original layout and the layout incorporating the human factors recommendations.

The results indicated that the interface design incorporating the human factors recommendations took less time for the participants to work out the correct sequence. This demonstrates that the interface is easier to use and has implications for time spent on training, time to use and reduced potential for error. Fewer errors were made using the revised interface and the subjective evaluation by the participants was more positive.

The proposed enhancement to the interface evaluated in this study were fed forward to the design team prior to detailed design. This results in cost and development savings, with the knowledge that the proposals had already undergone first level evaluation and gained approval from operators.

As with case study 2, additional benefits were realised within the project team through the ability to convey requirements clearly and visually, thus demonstrating their benefit to parties in different locations using different languages.

4 CONCLUSIONS

Human factors has a lot to offer to the design and development of all railway systems, not only rolling stock projects. Early identification of potential problems for operators, maintainers and passengers enables better user requirements to be defined and evaluated throughout the design lifecycle. The benefits of taking such an approach do not only include improved safety and operational performance enhancements. When applied appropriately, it can lead to significant cost and time savings for the project team and improvements in communication between different parties in the team. To achieve these results it is essential that the need for human factors analysis is identified at the start of the project and focused on those issues that will add value to the project. The requirements for integration in the project lifecycle can then be determined.

REFERENCES

BS EN-0126 1999. *Railway applications. The specification and demonstration of reliability, availability, maintainability and safety*. British Standards Institution.

ISO-13407 1999. *Human-centred design processes for interactive systems*. International Organization for Standardization.

Railway Safety and Standards Board 2007. *Engineering safety and management* (The Yellow Book), Issue 4. London: Railway Safety and Standards Board.

Railway Safety and Standards Board 2008. Understanding human factors—A guide for the railway industry. *http://www.rssb.co.uk/expertise/human_factors/human_factors_good_practices_guide.asp*.

Sanders, M.S. & McCormick, E.J. 1992. *Human Factors in Engineering and Design*. NY: McGraw-Hill Inc.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Creative constructing meaningful forms through gestalt-context approach

Tsai-Lin Yang
National Yunlin University of Science and Technology, Yunlin, Taiwan
Department of Product Design, Shu-Te University, Kaohsiung, Taiwan

Ming-Chyuan Ho
National Yunlin University of Science and Technology, Yunlin, Taiwan

ABSTRACT: Product form can be sorted into two categories: functional and decorative. The former has been thoroughly discussed and implemented through product semantics in industrial design. The latter has the potential to create 100 times of value for products with identical function performance. Realizing the significance of product form and less knowledge in this respect, this study develops a method for such purpose. The Gestalt-Context Approach is proposed, which consists of six steps: collecting meaningful data, selecting impressive elements, constructing contexts by association, identifying optimal gestalt level and mainstream storyline, defining components for fabrication, and implementation. Based on a famous Chinese novel-Dream of Red Chamber, the design method was successfully employed for designs of 12 series of accessories on 12 legendary female figures. Accordingly, prototypes were made and presented to a renowned researcher for validation. Five can be directly identified and six can be recognized with some hints, and the remaining one lacks of clear character in the novel. The six designs were submitted for competition and won six design awards. Moreover, the method was introduced to college students in product design major. The results were satisfactory and some of the students designs received awards from many design competitions. As a conclusion, it can be asserted that the new approach is feasible and effective.

Keywords: industrial design, product semantics, product form, Metaphor, topology

1 GENERAL INSTRUCTIONS

In the 1980s, semantics was a subfield derived from the Semiotics and had been introduced to the science and technology related industrial design field. After the implementation of semantics in this field, the product semantics emerged. Taiwan's design had been highly influenced by the Cranbrook Academy (United States of America) and hence Taiwan's design researches mostly were based on Charles E. Osgood's Semantic Differential Analysis (Lin, 2002). This method uses adjectives and multiple scale factor analysis to measure the product's forms and images in order to develop the "white box" application in the design field.

Product semantics mainly investigates on the meaning of a product delivered by its form, material and function, as well as how the user can relate, comprehend and use the product (Yang, 1998). Designers can design a new product by applying semantic analysis method to achieve a higher level of self-expression and interaction with the users. Generally, product form is divided into functional and decorative forms (Luh, 1996). The functional form emphasizes on the function and rational perception that relate to a system and logic of the science and technology; while the decorative form emphasizes on culture and emotions. Affective sequences and spiritual ideals are related to people with philosophy and literature knowledge. The current researches on product semantics put more focus on the functional forms and are fruitful; hence this research aims to focus on the product semantics for decorative forms.

Metaphor or analogy is the most common product semantic method applied in the product design. Analogy can be divided into five, four, or two styles. The four styles of analogy are the most widely applied in the product design. Explicit simile, implicit simile, synecdoche, and metonymy—all of these methods are applied on the icons and concepts to interpret or sculpture a new form of a product (Krippendorff, 1990). There is a clear logical boundary between the body comparison and the main body (Yang, 2003), which can help to understand the homologous model or rational perception of semantics and forms in a simple design or multi-layered creations.

Decorative form often involves multi-layered form metaphor. The meaning or function of the form is determined by the sequence, while the sequences can be the cognitive, usable, functional, situational or cultural ones. The representation or comprehension of the sequence cannot be deliberately separated in a single layer to interpret the design idea. In fact, it should be in the "Gestalt" concept. After implementing the Gestalt Concept in the forms and symbols, the emotional cognition model will build the whole effect and cultural effects in the product. By adopting Rudolf Arnhein's explanation of Gestalt Concept in which he took Brague as an example [1], it stated that "When lemon and orange are together, they were no longer called lemon and orange, they are fruits." When these two are together, the sequential relation or association will be generated. Lemon and orange build different layers of concept. If the way of positioning these two fruits can build an association, then it is possible to generate another Gestalt layer called "First Love" and hence the feeling of "sweet and sour" appears. Gestalt can be closed-ended or totally open-ended. McCoy (1990) presumed that the open-endedness would stimulate the audience to generate their own interpretations that will contribute to the contents of the idea.

In the open-ended Gestalt concept, the importance of sequence and layer lies on the relationship between the comparison body and the main body. The semantic of the functional form emphasizes on rational cognitive regarding functionality and is based on the coding to decide whether it is the key factor or not; whilst the decorative form emphasizes on the culture and emotion and is based on the sequence and layer to decide whether it is the core value or not. This research focuses on the application of semantics on product design, with an emphasis on decorative forms.

2 IDEA TREE

Form movement is a means of expression to communicate ideas with the second party. For designers, adjectives are the most frequently used keywords. Hence this research selects some adjectives that relate to the form as a fundamental basis to investigate on the Gestalt concept, and then using sequences and layers as tools to develop a new product semantic concept entitled the Idea Tree (Figure 1).

The layers in Idea Tree are derived from the design subject and each component should represent the information and design idea. This component is called "idea component." The adjectives in the idea components allow designers to communicate their ideas through words or symbols. Hence the idea components can be adjectives, allusions, metaphors, symbols, icons, and so on. In order to simplify the figure, the alphabets represent various idea components, whilst the abstract idea component is represented with "○" or a circle. Generally, the more components at the lower layer of the Idea Tree, the more layers can be established. This possibility depends on designers' creativity and ability of creating associations.

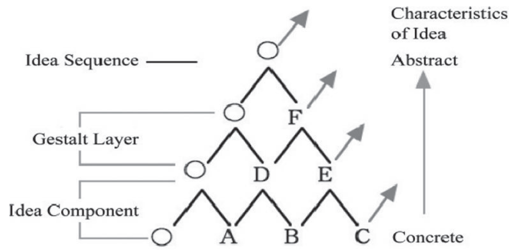


Figure 1. The sequence map of creative concept or "Idea Tree".

The Idea Tree consists of branches with different lengths. The bottom layer of idea branch is the designated idea component. Hence, idea component and branch of the Idea Tree exist based on the spatial concept. The Idea Tree can be continuously developed vertically and horizontally with infinite possibilities. The Gestalt Layer is positioned between the edge of idea branch in the idea component, e.g.: D is the Gestalt Layer of idea component A and B, while D is idea component to the upper level of Gestalt Layer F. By connecting different layers of idea components, the connected lines are called Idea Sequence.

Generally speaking, the idea sequence explains different layers of idea components or Gestalt Layers. However, it should follow the following perception: the lower the layer of an Idea Tree, the more concrete the idea component tends to be; whilst the higher, the more abstract it is. Hence, the Idea Tree can be perceived as a map to develop a creative concept that shows concrete to abstract, broad to narrow and clear layers and sequences of ideas. The designers can use this map as a medium of communicating ideas and the creative development that defines the creative boundaries and helps select the Gestalt Layers to build the design concept and semantic sequence of the idea components. This allows the designer to achieve design consensus, execution of idea component, work division and planning, collaboration and realization.

To make the Idea Tree clear and logical, an example based on Brague and "first love" is provided as illustrated on Figure 2. Since it is derived form

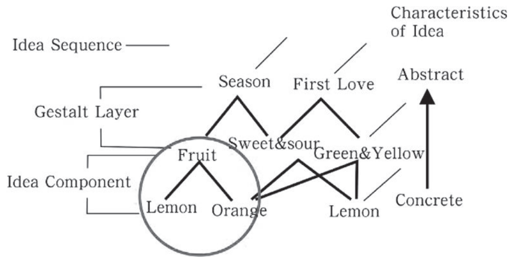


Figure 2. Application of Idea Tree based on "Fruit Lemon and Orange".

semiotics in literature, language thus plays key role in association and the following examples are sensitive to Chinese logic and literature. "Fruits" is the root of the other two idea components, "lemon" and "orange," as well as the branch of "Season." The development of Gestalt sequence of "First Love" is derived from the "Sweet and sour" and "Yellow and Green." The first idea component relates to the taste and the second one associate with the colors. The characteristics developed in this process are from the concrete appearance of "fruits" and then gradually develops into an abstract feeling of "First Love." In other words, the idea sequence can rationally systematize the associations in the Idea Tree.

3 THE CONSTRUCTION OF IDEA TREE

Take "Pearl Tear" which is the first author's creation of the metalwork design, "Dream Weaver in the Red Chamber" as the example and explanation. "Dream Weaver in the Red Chamber" is the set of accessories, which are the symbols of the twelve maidens mentioned in the Chinese classical literature, The Dream of the Red Chamber. "Pearl Tears" centralized on the leading actress "Lin Dai-Yue" as the semantic target. There are five reasons indicating why she was chosen as the example. First, the story, The Dream of the Red Chamber is framed generation by generation; therefore, each character had distinct personality and features. Second, the characters in the book had both abstract (fictitious characters) and concrete (concrete story) characteristics. Third, the personality of Lin Dai-Yue was unique and well-known, so the creation can diminish redundant explanation. Fourth, the design of hair accessory mainly focuses on the decorative form; it was not affected by the function model. Fifth, metalwork was equal to the final outcome. The creative conception would display more reality on the product design. The following was the main steps of the construction of creative trees:

1. Data collection and comprehension:
   Designers or creators collected the products' information to realize key subject of artistic creation. The concept of this artistic creation will focus on the Dream of the Red Chamber; therefore, the appearance, personality, and talent of the characters should keep to the original. The following citations mainly described the leading actress, Lin Dai-Yue in the Dream of the Red Chamber:
   a. People merely felt sorrow about both Hsue Pao-Chai's virtue and Lin Dai-Yue's talent. Although they are unique, the particularity they had still hidden in the forest and buried in the snow, they would not obtain the perfect ending in their life. —Chapter five, "The Dream of the Red Chamber".
   b. A blade of the Chiang Chu (purple pearl) grass grew by the side of the San Sheng (thrice-born) stone, which is on the bank of the Ling (spiritual) river in the west. Shen Ying, a fairy of the Ch'ih Hsia Palace day by day moistened its roots with sweet dew. This purple pearl grass, at the outset, therefore tarried for months and years. (At a later period imbued with the essence and luxuriance of heaven and earth, and having incessantly received the moisture and nurture of the sweet dew, the grass divested itself and turned into a human nature, which gradually became perfected into the person of a girl.) "I have been," Chiang Chu fairy would often commune within herself, "the recipient of the gracious bounty of rain and dew, but I do not possess such water to repay it! If the fairy need to descend into the world and turned into the form of a human being, I will also betake myself thither, along with it; and if I can make restitution to the fairy, with the tears of a whole lifetime, I will feel worthy about it" "This resolution evolved the descent into the world of so many pleasure-bound spirits of retribution and the experience of fantastic destinies; and this crimson pearl blade will also be among the number." It is indeed ridiculous," interposed the Taoist. "I never have before the restitution of the tears! —Chapter one, "The Dream of the Red Chamber".
   c. Her two arched eyebrows were thick as clustered smoke with a little frowning wrinkle. She had a pair of eyes, which possessed both cheerful and sad expression, overflowing with sentiment. Her face showed some sorrow stamped on her two dimpled cheeks. She was beautiful, but her whole frame was the prey of a hereditary disease. The tears in her eyes glistened like small specks. Her balmy breath

was so gentle. She was as demure as a lovely flower reflected in the water. Her gait resembled a frail willow, agitated by the wind. Her heart, compared with Pi Kan, had one more aperture of intelligence; even her beauty of ailment appearance exceeded the ailment of Hsi-Tzu who was also the famous four beauties in the Chinese history. —Chapter three, "The Dream of the Red Chamber"

d. (Vain knitting of the brows.) The one is a spiritual flower of Fairyland; the other is a beautiful jade without any blemishes. If their destiny was not remarkable, why he still has come to meet her again in this present life? If the destiny that you mentioned, why their love affair become a vain dream? The one with her loneliness merely sighed all the time; and the other yearn and crave hopelessly. The one would be like the reflection of the moon in water; the other would be like a flower reflected in a mirror. Could you image that how many drops of tears can there be in the eyes? And how could they continue to drop from autumn to winter and from spring till summer time? —Chapter five, "The Dream of the Red Chamber"

2. Extraction of the idea components:
From the searching information, the author selected some key conceptions and adjective terms for being the idea components. According to the description from the book, the author extracted the significant content about "Lin Dai-Yue" as the idea components and listed at the bottom of Idea tree in Figure 3.

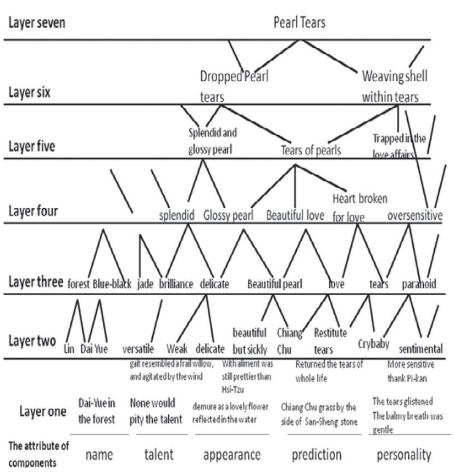


Figure 3. The construction of Idea Tree from "Lin Dai-Yue".

For example, "Jade girdle hanging in the forest" actually means "Lin", "Dai", and "Yue", the leading actress's name. "People would merely felt sorrow about the talent" extracted the "talent" as the creative component. "Her beauty of ailment appearance exceeded the ailment of Hsi-Tzu" extracted out the "beautiful but sickly" and "I can make restitution to the fairy, with the tears of a whole lifetime" extracted out the meaning of "restituted tear".

3. Stacking Gestalt Layers:
By applying metaphor and imagination and combing the semantic relationships between the creative components, the Gestalt Layer would be stacked in order. For example, according to "Dai (Blue-Black)" and "Yue (Jade)", people can imaged from the words to blue-black jade; from "Restituted tear" and "Crybaby" to the "Tears"; furthermore, from "Crybaby" and "Sentimental" to the "Paranoid" etc. If confronting the difficulty with stacking layers, the designers could follow the suggestion listed on the right side to arrange the idea components at the bottom. It would be beneficial for the development of Gestalt Layer. Or they can change the position of idea components at the bottom. Besides, shifting and adding components would be allowed.

4. Identify the best Gestalt Layer:
From the top to the bottom, the best innovative notion would be selected from the conception of Gestalt Layers. Basically, the more creative Gestalt Layer should get rid of general-standard reaction and comprise more idea components at the bottom. For example, "Blue-black jade" of the layer three generally represent "Lin Dai-Yue" at this Gestalt Layer; however, "Blue-black jade" merely was translated from her name, it could not contain other elements such as her talent, personality, and appearance etc. In Layer five, "Tears of pearls" could symbolize the fable, "Chiang Chu fairy restituted tears" and tender appearance, but it lacked for the world of devoted love affair and sentimental emotion. On the top of Gestalt Layer, "Pearl Tears" was decided as the final denomination. The meaning of "Pearl Tears" was wider and deeper but it would have some excluding part such as "Lin", "Dai", and "Yue" was the creative branches. The action that the author added "Lin Dai Yue", three words in the layer is acceptable. It might create higher Gestalt Layer. However, the complication of model accomplishment and difficulty of model reorganization would be increased after adding the words.

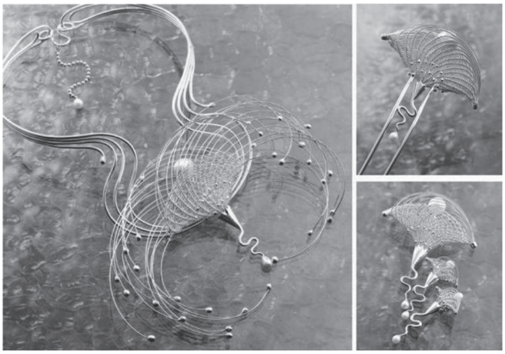


Figure 4. Part of the final artworks of “Pearl Tears”.

5. Definition of the form components and methods:
   According to the selected Gestalt Layers, the designers or creators can clarify the imagination range of creative branch through creative sequence; then, defining the modeling methods and components was the very step. For example, “Weaving shell within tears” indicated the weaving methods, “Trapped in the love affairs” expressed the density level of weave, and “Heart broken by the emotional affairs” implicate that pearls would represent tears.
6. Idea realization:
   Following the standard design procedures or methods of creation, the final artwork “Pearl Tears” (Figure 4) was finished.

4 ANALYSIS AND DISCUSSION

The construction of Idea tree applied on the conceptive creation of the twelve Maidens in the “Dream Weaver in the Red Chamber”. It would successfully obtain the Gestalt Layers of each character to develop new accessory designs, which are significantly different from the ones that ancient woman wore as people known. The proposed design approach possessed feasibility. Performing the creation model mentioned above, the designer would easily create a series of artworks and invented new weaving technique of metalwork design. The creation model would also contain usability.

The designer asked for advices about the a series of accomplishing works from the expert of the Dream of the Red Chamber, Professor Lin San-Ching’s who was the former dean of College of Liberal Arts. In the twelve artistic creations, Except the controversial of Hsue Pao-Chai’s characteristic, Prof. Lin could directly indicate five pieces and highly praised according to the interpreting characters, which include “Lin Dai-Yue”, “Shi Hsiang-Yun”, “Miao-Yue”, “Wang Hsi-Feng”, and “Chin Ke-Ching”. Other works also could be recognized and comprehended successfully by the brief cues. As the situation expressed, the construction of Idea tree and the design methods can be validated.

Afterwards, the series of artistic creations was exhibited publicly for a month. Observing the audiences’ reactions, people who know about the twelve maidens in the dream of the red chamber could recognize the works that corresponded to the subject characters without the cues or the brief explanation. “Pearl Tears” also was the entry in the handicraft works group of “The fifty-seventh Taiwan Province Fine Art exhibition”. It was nominated by the initial appraising and won the Gold award in the metalwork category. The entries of handicraft creation were total one hundred and twelve pieces. Thirty-three pieces were nominated. The rate of passing the first trial was 29%; furthermore, total entries in each artistic category were 2101 pieces and the entire nominated rate was 17.4% (Taiwan Art Exhibition, 2003). The above statistics indirectly support the method’s validity. Summarizing the points above, the proposed conception and creation model possess usability, feasibility and validity with certain levels.

5 CONCLUSIONS

Based on the discussions above, the following conclusions can be drawn:

1. In the open-ended Gestalt concept, sequences and layers are the crucial parts that establish the comparison body and the main body.
2. Idea Tree allows the development of rational and systematized structure to assist designers to acquire abundant amount of sequential ideas or Gestalt Layers which can function as a useful design tool.
3. Idea Tree can be implemented as a tool to select and communicate design ideas to help designers reach a consensus, work division and planning, and other related design tasks.
4. The structure and application of the Idea Tree are applicable on design procedures or creation model, as it can converge design ideas, help select the appropriate Gestalt Layer, establish design subject, define the form component and method, and so on.
5. The creation Model can be applied on the Classical Literature. In this case, it is a breakthrough on the application of product semantics.

6. “The Dreamweaver of the Red Chamber” is a series of accessories based on the “The Dream of the Red Chamber” and product semantics in the industrial design field. The design involves technical application of metal weaving. It not only possesses creative cultural characteristics and research fundamental, but also contributes to the interaction between the academics and arts and literature.

REFERENCES

Krippendorff, K. 1990. Product semantics: a triangulation and four design theories. *Proceedings of Product Semantics ’89 Conference*. University of Industrial Arts, Helsinki, 3–23.

Luh, D.B. 1996. The design theory of force development model. *Technical Journal* 12(1): 317.

Lin, M.H. 2002. Product design in form of encoding and decoding. *Design* 2: 39–52.

McCoy, K. & Michael. 1990. The new disclosure. In: *New cranbrook design disclosure*, New York: Rizzoli.

Taiwan Art Exhibition 2003. The whole province of the fifty-seventh province of Taiwan Art Exhibition. preliminary short-ist, retrieved 2003/3/13 from *http://www.tpg.gov.tw/dcea/action_show.asp?id = 78*

Yang, T.L. 2003. *A dream of red weaving - semantic method used in classical literature, jewelry creation*. Tainan National College of Art master’s thesis.

Yang, Y.F. 1998. *Design and cultural foundation*. Taipei: Asia-Pacific.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The relationship between computer self-efficacy and playfulness of mobile phone

Wei-Jen Chen & Li-Chieh Chen
Interaction Design Laboratory, Tatung University, Taipei, Taiwan

ABSTRACT: In this study, there were 57 participants join this survey. First, the Adult Playfulness Scale was used for exploratory factor analysis and three factors being generated which were 1. Use mobile phone as entertainment device. 2. Use mobile phone to acquire online information. 3. Use mobile phone to build social network. The regression analysis between three factors and Computer Self-Efficacy Scale indicated that people with stronger computer self-efficacy are more likely to use mobile phone as entertainment device such as music, game, eBook and movie. However, people with stronger computer self-efficacy are not stand for good interpersonal communication.

Keywords: mobile phone, playfulness, entertainment, computer self-efficacy

1 INTRODUCTION

From game to instant messaging, today's mobile phone entertainment is more and more popular. As we can see wild variety functions are added into the mobile phone such as GPS, micro-blogging, mobile online store … etc., There was a survey, Informa Telecoms & Media (2007), indicated that mobile game's global revenue in 2006 was more than $2.5 billion, and within five years, there will have three times revenue as much as in 2006. According to a survey, University students would like to spend extra money for important features of the mobile devices. These features includes such as 3G service, MP3 players and Bluetooth (Anastasios, 2009). This trend brings people more familiar with these new technology inventions, because they can play these functions right in the pocket. Both genders are positive with new mobile technology. Also females are more care about the mobile phone's appearance than males do (Anastasios, 2009).

Once mobile phones have more high tech features such as news and entertainment being installed, users will gradually accept these features and try to use them (Ran, 2008). Now, mobile phone not only for basic functions, but more entertainment contents being created. Besides making a phone call, mobile phones are commonly used for seeking news, web surfing and playing game (Ran, 2008). However, for user to download music, video and play game, specific skills are necessary such as install software or hardware know-how (Monica, Deborah, 2007). Without these particular skills, user might not play game, listening music or read eBook. Besides these digital contents, the mobile phone developer might need to consider what make user like to play or what content entertain the users. There is a research mentioned: When users are doing mobile entertainment, they can be affected by ambient noise or light, however these factors are ignored by designer (T. Serif, G. Ghinea, 2008). Therefore, these design issues need more concern and pay attention. Another situation might be interesting is that just like PC, mobile phones have more applications support. Despite many functions/services are added into the mobile phone, people are not automatically using all of them (Hannu et al., 2010). And what makes mobile device playfulness, the most important feature is to make it easy to use. Mobile phone may not as powerful as desktop computer, but it provide great mobility which make user love to use. (Gordon & Anand 2005).

Base on the literature review, there are three hypotheses to test the relationship between computer self-efficacy and mobile playfulness:

H1a: The stronger the computer self-efficacy that users possess, the more likely they will use mobile phone for entertainment.
H1b: The stronger the computer self-efficacy that users possess, the more likely they will use mobile phone to acquire online information.
H1c: The stronger the computer self-efficacy that users possess, the more likely they will use mobile phone to build social network.

2 METHODOLOGY

There were 57 participants (28 female and 29 male) join this survey. First, the questionnaire included demographic questions such as gender, age, education and marital status. The mean age in this survey was 29.7 (SD = 8.41) with a range from 21 to 44 years old. 72% had graduated from college and 45.6% were married. All participants have online experience in using mobile phone. 32.4% had applied 3G telecom service, and 11.7% had checked email via their mobile phones everyday. Second, we developed a questionnaire with "Adult Playfulness" Scale in mobile phone and the items included in the scale such as 'entertainment purposes', 'acquire information online'and 'social network activities'. There were ten statements based on five-point Likert scale which 1 meant "strongly disagree" and 5 meant "strongly agree."

We also employed "Computer Self-Efficacy" Scale in the questionnaire. Participants were asked to answer eight statements based on five-point Likert scale. The statements included such as 'I feel confident to solve my own computer problem', 'I can easily exchange file between mobile phone and computer' and 'I treat solving computer problem as self-educate opportunity. The reliability analysis of two scales' Cronbach α were 0.81 (Adult Playfulness) and 0.86 (Computer Self-Efficacy) which the internal consistency were approximately identical ($0.7 < \alpha \leq 0.9$).

3 RESULT

In order to simplified the variation and understand the factor structure, the "Adult Playfulness" Scale was subjected to an exploratory factor analysis. Before processing factor analysis, the data was subjected to Kaiser-Meyer-Olken (KMO) .71 and Bartlett's Test of Sphericity 14.23 (df = 7) to make sure the suitability. And the result for KMO and Bartlett's achieved statistical significance which indicated the data were suitable for processing exploratory factor analysis.

The ten statements were analyzed using exploratory factor analysis (principal axis factoring, PAF) with Varimax rotation. (Table 1) shows the 3-factor solution result, including the factor structure, variance explained, and eigenvalue. There were 3 components with eigenvalues greater than 1, the eigenvalues of 3.67, 2.19 and 1.76, respectively and the total variance explained 77.42% (31.64%, 24.83% and 20.95%, respectively). The result of PAF can well represent the original variance. According to the "Rotated component matrix" of Varimax rotation, The factor 1 included four statements that emphasize "personal" enjoyment and playfulness. Therefore, the factor 1 was labeled as "Use mobile phone as entertainment device". The factor 2 has three statements that were mainly indicating news gathering and to help users have instant information response. So, the factor 2 was labeled as "Use mobile phone to acquire online information." The third factor consisted of interpersonal network, instant message and micro-blogging, so we labeled the third factor as "Use mobile phone to build social network".

Table 1

Factor analysis matrix	Factor 1	Factor 2	Factor 3
Use mobile phone as …	Entertainment device	Acquire online information	Build social network
Listen music	**.87**	−.09	−.18
Play game	**.82**	.03	.26
Watch movie	**.71**	.16	.35
Read eBook	**.62**	.39	.07
Stock	.17	**.69**	.23
Weather	−.24	**.63**	−.04
Email	.29	**.57**	.28
Instant messaging	.21	.26	**.83**
BBS	.30	.37	**.67**
Facebook/Twitter	.24	.11	**.59**
Eigenvalues	3.67	2.19	1.76
Variance Explained (%)	31.64	24.83	20.95

Note: Higher factor loading appears in boldface.

Based on the "Correlation Matrix" (Table 2), the above-mentioned three factors were highly correlated with each other.

In the regression analysis, we set three factors as dependent variable and computer self-efficacy as predictors. In the first regression "Use mobile phone as entertainment device", the result indicated that computer self-efficacy = .41, t = 5.38, $p < .001$ (Table 3) was significant predictors and supported H1a. This accounted for 47.6% of the total variance. The second regression "Use mobile phone to acquire online information", the result indicated that computer self-efficacy = .26, t = 3.17, $p < .001$ was significant predictors and supported H1b. This accounted for 24.9% of the total variance. The third regression "Use mobile phone to build social network", the result indicated that computer self-efficacy = .07, t = 1.43 was not significant predictors, so the H1c is not supported. This accounted for 11.6% of the total variance.

Table 2. Correlation matrix.

Factor	1	2	3
1. Entertainment device	1.000	.713	.532
2. Acquire online information	.713	1.000	.492
3. Build social network	.532	.492	1.000

Table 3.

Computer self-efficacy	R^2	β
Factor 1	.476	.41***
Factor 2	.249	.26***
Factor 3	.116	.07

*$p < .05$
**$p < .01$
***$p < .001$

4 DISCUSSION

There are three types of mobile playfulness been generated in this survey. They were 1. Use mobile phone as entertainment device, 2. Use mobile phone to acquire online information, 3. Use mobile phone to build social network. Based on the regression analysis, two hypotheses were support. One is that people with stronger computer self-efficacy are more likely to use mobile phone as entertainment device such as music, game, eBook and movie. The other is that people with stronger computer self-efficacy are more likely to acquire online information such as instant stock feedback, weather and email. As for the third hypothesis, the stronger the computer self-efficacy that users possess, the more likely they will use mobile phone to build social network, was not supported. This result can be seen that people with good computer skill or knowledge are not stand for good interpersonal communication. For instance, computer geek might proficient in solving mobile phone issues, but not interested in using mobile Facebook connecting friends or building social network.

The H1a and H1b were supported can be explained that factor 1 and 2 were more related to personal hobby, information or privacy such as listen music, play game or check email/stock. These activities require certain computer skills to transfer digital date to mobile phone or install certain software to activate the function on mobile phone. Therefore, with stronger computer skill, users are more likely to transfer their computer professional ability onto the mobile phone; which also means, if they encounter certain problems, they are more likely to have abilities to solve by themselves instead of giving up. Overall this study revealed important facts of people who use mobile phone for playfulness and the fact between computer self-efficacy and mobile entertainment.

Further research might consider personal entertainment with mobile cognitive. The individuals with different real world hobby or interest might have variety of cognitive in mobile entertainment. In conclusion, mobile entertainment is growing more rapidly than ever before. In order to fulfill users' expectation, researchers in mobile entertainment field need further consideration in users' psychological interests and cognitive.

REFERENCES

Anastasios, A.E. 2009. Students' thoughts about the importance and costs of their mobile devices' features and services., *Telematics and Informatics* 26(1): 57–84.

Gordon, C. & Anand, K. 2005. Explaining consumer acceptance of handheld Internet devices. *Journal of Business Research* 58(5): 553–558.

Hannu, V., Carolina, L., Francisco, J. & Harry, B. 2010. Analysis of users and non-users of smartphone applications. *Telematics and Informatics* 27(3): 242–255.

Liang, T.P. & Yeh, Y.H. 2009. Situational effects on the usage intention of mobile games. In *Lecture notes in business information processing: Designing e-business systems, markets, services, and networks* 22(2): 51–59.

Monica, T. & Deborah, M. 2007. Online recreation: The relationship between loneliness, Internet self-efficacy and the use of the Internet for entertainment purposes. *Computers in Human Behavior* 23(3): 1435–1446.

Ran, W. 2008. Motivations for using the mobile phone for mass communications and entertainment. *Telematics and Informatics* 25(1): 36–46.

Serif, T. & Ghinea, G. 2008. Mobile information access in the real world: A story of three wireless devices. *Computers in Human Behavior* 24(4): 1385–1403.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Future architecture and past architecture in virtual reality

Kaihsiang Liang
Graduate Institute of Architecture in National Chiao Tung University, Hsinchu, Taiwan

ABSTRACT: Architecture media are the tools for the communication between architectural designers and the general public. Compared to the architectural non-digital media, digital media can provide more information, such as virtual reality. Virtual reality can make viewers feel being inside virtual architectural design rather than only observe it. Therefore, the public can experience the architectural designs which they can not experience in the past, including future architecture and past architecture. My problem is what the differences between the presentations of two architectural types with virtual reality are. The conclusion is because virtual reality itself can cause the screen to distort. To avoid the picture distortion, the lenses inside virtual reality are required careful adjustment. The contribution of this study provides a guideline for designers to easier create a future architecture or past architecture in virtual reality.

Keywords: virtual reality, architectural media, past architecture, future architecture

1 INTRODUCTION

Architecture media are the tools for the communication between architectural designers and the general public. Compared to the architectural non-digital media, digital media can provide more information, such as virtual reality (Bertol & Foell 1997). Virtual reality can make viewers feel being inside virtual architectural design (Kalay 2004), rather than observe it (Bridges & Charitos 1997; Burdea & Coiffet 2003; Wu 2009). Therefore, the public can experience the architectural designs which they can not experience in the past.

About the so-called "architectural designs which people can not experience in the past", this study divided them into three types. The first is future architecture which is not yet constructed (Liu 2007). The second is past architecture whose original design has been changed (Mitchell et al. 2007; Chang et al. 2007). The third is conceptual architecture. Because different conceptual architectures vary greatly, this research chooses to focus on the future architecture and past architecture, and discuss the differences between the presentations of two architectural media in virtual reality. The objective is to discover how people determine what kind of architecture is future architecture and past architecture, and what affects their judgments?

2 METHODOLOGY AND STEPS

My experiments are divided into two phases: the experiment of "one whole space", and the experiment of "spatial details." The first phase, experiment of one whole space in virtual reality, is to test how people determine what kind of one whole space is future architecture or past architecture. The second phase, experiment of the spatial details in virtual reality, is to test how people determine what kind of spatial details belongs to future architecture or past architecture.

2.1 *Experimental devices*

Experimental devices mainly include a rear projection screen which can demonstrate virtual reality, whose size is 120 inches and one joystick which can make users easily move in virtual reality (Fig. 1).

2.2 *Experimental material*

There are two experimental materials, each of which is represented future architecture and past architecture. The future architecture is the museum of art in National Chiao Tung University, which is not yet built (Fig. 2); the past architecture is the consul's residence in Fort San Domingo, built in 1981 (Fig. 3). The museum of art in National Chiao Tung University is designed by Tadao Ando. Because of lacking funding, it will not be completed in the recent years. The consul's residence in Fort San Domingo belongs to Grade I of historic monuments in Taiwan. However, it had been rebuilt several times. The experimental material is from original design. I have created the 3d models by floor plans, elevations, sections, and physical models. The 3D models are available for

Figure 1. Experimental devices.

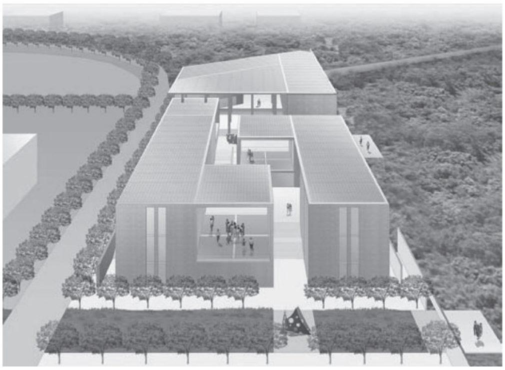


Figure 2. The rendering of the museum of art in National Chiao Tung University.

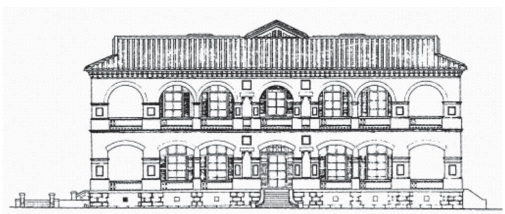


Figure 3. The elevation of the consul's residence in Fort San Domingo.

showing three-dimensional rendering of the low numbers of digital surface in real time, including render to texture.

2.3 *Experimental steps*

In the warm-up experiment, subjects freely moved in a non-experimental space to make themselves understand the mode of operation in virtual reality. In the formal experiment, subjects browsed the designated space first. This study chose six rooms to be experimental space, three from the museum of art in National Chiao Tung University (Fig. 4) and three from the consul's residence in Fort San Domingo (Fig. 5). Then, I asked subjects to judge that his location is future architecture or past architecture by four spacial factors and their sub-factors. Then I asked the subjects what factors affect them to judge. All data which the subjects answered are recorded by Likert scale.

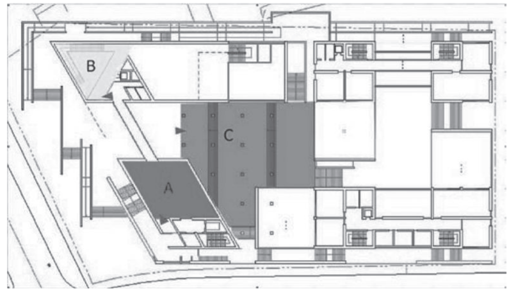


Figure 4. The floor plan of the museum of art in National Chiao Tung University.

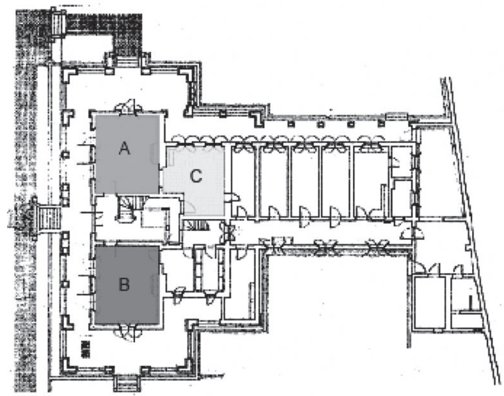


Figure 5. The floor plan of the consul's residence in Fort San Domingo.

3 THE EXPERIMENT OF "ONE WHOLE SPACE"

3.1 *The spacial factors and sub-factors*

There are four spacial factors, which are spacial form, spacial scale, spacial circulation, and spacial organization. The sub-factors of spacial form are shape, edge, surface, and opening (Ching 1996). The sub-factors of spacial scale are reference object, perspective, distance (Ching 1996), size of texture, point of view (Danby 1962). The sub-factors of spacial circulation are path, stair (Krier 1988, Ching 1996), ramp (Danby 1962), and square (Bloomer & Moore 1977). The sub-factors of spacial organization are spacial property, entrance (Ching 1996; Liu 1996), spacial layout (Krier 1988, Ching 1996), spacial adjacent (Ching 1996).

The sub-factor, Shape, means the shape of the floor plan, where the view locates or look at. The sub-factor, Edge, means the edges in the space. The sub-factor, Surface, means the skin of walls, floors, and ceilings. The sub-factor, Opening, means the opening of walls, floors, or ceilings, such as windows and skylights.

Table 1. The spacial factors and sub-factors of the experiment of one whole space.

Main spacial factors	Spacial sub-factors
Spacial form	Shape
Edge	
Surface	
Opening	
Spacial scale	Reference object
Perspective	
Distance	
Size of texture	
Point of view	
Spacial circulation	Path
Stair	
Ramp	
Square	
Spacial organization	Spacial property
Entrance	
Spacial layout	
Spacial adjacent	

The sub-factor, Reference object, means a reference scale. If there is a window inside the space, it is easy for people to judge the size of the space. The sub-factor, Perspective, means the perspective angle of lens. The sub-factor, Distance, means viewing distance. If people stand too far away the window, it is more difficult to judge the size of the space. The sub-factor, Size of texture, means the size of the architectural materials. The sub-factor, Point of view, means viewing location in the space.

The sub-factor, Path, means the passageway. The sub-factor, Stair, means one step of a series for ascending or descending to a different level. The sub-factor, Ramp, means an inclined plane serving as a communication between different interior levels. The sub-factor, Square, means a public square with room for pedestrians.

The sub-factor, Spacial property, means the property of individual space, such as public space and private space. The sub-factor, Entrance, means a passage, door, or gate, for entering. The sub-factor, Spacial layout, means the layout of a whole architecture, such as concentration, linear, grid, etc. The sub-factor, Spacial adjacent, means the connection of space and space.

3.2 *The experimental analysis*

In terms of spacial form, the subjects generally considered past architecture has a simple form, and its space should be orthogonal; the contrary, subjects thought the polygonal floor plans should belong to future architecture. In addition, the experimental results show that subjects will misjudge the shape of floor plans, because virtual reality environment would cause a distortion of the presentation. An example is that subjects misjudged that an orthogonal space becomes a bevel space.

In terms of spatial scale, the main sub-factor impacting on subjects is openings. My speculation is that subjects can easily observe the openings in their daily life. If the shape of opening is irregular shape, the subjects tend to judge the space belongs to future architecture.

The contrary is past architecture. In terms of both spacial circulation and spacial organization, the subjects tend to judge the complex is future architecture. On the contrary, the simple is past architecture. Although spacial factors can reveal some differences between the two, but overall the future architecture and past architecture have only a few differences.

4 THE EXPERIMENT OF "SPATIAL DETAILS"

4.1 *The spacial factors and sub-factors*

There are four spacial factors, which are material effect, texture of surface, historic effect, and architectural element. The sub-factors of material effect are material shape (Danby 1962), material color (Franz 2006), material proportion (Legakis et al. 2001). The sub-factors of texture of surface are material tiling, light and shadow, material bump (Kolarevic 2006). The sub-factors of historic effect are manufacturing era (Giedion 1963), manufacturing method (Allen & Agrest 2000), and architectural style. The sub-factors of architectural element are opening, wall, structure (Liang 2009).

The sub-factor, Material shape, means the shape of one material unit. The sub-factor, Material color, means the colors of one material unit. The sub-factor, Material proportion, means the proportion one material unit.

The sub-factor, Material tiling, means segmentation line of the surface, such as floors, walls, and ceilings. The sub-factor, Light and shadow, means the expression of light and shadow. The sub-factor, Material bump, means the bump of the surface.

The sub-factor, Manufacturing era, means the feeling of old or new. The sub-factor, Manufacturing method, means the method by artificial system of mechanical system. The sub-factor, Architectural style, means the mode of expressing thought the form of architecture.

The sub-factor, Opening, means a place which is open. The sub-factor, Wall, means an architectural

Table 2. The spacial factors and sub-factors of the experiment of spatial details.

Main spacial factors	Spacial sub-factors
Material effect Material color Material proportion	Material shape
Texture of surface Light and shadow Material bump	Material tiling
Historic effect Manufacturing method Architectural style	Manufacturing era
Architectural element Wall Structure	Opening

structure, raised to some height, and also, one of the upright inclosing parts of the space. The sub-factor, Structure, means the manner of one architecture.

4.2 *The experimental analysis*

In terms of material effect, the subjects think that past architecture has a diversity of architectural materials, and relatively future architecture has fewer materials. In terms of texture of surface, the subjects think the architecture making them feel strongly is past architecture; the contrary is future architecture. About these two findings, the study speculated that the subjects are affected by their past experience. In the real world because the materials become dirty, the colors of materials will be mottled. Therefore, if wanting to design a past architecture in virtual reality, this study recommends designers to use a variety of colors.

In terms of historic effect, the space with more edges makes subjects feel it is old; the other hand, the more simple space is easy for the subjects considered to be the future architecture. That is because subjects think past architecture has more decorative details. However, some spaces in virtual reality have the limitation of mapping, so they are difficult to distinguish between past architecture and future architecture.

In terms of architectural elements, the subjects think future architecture either hides its structure. It is because new technology makes structure light and thin. If an architectural structure is purely functional, the subject will see it as past architecture.

5 CONCLUSIONS

The conclusion is because virtual reality itself can cause the screen to distort. If the distortion is more, it is easier for the viewers think the architecture on the screen is the future architecture. To avoid the picture distortion, the lenses inside virtual reality are required careful adjustment. When judging architecture is future architecture or past architecture in virtual reality, people generally think a future architecture is polygonal floor plans, irregular opening, complex spacial circulation and organization, fewer materials, variety of colors, simple space, and hidden structure. The Contrary is past architecture. The contribution of this study provides a guideline for designers to easier create a future architecture or past architecture in virtual reality. The future research will be adding real-time interactive virtual reality features.

REFERENCES

Allen, S. & Agrest, D. 2000. *Practice: architecture, technique and representation*. ISBN 90-5701-032-1. Amsterdam: G+B Arts International.

Bertol, D. & Foell, D. 1997. *Designing Digital Space: An architect's guide to virtual reality.* John Wiley & Sons, Inc.

Bloomer, K.C. & Moore, C.W. 1977. *Body, Memory and Architecture*, Yale University Press.

Bridges, A. & Charitos, D. 1997. On architectural design in virtual environments, *Design Studies* 18(2): 143–154.

Burdea, G.C. & Coiffet, P. 2003. *Virtual Reality Technology*. A Wiley-International publication.

Chang, L.H., Hou, J.H. & Liu, Y.T. 2007. New media in architectural heritage—process, standards and the database platform. In *Proceedings of DACH 2007 (NCRPCP)*, 191–202.

Ching, F.D.K. 1996. *Architecture-form, space, & order*. John Wiley & Sons.

Danby, M. 1963. *Grammar of architectural design: with special reference to the tropics*. Oxford University Press.

Franz, G. 2006. Space, color, and perceived qualities of indoor environments. In M. Tolba, S. Soliman and A. Abdel-Hadi (Eds.) *Environment, Health, and Sustainable Development - IAPS 19*.

Giedion, S. 1963. Space, time and architecture: the growth of a new tradition. *Cambridge, MA: Harvard University Press.*

Kalay, Y.E. 2004. *Architecture's new media:principles, theories, and methods of computer-aided design.* MIT Press.

Krier, R. 1988. *Architectural Composition.* Rizzoli, New York.

Kolarevic, B. 2006. Manufacturing surface effects, synthetic landscapes, In *Proceedings of the 25th Annual Conference of the Association for Computer-Aided Design in Architecture,* 95–103.

Liang, K.S. 2009. Redefining architectural elements by digital media. In *Human Centered Design – HCD 2009 (HCII 2009)*: 995–1002.

Liu, Y.T. 1996. *Understanding of architecture in the computer era.* Hu's: Taipei.

Liu, Y.T. 2007. *Tomorrow architecture Exhibition-physical, virtual, interactive, space*. Taipei Fine Arts Museum Exhibition Album.

Rasmussen, S.E. 1964. *Experiencing Architecture*. The MIT Press.

Mitchell, W.J., Beamish, A. & Allibhai, S. 2007. Preserving architectural heritage in the digital era - archNet's lessons and future directions. In *Proceedings of DACH 2007*: 3–28.

Wu, Y. 2009. Some Phenomena in The Spatial Representation of Virtual Reality, In *Proceedings of the 14th International Conference on Computer Aided Architectural Design Research in Asia*, 143–152.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Difference in designers' ideation process between referring to printed stimuli and computerized stimuli

Pei-Jung Cheng
Asia University, Taichung, Taiwan

Li-Hao Chen
Fu Jen Catholic University, Taipei, Taiwan

ABSTRACT: The main purpose of this study is to explore the characters and differences between those designers who refer to relevant stimuli from printed reference and from computerized reference on-line. For this purpose, the observations with four graphic design practitioners were conducted in this study. Finally, two different thinking ways are utilized to individually describe the ideation process of the designers with two different referring behaviors. Besides, the characters and differences are found out from the ideation process of the designers who inspired by printed stimuli and inspired by computerized reference on-line. And two different idea development methods are also found out from the idea transforming process of the designers who inspired by the two different kinds of reference.

Keywords: referring behavior, graphic design, ideation process

1 INTRODUCTION

Most researchers argue that visual data can trigger designers' mental image and help them to develop appropriate solutions in the design process (Dorst & Cross 2001; Suwa et al. 2000; Verstijnem et al.1998; McGown et al. 1998; Goldschmidt 1994; Herbert 1993; Schön & Wiggins 1992). That is to say, visual stimuli influence on designers' performance heavily when they are facing with various design problems (Goldschmidt & Smolkov 2006; McCoy & Evans 2002). The result found in related research shows that designers do need to refer to relevant information in their ideation and most of designers spend over one-third of their ideation time on searching for related reference to come up with the design problem (Cheng & Yen 2008). Moreover, Petre et al. (2006) found from observing designers working process to propose that designers tend to select and adapt a given source or incorporate the source into their imagination to produce indirect or unexpected ideas. Thus, various references absolutely play the significant role in designers' ideation process and should be further considered about.

However, nowadays most designers are getting used to retrieving relevant information or visual data on-line for finding some inspirational data during their ideation process, rather than just searching for the relevant information through reference books. In terms of the important change in the digital environment, we conjunct that the habit of referring to some computerized data by keying in keywords on-line could change designers' thinking path and idea generating way. In this study, we assume that designers refer to the printed reference they searched for from books and the computerized reference they retrieved on-line may have different influences on designers' ideation process. Therefore, the purpose of this study is to explore: 1) The character that designers refer to relevant information through searching for it from printed reference; 2) The character that designers refer to relevant information through retrieving it on-line; 3) The difference between the two types of referring behaviors in designers' idea transforming process; and 4) The main influence on designers' ideation process by the inspirational reference they searched for from books and they retrieved on-line.

2 METHOD

Based on the research hypothesis and purpose of this study, we conducted observations with four graphic design practitioners in their workspace to explore the major characters and differences in the ideation process between the two types of referring behaviors. Among the four participants, two designers tend to search for visual stimuli from some books and the other two designers tend to seek for some needed reference on-line. The comparison and analysis of the characters and differences between the two types of referring behaviors are made after the observations are carried out.

2.1 *Participants and tasks*

Every graphic practicing designer in this study had at least two-year work experience and was not aware of the purpose of the research before taking the task. Three of those participants individually work for a design firm and another one works in his graphic design studio. The four participants' background and information are presented in Table 1.

For exploring the characters and differences in the two types of referring behavior, the study made an arrangement on the group and the task for each participant in the observations. We divided the four participants into two groups, by which every group includes one participant who prefers referring to printed reference in the ideation process but another one prefers referring to computerized data on-line. Besides, two participants with different referring behavior in one group were assigned the same task. For instance, the members in group 1 includes participant A and B, one is used to searching for stimuli from books in his ideation but another is used to retrieving relevant data on-line, and both of them should take the same task to 'design a DM for the New Year firework show'.

By making the arrangement, we may observe that whether the similar characters could be captured during the ideation process of the two participants, with the same referring behavior in different group, even if they are performing a different task. Besides, we may also observe that whether the differences could be captured during the ideation process of the two participants, who take the same task in one group, while they are having different referring behavior. For completely showing the arrangement about each participant's group and assigned task in this study, Table 2 is set.

Moreover, the study draws two behavioral codes from previous studies (Cheng 2010; Cheng & Yen 2009; Cheng & Yen 2008) to show the two types of referring behavior, which are the codes 'SI' and 'RI' (refer to Tab. 2). The code 'SI' marked in Table 2 shows the participants are used to searching relevant information from books and only had this kind of referring behavior during their ideation when taking their task. The code 'RI' means that the participants are used to retrieving relevant information on-line and only had it during their ideation.

Table 1. Four participants' individual information.

Participant	Sex	Experience	Type of company
A	Male	2 years	Design firm
B	Male	6 years	Individual studio
C	Female	9 years	Design firm
D	Female	2 years	Design firm

Table 2. Arrangement of four participants' group and task.

Group	Participant	Code	Task
1	A B	SI RI	To design a DM for the New Year firework show
2	C D	SI RI	To design a logo for a traditional food restaurant, named 'Original Taste'

*'SI' shows the behavior of searching for information from books; 'RI' shows the behavior of retrieving information on-line.

2.2 *Procedure of the observations*

In the observational process of this study, each participant was asked to take an assigned task in his/her working space. Designers participated in the observation were given the greatest freedom and unrestricted searching for relevant data to perform the task as the situated designing happened in their everyday works. Besides, they were given the needed time to perform the assigned task in spite of the progress rate they achieved. Since designers may be reticent or unable to verbalize simultaneously whilst generating ideas (Nagai & Noguchi 2002), we merely monitored and captured their nonverbal behaviors by a digital camera and collected the idea sketches they have done during their ideation process. The digital camera was placed to right-back or left-back side of the designer (see Fig. 1). We started recording their ideation process after they read a provided task exposition and finally stopped it when the participant informed us an expected idea has been made to serve as the final representation.

2.3 *Data analysis*

The observational data and designers' idea sketches then were analyzed and examined by three researchers who are in design field. The analytic process includes two sections: firstly, the three researchers have to go through each participant's

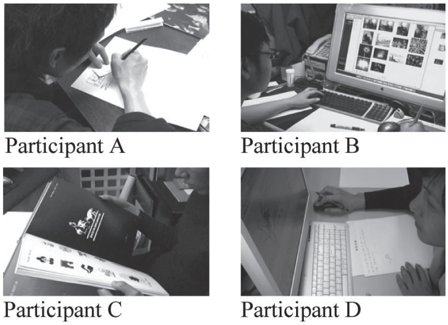


Figure 1. Monitoring method in the observations.

video of taking the assigned task for noting down the characters of the two referring behaviors they have found from those data; secondly, they have to sort each participant's idea sketches into different categories. In doing so, we could make the comparison of the participants' ideation process in referring to printed reference and referring to computerized stimuli. Besides, the two different referring behaviors' influences on designers' presentation could be figured out as well after examining the sorted categories of the four participants' idea sketches.

3 RESULTS AND DISCUSSIONS

In the analytic section, three researchers totally went through about 240-minute length videoed data, in which the length of participant A's record is about 66 minutes, participant B's is about 63 minutes, participant C's is about 65 minutes and participant D's is about 46 minutes. Moreover, the statement should be made before we describing the characters and differences in participants' SI and RI behavior. That is, the analytic results show that: 1) the ideation process of the participants who had the same referring behavior (SI: participant A & C; and RI: participant B & D) do exist similar characters even though they were assigned a different task; and 2) differences in the ideation process between the two participants who took the same task in a group (participant A & B or participant C & D) are found as well. Therefore, the arrangement for those participants' task and group in the observations is absolutely effective for interpreting and discussing the results.

3.1 *Characters and differences of the two referring behaviors*

By examining and making a comparison of the three researchers' notes, several characters and differences in the two kinds of referring behavior, SI and RI, should be noticed and described. We then address them individually as follows.

1. Characters and differences in the ideation process of the participants who only had SI behavior.

   There are five main characters captured from the videoed data of participant A and C by the three researchers of this study. Firstly, after the two designers understood their task, they started with flipping through the books until the inspirational images were found out. Secondly, they like to paste memory stickers on the targets they found out from the printed material or books. Thirdly, their ideation flow shows that they were frequently creating new idea sketches while searching for relevant information from books. That is, their SI referring behavior is found that has a strong interaction with their behavior of making new idea sketches. Fourth, they tend to extract several specific elements from the marked targets in the printed material or books and combine them together to form a new idea sketch. Lastly, in the two participants' ideation process, there is no idea was noted down in written form as the participants with RI behavior did.

2. Characters and differences in the ideation process of the participants who only had RI behavior.

   There are four characters captured from the videoed data of participant B and D by the three researchers. Firstly, they started with generating ideas and noted them down in written form after they understood the task. Then some of the written ideas are the keywords they inputted in the search engine later for retrieving some inspirational images or information on-line. Secondly, they seem to continually get next keyword by referring to some information or images they have retrieved. Thirdly, they tended to complement their written ideas while they were referring to the retrieved data on-line. Lastly, they would like to revise previous idea sketches or add some other elements on them rather than frequently create a new one as the participants with SI behavior did. Besides, some elements they added on the previous idea sketch were inspired by the images they found in some web pages.

3.2 *Discussions on the differences between the two kinds of referring behavior*

Based on the results we have described by last section, two types of thinking way should be considered about and further discussed in this section.

So that the 'verbal thinking' and 'visual thinking' were brought to individually describe the ideation processes of the designers who had SI behavior and the designers who had RI behavior. The brief illustrations (see Figs. 2–3) and descriptions of the two kinds of ideation procedures are shown as follows:

The brief ideation procedure of the designers with SI behavior could be regarded as the majority of 'visual thinking'. During these two designers' (participant A & C) ideation process, a great deal of new idea sketches was made while they were referring to some references in the books or printed material. Especially, we could not find out any written ideas they had during their ideation process but the designers with RI behavior did. Therefore, the influence of 'images' has on the two designers' ideation process should be taken seriously in this study. That is to say, according to the observational results, the designers who had SI behaviors in their ideation did particularly show a major 'visual thinking' process. In terms of the above explanation, the study simply illustrates the two designers' ideation process as the top of Figure 2.

Besides, we have mentioned about the characters in advance that the designers with SI behavior tend to mark some possible stimuli in the books for referring to later. Those visual stimuli may then be thought about or partly be combined together to form a new idea. So that we made a figure to interpret the two designers' idea development process shown as the bottom of Figure 2. The figure indicates that the two designers' ideas (see I.1–I.3) were inspired by the images (see S.1–S.6) they referred to and developed by combining those different stimuli. For instance, the designer may combine the stimuli S.1 and S.2 to form the idea sketch I.1, as well as to combine the stimuli S.4, S.5 and S.6 to develop the idea sketch I.3. The designers with SI behavior extracted some specific features from those printed stimuli they found out to take as the elements of their idea sketches. The interpretation also emphasizes the influence of referring to the printed visual stimuli on the idea development during these two designers' ideation process.

However, the brief ideation procedure of the designers with RI behavior involves both 'visual thinking' and 'verbal thinking' processes. The two designers with RI behavior (participant B & D) started with generating lots of ideas from their mind and noted them down in written form in the beginning of their ideation process. Then they tried to pick up some appropriate verbal concepts to be the keywords for inputting in a search engine to retrieve the computerized stimuli on-line. During the period, we found lots of ideas have been written down on the paper before they retrieved information on-line. The continual word noting and keying actions can be regarded as the external presentation of the two designers' 'verbal thinking' process. Thus, the heavy influence on the two designers' idea development by 'words' during the period should be taken seriously. However, part of the two designers' ideation process was also influenced by 'images'. We have found that the designers with RI behavior were creating several new idea sketches while they were referring

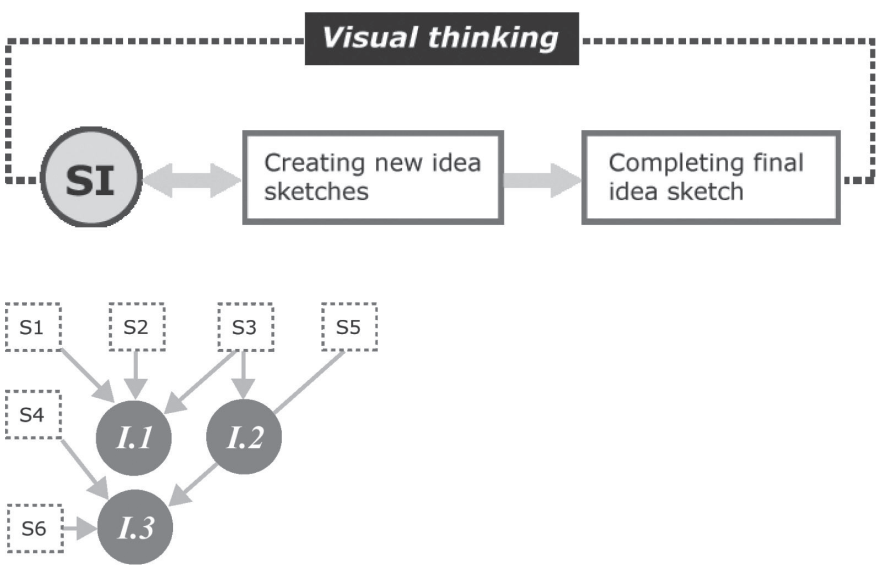


Figure 2. Ideation process (Top) and idea development (Bottom) of the SI behavior designers.

to the computerized visual stimuli. Perhaps we should say that they seemed to keep adding some elements or revising previous idea sketches when they inspired by those computerized stimuli on-line. Therefore, this part of their ideation process should be regarded as a 'visual thinking' process. In terms of the above explanation, the study simply illustrates the two designers' ideation process as the top of Figure 3.

Besides, we have mentioned about the characters in advance that the designers with RI behavior tend to revise previous idea sketches and add some elements after they referred to the computerized stimuli. Thus, those visual stimuli may be the inspirational images to the designers for adjusting previous idea. And they also added some elements to the idea sketch they have made after they found out some appropriate images from the computerized data. So that we made a figure to interpret the two designers' idea development process shown as the bottom of Figure 3. The figure indicates that the two designers' ideas (see I.1–I.4) were inspired by the images (see S.1–S.6) they referred to and especially developed by revising previous idea. That is, the designer may be influenced by the stimuli S.1 to draw the idea sketch S.1. Then the designer may get the concept by referring to the stimuli S.2 and S.3 to revise idea sketch I.1 and add some elements on it to developing the idea sketch I.2. And the designer may get the final idea sketch I.4 through continually revising and adding actions while referring to the computerized stimuli.

3.3 *Categories of designers' idea sketches with the two referring behaviors*

This section mainly addresses the result got from the sorting task of the three researchers' data analysis. The result shows that more categories of idea, as well as more idea sketches, can be found in the ideation process of the participants with SI referring behavior than the participants with RI referring behavior. We list each participant's idea sketches by Table 3 and describe the results as follows. Firstly, in terms of the number of sketches, participant A has 10 idea sketches in his ideation, participant B has 3, participant C has 7 and participant D has 4. The result indicates that designers with SI behavior (participant A & C) generated more idea sketches than designers with RI behavior (participant B & D), even though they took a different task. That is to say, designers could get more ideas while they are referring to printed stimuli, rather than referring to computerized stimuli.

Secondly, in terms of the category of each participant's idea sketch, participant A's sketches are sorted into 6 categories, participant B's sketches are sorted into 1 categories, participant C's sketches are sorted into 5 categories. The result shows that designers with SI behavior (participant A & C) had more different directions of idea during their ideation than designers with RI behavior. That is, designers could generate more directions of idea while they are referring to printed stimuli or books, rather than referring to computerized stimuli on-line.

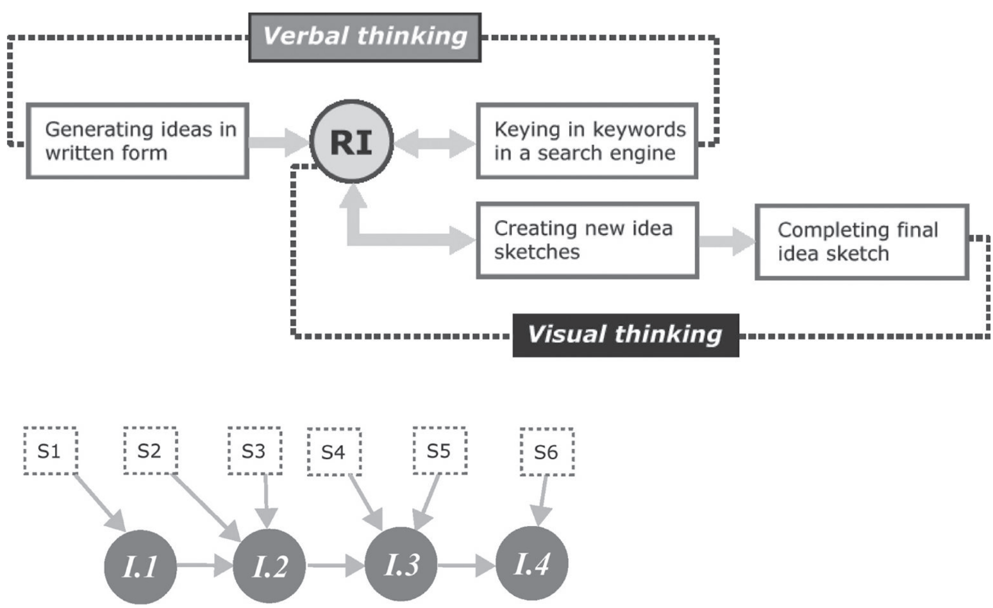
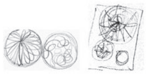
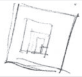
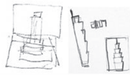
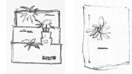
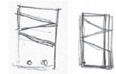
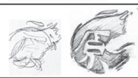


Figure 3. Ideation process (Top) and idea development (Bottom) of the RI behavior designers.

Table 3. Categories of each participant's idea sketch.

	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6
A						
B						
C						
D						

Besides, we then compare the idea sketches of the designers who had SI behavior with the designers who had RI behavior. The result indicates that the designers with SI behavior can develop more different dimensions of concept when inspired by the printed stimuli. We take participant A's idea sketches as example. He did have six different dimensions of idea for carrying out his task, such as high building, fireworks, the year … and so on. And those ideas are more different to each other. In another word, the designers with SI behavior tend to generate ideas with horizontal development during their ideation process.

However, the designers with RI behavior may develop more similar ideas when inspired by the computerized stimuli on-line. We take participant B's idea sketches as example. He only has one dimension of idea to take his task, which is firework show. Most time of his ideation process, participant B was thinking of and noting down his ideas on the paper when referring to the computerized stimuli on-line. Thus, he might get a more concrete idea to meet the design problem before he started to draw idea sketches. So we conjecture that he was trying to find out some sort of stimuli on-line for concretizing his idea in mind. It also does explain the result in section 3.2 that the designers with RI behavior would like to continually revise previous idea sketch until the final work came out. Therefore, the existed idea in his mind makes several similar idea sketches show up. By this perspective, we think that the designers with RI behavior tend to get ideas with vertical development during their ideation process.

4 CONCLUSIONS

Three main findings in this study are: first, the two different thinking ways, 'visual thinking' and 'verbal thinking', are utilized to individually describe the ideation process of the designers who referred to printed stimuli and the designers who referred to computerized stimuli on-line. Second, the designers who tend to refer to printed material could generate more different dimensions of idea sketches than the designers who tend to retrieve computerized reference on-line. In other words, the designers generate ideas mainly with horizontal development when inspiring by printed stimuli. However, the designers generate ideas with vertical development when inspiring by computerized stimuli on-line. Third, the ideation process of the designers who referred to printed reference show a stronger influence on their ideation by 'images', but the ideation process of the designers who referred to computerized data on-line shows a stronger influence on their ideation by 'words'.

Besides, the features of the designers who tend to refer to computerized reference on-line are worth to be further considered about in the future studies related to the research of design support system. However, the number of participants participated in the study is not so sufficient to confirm the final conclusions. Therefore, more participants should absolutely be concerned about in the related studies.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support for this research provided by the National Science Council under Grants No. NSC-97-2917-I-224-101. The authors also wish to thank the four graphic designers who participated in this study, Han-Zhao Lai, Zhao-Wei Chen, Zhi-Ming Zhuang and Li-Lin Zhou. As well as the three researchers participated in the analytic section, Hui-Jun Hu, Hui-Ping Lu and Shinfu Huang.

REFERENCES

Cheng, P.J. & Yen, J. 2008. Study on searching-retrieving behaviour in designers' ideation process. *Bulletin of Japanese Society for the Science of Design* 55(3): 91–98.

Cheng, P.J. & Yen, J. 2009. Graphic Designers' Computerized Retrieving Behavior in the Ideation Process. *Proceeding of HCI International 2009 CD-ROM: ISBN 978-3-642-02884-7*. San Diego, CA, USA.

Cheng, P.J. 2010. *A Study on Designers' Searching-retrieving Behavior in the Ideation Process.* Unpublished doctoral thesis. National Yunlin University of Science & Technology, Taiwan.

Dorst, K. & Cross, N. 2001. Creativity in the design process: co-evolution of problem-solution. *Design Studies* 22(5): 425–437.

Goldschmidt, G. & Smolkov, M. 2006. Variances in the impact of visual stimuli on design problem solving performance. *Design Studies* 27(5): 549–569.

Goldschmidt, G. 1994. On visual design thinking: the vis kids of architecture. *Design Studies* 15(2): 159–174.

Herbert, D. 1993. *Architectural and study drawings*. New York: John Wiley & Sons.

McCoy, M.J. & Evans, G.W. 2002. The potential role of the physical environment in fostering creativity. *Creativity Research Journal* 14(3–4): 409–426.

McGown, A., Green, G. & Rodgers, P. 1998. Visible ideas: information patterns of conceptual sketch activity. *Design Studies* 19(4): 431–453.

Nagai, Y. & Noguchi, H. 2002. How designers transform keywords into visual images. *Proceedings of the 4th Conference on Creativity and Cognition:* 18–125.

Petre, M., Sharp, H. & Johnson, J.H. 2006. Complexity through combination: an account of knitwear design. *Design studies* 27(2): 183–222.

Schön, D.A. & Wiggins, G. 1992. Kinds of seeing and their function in designing. *Design Studies* 13(2): 135–156.

Suwa, M., Gero, J. & Purcell, T. 2000. Unexpected discoveries and S-invention of design requirements: important vehicles for a design process. *Design Studies* 21(6): 539–567.

Verstijnem, I.M., Hennessey, J., Leeuwen, C., Hamel, R. & Goldschmidt, G. 1998. Sketching and creative discovery. *Design Studies* 19(4): 519–546.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Conceptual model to assist Prevention through Design (PtD)

Mauricio López-Acosta & Aarón F. Quirós-Morales
Instituto Tecnológico de Orizaba, Orizaba, Veracruz, México
Instituto Tecnológico de Sonora, Navojoa, Sonora, México

Guillermo Cortés-Robles & Alberto A. Aguilar Lasserre
Instituto Tecnológico de Orizaba, Orizaba, Veracruz, México

Stéphane Negny
Institut National Polytechnique de Toulouse, Toulouse, France

ABSTRACT: Prevention through Design (PtD) means to minimize or eliminate risks during early design stages of processes and products. This approach is one—and maybe the best—way to prevent injuries, illnesses and occupational deaths. Design issues are essential in risk management because these issues play an important role in fatal injuries and illness. Design by itself is a creative and knowledge based process. Thus a model where creativity, knowledge and tools for risk assessment are combined in an equilibrated way is highly desirable. In this paper is presented a conceptual model that combines effectively creativity, knowledge and risk management tools. The model combines the capacities of the Case-Based Reasoning (CBR) process, the MAPFRE method and the Theory of Inventive Problem Solving (TRIZ theory). The CBR has the ability to capture, index, store and reuse past experience to face new problems. The MAPFRE method is useful for risk evaluation and the TRIZ theory can guide creative effort when facing inventive design problems.

Keywords: PtD, TRIZ theory, Case-Based Reasoning, MAPFRE Method

1 INTRODUCTION

The global burden of occupational injury, illness, and death is still significant. Statistics reveal 2 million lives are lost every year around the world, 268 million of non-mortal accidents and 160 million professional diseases. Furthermore 4 percent of the world-wide gross internal product (1.251.353 million USA dollars) is lost by the cost in absences of work, treatments of diseases, incapacities, and benefits of survivors, originated by injuries, the deaths and the diseases (OMS, 2005). To solve this problem, several initiatives have been undertaken, but in the present context, the most effective way to prevent and control occupational injuries, illnesses, and fatalities is to design for minimizing hazards and risks in the early design stages (Howard, 2008).

Very few studies describe the proportions of injuries caused by poorly designed or malfunctioning equipment in the workplace. However, the main conclusion extracted from the small number that had evaluated the impact of design in risk materialization reveals that is clear that poorly designed machinery, safety measures and/or workplaces play a significant role in elevating the overall risk of occupational injuries (ASCC, 2005).

Design issues are rarely considered comprehensively in Occupational Health and Safety (OHS) research. The Australian Safety and Compensation Council (ASCC, 2005) presents a detailed consideration of the role of design issues in fatal work-related injury in Australia, with emphasis on recurrent circumstances and key industries. The analysis has shown that: (1) Similar design problems are involved in many fatal incidents. (2) Design is an important contributor to fatal injury in many industries. (3) Solutions already exist for most of the identified design problems.

Thus, design issues make an important contribution to the occurrence of fatal and serious non-fatal work-related injury in Australia. Measurement and monitoring of the extent and nature of the contribution of design issues should be able to be improved through:

1. The development of an agreed definition or set of definition.
2. Improvement of current data sources and/or introduction of new data sources.
3. The use of new research approaches.

In this paper is proposed a conceptual model that aims to manage risk in the early designs phases. This model creates a synergy among Case-Based Reasoning process, a modification of the MAPFRE methodology and the Theory of Inventive Problem Solving (TRIZ). Next section briefly introduces the Prevention through Design (PtD) approach which plays an important role in this paper.

2 PTD FOUNDATIONS

The essence of the Prevention through Design (PtD) initiative can be defined as: "*The practice of anticipating and "designing out" potential occupational safety and health hazards and risks associated with new processes, structures, equipment, or tools, and organizing work, such that it takes into consideration the construction, maintenance, decommissioning, and disposal/recycling of waste material, and recognizing the business and social benefits of doing so*" (Schulte et al., 2008). A Safety through Design definition explain that it is *"The integration of hazard analysis and risk assessment methods early in the design and engineering stages, and taking the actions necessary so that risks of injury or damage are at an accep*table *level"* Manuele (2008). PtD addresses occupational safety and health needs in the design process in order to prevent or minimize the work-related hazards and risks associated with the construction, manufacture, use, maintenance, and disposal of facilities, materials, and equipment (Howard, 2008). In this paper the term Prevention through Design (PtD) is employed as a synonymous of safety through design. According to Manuele (2008), the advantages obtained when deploying a PtD strategy are:

- Improved productivity
- Decreased operating costs
- Significant risk reduction
- Avoidance of expensive retrofitting.

Once presented the essential definitions and basic concepts, it is important to briefly introduce some advanced studies in this field. Prevention through Design (PtD) has been impelled by two approaches: PtD in Australia and PtD in USA. In 2000, two reports prepared for the Australian National Occupational Health and Safety Commission (NOHSC) underline the importance of PtD. The first one was a review of safe design literature. This report identified areas of deficiency in the understanding of legal requirements and knowledge related to safe design, and was a significant influence on the direction of future work in this area. The second report was an analysis of 225 fatality, involving machinery and fixed plant in Australia. Among those fatalities 52% were found to have at least one design factor contributing to the outcome. This report revealed that inadequate consideration at the design stage led to unsatisfactory safety outcomes (Creaser, 2008).

A similar report on design-related work injuries (NOHSC, 2004) found that:

- 37% of 210 workplace fatalities studied definitely or probably had design-related issues involved.
- In another 14%, circumstances suggested design issues were involved.
- 90% of the fatalities involving machinery and fixed plant appeared to be due in part to design issues.
- Design issues appeared to contribute to at least 30% of injuries.

Resulting from this effort, a national strategy was proposed. In this strategy five points were stated as a national priority: (1) Reduce high incidence/severity of risks. (2) Improve the capacity of business operators to manage OHS effectively. (3) Prevent occupational disease more effectively. (4) Eliminate hazards at the design stage. (5) Strengthen the capacity of the government to influence OHS outcomes. The fourth priority reflects the importance that safe design had taken in the policy context in Australia, and it has ensured that "safe design" maintains a high profile as an OHS policy issue. In 2006 the Australian Safety and Compensation Council (ASCC) published a model for safe design (Figure 1).

This model is a five stages design process that starts with the pre-design stage and concludes with a validated design or Design completion.

In USA, it was the National Institute for Occupational Safety and Health (NIOSH) the organism responsible for the creation of a sustainable national strategy for PtD. This national

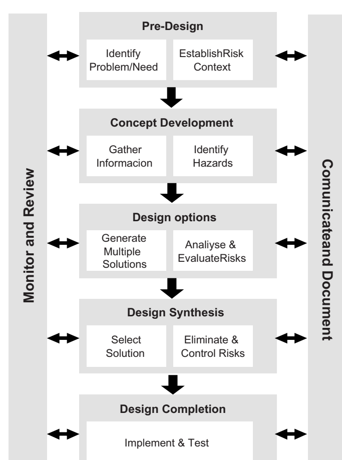


Figure 1. Safe design model (ASCC, 2006).

initiative was launched in 2007 and its objective was "*The elimination of occupational hazards and risk control to "workers "at the source" or as early as possible in the life cycle of items or workplaces*". As early means when machinery, operation or processes are in its design stage. A framework to manage this initiative was established as a national strategy to reach this objective (Figure 2).

In both initiatives there are several obstacles to surmount. Gambatese (2008) describe what are the main barriers to conducting research in the PtD field (at organizational and industrial level). Among the most relevant in the context of this article are:

- The large size, complexity, and fragmentation of the industry sectors.
- Difficulties associated with conducting occupational safety and health research.
- A lack of a standardized PtD research methodology.
- Difficulties in analyzing safety and health hazards and identifying the design as a causal factor.
- An absence of reliable data for analysis.
- Competing design priorities (e.g., cost, productivity, quality).
- An incomplete communication network for disseminating PtD knowledge.
- A lack of PtD education and training for designers.
- Difficulties in translating research to something that industry can understand.
- Another barrier not explored in literature was the creative design component of PtD. Frequently when managing risk in early design stages the process, object or service does not exists. Under this condition, risk management becomes an inventive activity where failures or risk scenarios should be invented (Lopez et al., 2008). In order to overcome these barriers a synergy among government, researches and industries should be proposed. The same strategy has been proposed in similar studies (Etzkowitz & Leydesdorff, 2000). Next section describes a conceptual model that assists design at minimal risk.

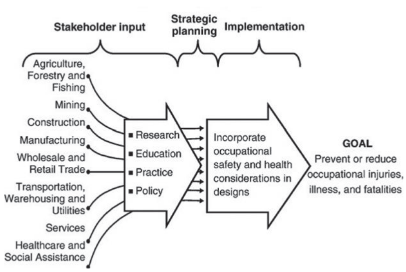


Figure 2. Framework of the PtD initiative (Schulte et al., 2008).

3 SAFE DESIGN MODEL

The identified opportunities (barriers) detected in some PtD approaches are the basis for a conceptual model to support PtD activities. This model involves different techniques and methodologies that could reduce the negative impact of the barriers listed in above section. The main components in the model are: (1) Case-Based Reasoning (CBR) as a knowledge capitalization approach that offers an excellent support for collecting, organizing, refining and distribution of knowledge. In the CBR process, problems are solved by reusing earlier experiences. In this process, a target problem is compared with a set of specific solved problems encountered in the past (called cases), to establish if one of the earlier experiences can provide a solution. If a similar case or set of cases exist, their associated solutions must be evaluated and adapted to find a new one. This approach has proved its utility to support design activities, equipment selection and also knowledge management activities among others (Avramenko et al., 2004). (2) Risk Management by incorporating the basic elements of the MAPFRE method as one of the most employed method for evaluating ergonomic issues related to workstations and work environment and finally, (3) an approach for solving inventive problems through the Theory of Inventive Problem Solving (TRIZ), TRIZ has several advantages over traditional methods, particularly when it's applied in the early design stages. The main advantages are: a) offers an important collection of knowledge extracted from several domains. This capacity produces an environment where knowledge could be used in a transversal way. As a consequence, the application of TRIZ is not restricted to a single technical domain. b) TRIZ is a more equilibrated approach that combines, in the same environment, a psychological and technical creativity dimension (Cortes et al., 2009). The integration of CBR, Risk Management and TRIZ theory produces an eight stages model. Table 1, Figure 3.

3.1 *Safe design model: Phases*

3.1.1 *Characterization*

This step generates a detailed description of the functions that must be performed by the system under design or redesign. These functions will be evaluated in the next step with the goal to identify potential risks. López et al. (2008) propose a preliminary inspection of risks through an in situ evaluation. The three steps evaluation takes in account

Table 1. Relations stages-solving approach/tool.

Stage	Solving approach/tool
1. Characterization	CBR
2. Risk identification	MAPFRE
3. Retrieval	CBR
4. Reuse/adaptation	CBR
5. Revise	CBR
6. Memorize	CBR
7. Risk Evaluation	Risk matrix
8. Risk assessment	TRIZ theory

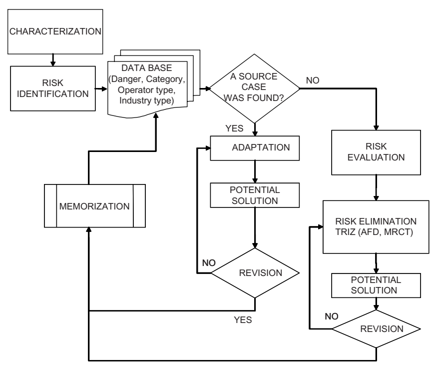


Figure 3. Shows the subjacent logic of the model.

available information regarding the situation, personnel, equipment and existing spaces of the system. In the first step it is necessary to evaluate the chart of process operations, layout distribution and other areas such as production, storehouse, among others. The goal is to obtain a wide perspective of the general situation of the company and its various production processes. In second step it is necessary to generate a classification of industrial activities. Finally in third step is needed to produce a job classification in order to characterize the risks in the productive chain. Thus, the model proposes in the characterization stage, an approach based on three elements: (1) Characterization of the company (division, group and fraction). (2) All needed activities to obtain the process output and (3) The category, class and type of operators involved in the process. These data are standard elements of the Mexican Social Security Institute which means that data are available and accessible. With this information is possible to generate the classification of cases. Those cases could be employed in a very transversal way decreasing complexity (exposed as a barrier or limitation of PtD).

3.1.2 *Risk identification*

This stage is carried out a via the MAPFRE method (Alvarez, 1994). This method offers a systematic process that allows identification to potential risks. The hierarchical classification of potential risks is the basis for retrieving a case in the case-memory. The MAPFRE method was adapted to match some Mexican safety regulations in order to facilitate standardization of cases and to increase its capacity to manage a greater number of potential risks. Finally the identification stage takes in account at least next type of potential risks:

1. Equipment Workspace layout
   a. Anthropometric Elements
   b. Other physical elements of the Workstation
2. Physical workload static posture
3. Physical workload Dynamic
4. Repetitiveness
5. Noise
6. Illumination
7. Temperatures
8. Chemicals

3.1.3 *Retrieval stage*

All the information obtained in the last stage is used to compare the present problem (target problem) with the cases stored in the memory (source problems). The identification of a source problem in the memory depends to the structure of the case-memory. The arborescence used to build the case-memory enables the search by using the KNN algorithm (K Nearest Neighbor). The structure of a case utilizes three components: (1) Company: division, group and fraction. (2) Worker: category, class and type. (3) Risk: class and type. A similarity function is deployed to classify the cases and to calculate its utility. This means that if two problems have a high level of similarity, then less effort will be needed to adapt the known solution. Inversely, high dissimilarity will increase effort to propose a new solution. Thus, a condition to launch the retrieval process states that this process will be launched only and only if the type of risk is imperatively the same. The weight associated to the case components was a subjective evaluation defined by an expert panel: (1) Characterization of the company (Division = 3%, Group = 3%, Fraction = 4%). (2) Worker (Category = 5%, Class = 5% and Type = 10%). (3) Risk (Class = 30% and Type = 40%).

Note: The criteria for considering a case needs at least 70% of similarity. This means to retrieve a case with the same type and class of risk, but detected in a different industry or a different type of operator. A case with this minimal similarity level will be considered as an acceptable one. If a source problem is retrieved, the solver will start the adaptation process, otherwise the risk evaluation should be launched.

3.1.4 *Reuse or adaptation*

Once a source problem is identified, the associated solution is evaluated to identify its pertinence and efficacy (constraint satisfaction) on the target problem. At the end of this stage, there is a potential solution for the target problem. Thus the minimal elements in a case are: [Target Problem, (Source Case, Solution (Source))]→ Solution (Target).

3.1.5 *Revise*

The solution generated is verified to check whether or not it meets the requirements of the target problem. If the test is negative, the solution should be modified until requirements are satisfied (Validate solution = positive test).

3.1.6 *Memorization*

Once the solution to the goal problem has been validated, the new experience or event can be recorded and stored in the case-memory. This process is essential to produce coordination between available knowledge and the evolution of the case-memory.

3.1.7 *Assessing the risks*

An important aspect to consider is that there are many techniques for risk management, and each sectors of production may be regulated by different systems of risk analysis, compliance with the basic premises of risk management:

a. It should be one—or a set of—criterion and also a well defined structure of analysis to compare a risk.
b. To develop control mechanisms for analyzing risk according its consequences and probabilities in a well defined context.
c. The analysis should consider the range of potential consequences and how those consequences could happen.

Kulez (2003), define some criterions to classify risk based on an event probability and consequence. This information is organized in a square matrix and a risk consequence table, useful to classify risk (Figure 4 and Table 2). In this evaluation system, the risk probability is defined as follows:

And the consequence can be measured in technical performance, cost, impact (Equipment or people) and lesion level. Five levels are proposed to accomplish this objective:

A texture code helps people to evaluate the risk level. Three elements are employed:

– High—unacceptable. Major disruption likely. Different approach required. Priority management attention required.
– Moderate—some disruption. Different approach may be required. Additional management attention may be needed.
– Low—Minimum impact. Minimum oversight needed to ensure risk remains low

3.1.8 *Elimination/reduction of risk*

The objective of this stage is to eliminate or reduce the risk prioritized in the previous phase to an acceptable level. This is a creative stage. It is thus a stage where concept generation and evaluation becomes critical in order to propose solutions that will satisfy constraints and requirements. The approach applied to fulfill this stage is the theory of inventive problem solving or TRIZ theory. According to Manuele (2005), in this stage is very

a – Remote
b – Improbable
c – Probable
d – Highly Probable
e – Security

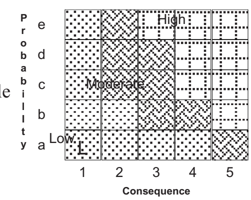


Figure 4. Evaluation matrix.

Table 2. Risk consequence levels.

Level	Technical performance	Cost	Other equipment impact / people	Lesion level
1.	Minimum / Not impact at all	Minimum / not impact at all	Non	Not applicable
2.	Acceptable some with reduction	<5%	Some impact	Medical treatment
3.	Acceptable with Significant reduction	5–7%	Moderate impact	Hospitalization, temporary disability
4	Acceptable without remaining margin	>7–10%	Mayor impact	Serious health condition, permanent disability
5.	Unacceptable	>10%	Unacceptable	Death

important to take in account the control hierarchy of risk management:

a. Eliminate hazards and risks through system design and redesign
b. Reduced risks by substituting less-hazardous methods or materials
c. Incorporated safety devices
d. Provide warning systems
e. Apply administrative controls
f. Provide Personal Protective Equipment (PPE)

When a potential solution is available, the revise stage should be launched (see 2.1.5). If the test is positive this solution is stored in the case-memory (see 2.1.6 memorization process). Otherwise, the solution is repaired and tested again until the test is positive. Once the solution to the target problem has been validated, the new experience or event can be recorded and the case-memory is updated finalizing the knowledge cycle.

3.1.9 *Advantages and Limits*

The process described above creates a strong link between knowledge and crucial problem requirements. This link is very important in a particular context: when managing risk in order to prevent accidents and illnesses in early design stages of processes, products or work procedures. The capacity to remember past solutions (errors and success) to solve new problems is provided by the CBR solving process—enabling organizational learning-. Thus the efficacy of the process is related to the content case memory—which needs too much time and effort-. This reveals on of the most important drawback of the process: what if the target problem has never been faced in the past? The element to surmount this condition is the TRIZ theory approach that can guide creative effort when solving problems never faced in the past (inventive problems).

4 CONCLUSION

This paper proposes a conceptual model to overcome some of the barriers in the PtD field. The most relevant in the context of this article are: The model could be a standard method for safe design. It allows analysis of safety and health hazards and is useful to define and solve design problems by combining the CBR and the TRIZ theory approaches. The Model is also useful to capitalize knowledge, ability that creates a collaborative knowledge network. This model can to translate PtD research in something that industry can understand, revealing the main advantages of safe design and it is potential to improve productivity.

REFERENCES

Australian Safety and Compensation Council. 2005. *Design issues in work-related serious injuries*.

Australian Safety and Compensation Council 2006. Guidance on the principles of safe design for work. URL *http://www.ascc.gov.au/ascc/HealthSafety/SafeDesign/SafeDesignforEngineering StudentsSDES.htm*

Avramenko, Y., Nyström, L. & Kraslawski, A. 2004. Selection of internals for reactive distillation column-case-based reasoning approach. *Computers & Chemical Engineering* 28(1): 37–44.

Cortes Robles, G., Negny, S. & LeLann, J.M. 2009. Case-based reasoning and TRIZ: A coupling for innovative conception in Chemical Engineering. *Chemical Engineering and Processing: Process Intensification* 48: 239–249.

Creaser, W. 2008. Prevention through Design (PtD) safe design from an Australian perspective. *Journal of Safety Research* 39: 131–134.

Etzkowitz, H. & Leydesdorff, L. 2000. The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations, *Research Policy* 29(2): 109–123.

Gambatese, J.A. 2008. Research Issues in Prevention through Design, *Journal of Safety Research* 39: 153–156.

Gonzales, B. & CarolInche, M.J. 2004. Modelo de análisis y evaluación de riesgos de accidentes en el trabajo para una empresa textil. *Industrial Data* 7: 33–41.

Howard, J. 2008. Prevention through design-introduction. *Journal of Safety Research* 39: 113.

Kulez, J. 2003. Risk integration approaches and results for the Virginia class submarine. *NASA Risk Management Colloquium* URL http://www.atc.nasa.gov/hostedEvents/rmc4/presentations/Day%201%209–403/3%20pm%20 Kulesz.pp t#5

López Acosta, M., Cortes Robles, G., Flores Ávila, L.C. & Negny, S. 2008. Integración de la administración del conocimiento y la teoría TRIZ aplicada durante el diseño de sistemas de bajo riesgo (ed.), *6º Congreso Internacional en Innovación y Desarrollo Tecnológico*; *Proc. intern. symp., Cuernavaca, 8–10 Octubre 2008*. México.

Manuele, Fred A. 2005. Risk assessment & hierarchies of control. *Professional Safety* 50: 33–39.

Manuele, Fred A. 2008. Prevention through Design (PtD): History and future. *Journal of Safety Research* 39: 127–130.

National Occupational Health and Safety Commission 2004. The role of design issues in work-related injuries in Australia 1997–2002. *National Occupational Health and Safety Commission*. URL http://www.safeworkaustralia.gov.au/ NR/rdonlyres/75 AA2 A84–57 A0–405 A-9C70-E5BB00BC1F 3D/0/SafeDesignWorkInjuries.pdf

OMS/Oficina Internacional del Trabajo 2005. Comunicado: El número de accidentes y enfermedades relacionados con el trabajo sigue aumentando. URL http://www.ilo.org/public/spanish/region/ampro/cinterfor/newsroom/hechos/safew05.htm

Schulte, P.A., Rinehart, R., Okun, A., Geraci, C.L. & Heidel D.S. 2008. National Prevention through Design (PtD) initiative. *Journal of Safety Research* 39: 115–121.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Study of prediction model of seat comfort in seat design

Kuen-Meau Chen, Yi-Hwan Shyr & Hsu Lee
Department of Industrial Design, National United University, Taiwan

ABSTRACT: This research examines the interaction between the parameters of seat size and its effect on physical comfort. Forty persons whose BMI values ranged from 18.5 to 24 participated in the experiment. An equal number of participants were males and females. Participants' heights ranged between 5% and 95% of the average heights of Taiwanese males and females. Using adjustable chairs in the experiment, this study asked participants to evaluate the degree of comfort of 27 different sizes of seats with three settings each for height, inclination angle of the seat, and height of the lumbar support. This experiment obtained reactions to different seat settings to provide designers with quick and convenient recommendations to revise seat designs. The results of the experiment suggest (1) that increasing the inclination angle of the seat back reduces discomfort due to the lack of lumbar support; (2) when the inclination angle of the seat back is almost 90 degrees, adding lumbar support significantly improves the degree of comfort; (3) when the inclination angle is around 120 degrees, seat surfaces should be lowered to increase comfort; (4) no correlation was found between the heights of chair surfaces and the heights of lumbar support. We expect to help designers assess the effects of chair size on the comfort experienced by a specific group based on the results, in addition to their own design preferences, to prevent major flaws and increase comfort for users.

Keywords: seat design, degree of comfort, interaction design

1 INTRODUCTION

If you have experienced extremely uncomfortable cinema seats with lumbar support too low, the cushion too soft, the inclination angle of the back too great or too small, you know there is room for improvement in seat design. Alternatively, restaurants seats that are so high your feet cannot touch the ground or whose backs are so hard you lose your appetite. Yes, there is innovative ability to improve seat design. This may result from designers focusing solely on appearance, or from speculation based on their own feelings, or from confusion due to others' recommendations concerning size. A user should adjust seat settings based on different seating purposes. In addition, designers cannot design every seat based on the suggested height and angle by the same group of people. If they do, their designs may be less comfortable for the public or may have major flaws.

While many factors influence the degree of seat comfort, few simple and practical tools are available for designers to assess seat design. The most important consideration when designing a non-adjustable chair is the comfort experienced by the targeted group. If these recommended guidelines are developed to help designers examine comfort in seat design, many users will benefit from better lumbar support by avoiding possible discomfort and injuries.

Many researchers have studied seat design and size recommendations, but there is little information about the interaction of various parameters related to seat size in the literature. For example, a seat with an angle close to 90 degrees between its back and surface usually causes strong discomfort to users; however, the degree of comfort can be improved if lumbar support is added. In addition, a seat that tilts back more than 10 degrees is worse for sitting in front of a low writing desk because it is difficult for users to lift their upper body when writing. This research should find other similar dynamic interactions to provide designers with simple guidelines to avoid reduced comfort when focusing on appearance. In addition, this study hope to understand the extremes of uncomfortable chairs, seat parameters' proportional effects on the degree of comfort, make an integrated prediction about comfort assessment, and further develop human factor design standards for the interaction of seat parameters.

2 LITERATURE REVIEW

People now spend long hours in a seat for work and leisure that results in musculoskeletal pain and discomfort (Ebe and Griffin, 2001). To date many researchers have studied recommended sizes for seat design for uses such as with a Visual Display Terminal (VDT) (Liao and Drury, 2000). However, the fit of seat size to body size does not necessarily determine comfort. Other factors include individual habits, seat pressure (Gyi and Porter, 1999), types of tasks conducted when sitting, the muscles of different body parts, BMI, body fat percentage, allocation of body fat (Nick et al., 1998), thermal comfort resulting from the texture of the seat surface, and even microclimatic effects between the seat surface and the individual (Diebschlag et al., 1988). Many scholars have proposed various research perspectives (Zhang et al., 1996; Yamazaki, 1992) and contributed valuable data for designers.

Because regular office chairs and car seats are concerned with different factors (Andreoni et al., 2002), car seats require designers to consider more limitations and to conduct several operations. These limitations can be applied to other non-vehicle seats, such as seat pressure (Kamijo et al., 1982; Hertzberg, 1972), inclination angle of the seat back and depth of the seat surface. However, car seats must meet certain requirements different from specifications for general usage chairs, such as type of task, possible vibrating environments, and narrower spaces.

Anthropometric data determines the design of many seats. For example, knee heights determine seat height while the length of upper legs when seated determines seat depth. However, users become used to the height of lumbar support on seats they have used for a long period, so that the presented data does not correspond to anthropometric measurements (Stevenson, 1991). According to the research of Reed et al. (1994), adjusting the height of lumbar support so the peak points of lumbar outlines are located approximately 105 to 150 mm from the H-Point greatly improves lumbar comfort. If a seat is too deep, surface pressure may center on the bottom of the upper legs just above the knees, causing discomfort (Reed et al., 1994). Eklund and Corlett (1987) compared two different types of seats and five different tasks. They found that seats or tasks that require individuals to bend the body-trunk often lead to more discomfort. Zhao and Tang (1994) pointed out that better matching seat backs to users' backs and lumbar outlines significantly improves comfort.

Relevant studies suggest that comfort and discomfort are two independent indicators for evaluation (Sauter et al., 2005; Zhang et al., 1996). Many studies only include subjective measurements of discomfort (e.g., Hsu and Wang, 2003; Jung and Choe, 1996; LeBlanc et al., 2003; Straker et al., 1997). Some studies treat comfort and discomfort as opposite directions in one dimension (e.g., Genaidy et al., 1995; Genaidy and Karwowski, 1993; Kee and Karwowski, 2001, 2003 and 2004). De Looze et al. (2003) noticed that most researchers do not separate the comfort and discomfort but combine them when including comfort and discomfort as evaluation indicators.

Therefore, this research refers to the results of Kyunga and Nussbauma in 2008 and of Gyouhyung et al. in 2008. This study believes that measuring the concept of comfort, one can identify differences in comfort more effectively than differences in discomfort.

This research focused on the interaction of seat parameters and proposed a prediction model for degrees of comfort. Why is the interaction between parameters so important? Relevant studies point out that the contents of tasks rather than the types of seats influence muscular tension (Van Dieen et al., 2001). This indicates that in addition to seat size, guidelines must consider the type of task. When carrying out different VDT tasks, the different heights of screens may cause physical discomfort (Seghers et al., 2003). Accordingly, seat comfort not only emphasizes physical size, but also involves the interaction with task environment and content. Consequently, to achieve a more comprehensive prediction, researchers should adopt a biomechanical model for calculating the lumbar pressure caused by certain postures (Kayis & Hoang, 1999) and assess the surface pressure of seats (Hertzberg, 1972; Kohara and Sugi, 1972; Chow and Odell, 1978; Kamijo et al., 1982; Bader et al., 1986; Diebschlag et al., 1988).

3 RESEARCH METHOD

3.1 *Participants*

This experiment used adjustable chairs and asked participants to evaluate the degree of comfort of 27 differently sized seats based on three parameters: seat height, inclination angle of seat back, and height of lumbar support. By doing this, this study intended to discover the most comfortable seat and the interaction between parameters. During the experiment, 27 combinations were randomly selected to prevent user memorization and bias toward any particular combination of parameters.

Forty persons participated in the experiment with an equal numbers of males and females. Table 1 presents specific requirements for participants.

1. Participants were required to have a BMI value between 18.5 and 24.
2. Participants' heights ranged between 5% and 95% of the average heights of Taiwanese males

Table 1. Basic information about participants.

	Average	Minimum	Maximum
Male height	172.5 (cm)	161 (cm)	180 (cm)
Female height	159 (cm)	150 (cm)	166 (cm)
Age	21.9 yr.	18 yr.	30 yr.
Ave. height	166.8 (cm)	150 (cm)	180 (cm)

and females, with one participant at every 5% along the scale. Female participants' heights ranged between 148 cm and 165 cm and male participants' heights ranged between 160 cm and 179 cm.

3. Participants had no history of major injuries or diseases in the lumbar area.

3.2 *Preparation before the experiment*

Before the experiment, participants were asked to stand for 10 minutes to relieve the impact of sitting for a long period. This study measured participant knee height and the distance from the knees to the back of the buttocks. Participants were asked to take off shoes and remove belts and large items from pockets. Participants were asked to try out three different combinations of the parameters to understand the range of seat sizes and to prevent biased feelings during the experiment. Participants were told to evaluate different combinations of seat parameters based on preferred posture. Finally, terminology and evaluation standards were explained to the participants.

3.3 *Combinations of the seat parameters*

The first phase of the experiment involved three parameters discussed the most in the literature. The selected sizes for the experiment were:

Factor 1 includes seat height at 33 cm, 39 cm, and 45 cm.

Factor 2 includes lumbar support as 15 cm, 27 cm, and no lumbar support.

Factor 3 includes inclination angle of seat back as 90 degrees, 105 degrees, and 120 degrees.

The three factors were combined and resulted in 27 different sizes. The 27 combinations were presented randomly for participants to evaluate. Each evaluation included four aspects of comfort: overall feeling, seat height, lumbar support, and inclination angle of seat back. Other controlled parameters included the cushion of 5 cm, the seat surface of 40 cm^2, the included angle to the horizon of 3 degrees, the 45 cm height of the seat back with a width of 40 cm. The cushion for the back was 5 cm thick.

3.4 *Questionnaire design*

Before the experiment, participants were asked about any symptoms of lower back pain and its frequency, whether they have had lumbar injuries, if they pay special attention to lumbar posture, and about any problems they encounter when seated. Previous studies suggest it is better to identify differences between seats using the concept of comfort rather than discomfort. Thus, this experiment adopted the concept of comfort to evaluate the interaction between the three parameters. During the experiment, participants were required to evaluate the combinations of the parameters in terms of four aspects of comfort: overall feeling, seat height, lumbar support, and inclination angle of the seat back. Comfort was measured using a 7-point Likert scale, that with 1 being the least comfortable and 7 the most comfortable.

Whenever a combination was scored lower than 3 points on overall feeling, the reasons why participants considered it less comfortable were documented for future analysis.

Definitions of each aspect of comfort are described as follows:

1. Overall feeling: the overall comfort experienced by users when seated.
2. Seat height: the comfort resulting from different heights from the floor to the lowest point where the ischium is located.
3. Lumbar support: the comfort resulting from different heights of the lumbar support and from the existence of lumbar support.
4. Inclination angle of seat back: the comfort resulting from different angles between seat back and seat surface.

4 RESULTS AND DISCUSSION

4.1 *The significance of factor impact and the interaction between factors*

As presented in Table 2, this multivariate factor analysis reveals that the three major factors of seat height ($F(2,944) = 4.404$, $P < 0.05$), height of lumbar support ($F(2,944) = 12.130$, $P < 0.05$), and inclination angle of seat back ($F(2,944) = 47.297$, $P < 0.05$) all have significant impact on comfort, and that there is significant interaction between the inclination angle of seat back and the height of lumbar support ($F(4,944) = 10.784$, $P < 0.05$). In addition, Fig. 1 suggests that, when the inclination angle of the seat back is 90 degrees, lumbar support of 15 cm significantly increases comfort. When the inclination angle is 120 degrees, lumbar support should be removed to increase comfort. When the inclination angle is 90 degrees, seats without

Table 2. Three-way ANOVA analysis.

Variables	SS	Degree of freedom	MS	F	P value
Seat height	11.786	2	5.893	4.404	0.012*
Height of lumbar support	32.459	2	16.230	12.130	0.000*
Inclination angle of the seat back	126.567	2	63.284	47.297	0.000*
Seat height * height of lumbar support	8.690	4	2.172	1.624	0.166
Inclination angle of the seat back * seat height	14.792	4	3.698	2.764	0.027*
Height of lumbar support * inclination angle of seat back	57.719	4	14.430	10.784	0.000*
Seat height * height of lumbar support *inclination angle of the back	5.475	8	.684	.512	0.848
Error	1228.286	918	1.338		
Sum	1485.774	944			

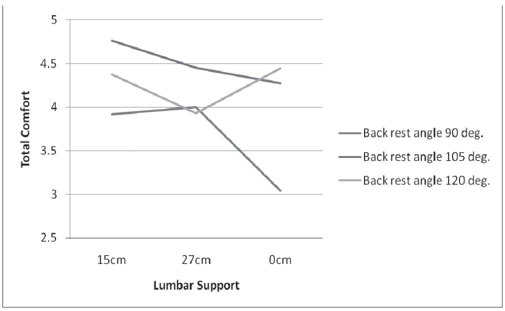


Figure 1. The interaction between lumbar support and inclination angle of seat back.

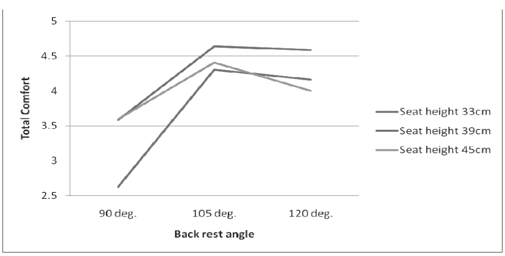


Figure 2. The interaction between seat height and inclination angle of seat back.

lumbar support have the lowest comfort score. Table 2 reveals the significant interaction between seat height and the inclination angle of the seat back ($F(4,944) = 2.764$, $P < 0.05$). Figure 2 shows that when the seat height is 33 cm and the inclination angle is 90 degrees, the seat is the least comfortable. Making the inclination angle larger than 105 degrees significantly improves comfort.

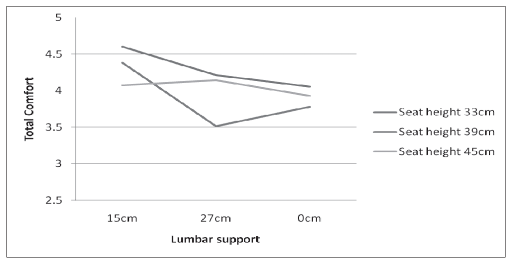


Figure 3. No interaction between lumbar support and seat height.

Another important finding is a higher (45 cm) seat causes less discomfort than a seat with a lower (33 cm) surface, even if the inclination angle is 90 degrees. A higher seat causes the least discomfort compared to seats with different heights so long as the inclination angle is larger (120 degrees). No significant difference was found between seat height and the height of lumbar support in the multivariate factor analysis ($F(4,944) = 1.624$, $P > 0.05$). Figure 3 reveals a basic trend that a seat is the most comfortable when the seat height is 39 cm and the height of lumbar support is 15 cm. No interaction was found between other factors.

4.2 *Average values of factor levels and variation significance*

Table 3 compares the average values of the three levels of seat height, lumbar support, and inclination angle of the seat back. The result of a Bonferroni multiple comparison suggested there is no significant difference between a seat height of 45 cm and 33 cm ($P > 0.05$) while the difference

Table 3. Average comfort scores of different factors on various levels.

Factors	Level	Average (SD)
Seat height	33 cm	4.02 (1.20)
	39 cm	4.26 (1.22)
	45 cm	4.03 (1.32)
Lumbar support	15 cm	4.33 (1.29)
	27 cm	4.10 (1.17)
	No	3.88 (1.25)
Inclination angle of seat back	90 degree	3.60 (1.24)
	105 degree	4.47 (1.16)
	120 degree	4.24 (1.19)
Average		4.10 (1.25)

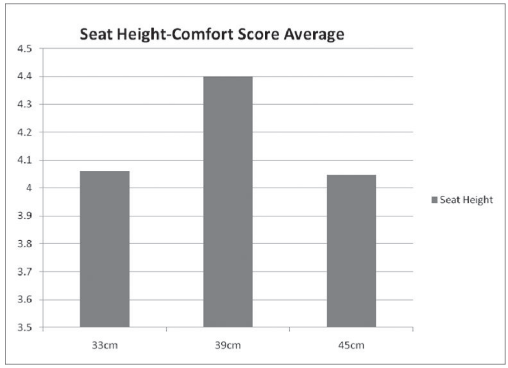


Figure 4. Average comfort scores ith respect to seat height.

Table 4. Significant differences between diffferent factors on various levels.

Factors	Comparison between levels	P value
Seat height	33 cm–39 cm	0.056
	39 cm–45 cm	0.043*
	45 cm–33 cm	1.0
Lumbar support	15 cm–27 cm	0.43*
	27 cm–no	0.49*
	no–15 cm	0.000*
Inclination angle of seat back	90 deg–105 deg	0.000*
	105 deg–120 deg	0.043*
	120 deg–90 deg	0.000*

between a seat height of 33 cm and 39 cm is on the edge significance ($0.06 > P > 0.05$). Table 4 shows the differences between other factors are all significant. A seat height of 39 cm is the most comfortable to users while heights of 33 cm and 45 cm are the least comfortable that previous recommended figures still apply in this experiment, as Fig. 4 shows.

Lumbar support with a peak height of 15 cm is the most comfortable, followed by a height of 27 cm that garners an average comfort score. Seats with no lumbar support are the least comfortable. Figure 5 shows the comfort score of such seats is about 0.2 points less than the average score. Seats with an inclination angle of 90 degrees are the least comfortable. The score is approximately 3.6 and is 0.5 points less than the average of 4.1. Seats with an angle of 105 degree have a score of 4.47 that is the highest among all factor parameters (see Fig. 6).

This study can conclude from the above data that seats that are too high or too low, have no lumbar support, and have an inclination angle of 90 degrees greatly reduce comfort.

5 CONCLUSION

The three factors adopted in this experiment are the most common and important parameters when designing seats. According to the experimental data, this study there is a significant interaction between seat height and seat back inclination angle. The result of the interaction is obvious when there is no lumbar support (Fig. 4). Seats with an inclination angle of 90 degrees but have no lumbar support receive low comfort scores. The discomfort may result from having to excessively bend the body trunk causing pressure on the abdominal cavity and the lumbar vertebra cannot retain a concave posture. By contrast, when the inclination angle is 120 degrees, seats without lumbar support are more comfortable. A possible explanation is that when the upper body remains straight, the lumbar outline forms a deeper concave; however, when the angle of the back is 120 degrees, the outline of the back is changed, and the concave of the lumbar vertebra is no longer obvious. Lumbar support of the same thickness may cause pressure to some tissues and lead to discomfort. From the perspective of lumbar support and seat back inclination angle, this study makes two observations. First, when the inclination angle is close to 90 degrees, increasing lumbar support can improve comfort. Second, when the inclination angle is larger than 120 degrees, removing lumbar support or reducing the thickness of the lumbar cushion relieves pressure on the lumbar area.

The interaction between seat height and seat back inclination angle is most apparent when the angle is 90 degrees and the height is 33 cm. These settings significantly reduce comfort, probably because of the discomfort caused by bending the body trunk. However, when the angle is larger than 120 degrees, seats of 45 cm height are less comfortable than seats of 33 cm. One can imagine the embarrassing situation when shorter individuals sit on seats that are too high and the angles too large, that lowers the center of body gravity. Consequently, based on the interaction between seat height and seat back, this study makes two observations. First, when the inclination angle must be close to 90 degrees, elevating the seat height reduces the amount of body trunk bending and increases comfort. Second, when the inclination angle is close to 120 degrees, lowering the seat height increases comfort.

6 FUTURE RESEARCH

Many variables must be controlled when designing seats, such as the thickness of cushions, the texture of seat surfaces, and the types of tasks conducted on seats. Unfortunately, researchers have not comprehensively explored the proportional impact of these factors on comfort and their interactions. In addition, to control the variables, this study simplified the experiment by asking users to evaluate seat comfort when not carrying out any tasks. However, most of the time when this study are seated, this study need to conduct specific tasks, such as using computers, drawing, dining, and writing. This study suggest that future studies include an evaluation of comfort when carrying out tasks, particularly including the parameter of desk height because desk height has a great impact on the posture of the upper body. Furthermore, other parameters of seats should be taken into consideration, such as seat depth and neck support.

The categories of participants' weight should be increased. For example, groups whose BMI values are smaller than 18.5 or larger than 24 should be recruited to exhaust the effects of comfort research on design decisions.

ACKNOWLEDGEMENT

This study thanks the Ministry of Economic Affairs for the subsidies as a part of the collaborative research project. The number of the project is 98-EC-17-A-29-S2-0150.

REFERENCES

Andreoni, G., Santambrogio, G.C., Rabuffetti, M. & Pedotti, A. 2002. Method for the analysis of posture and interface pressure of car drivers. *Applied Ergonomics* 33(6): 511–522.

Bader, D.L., Barnhill, R.L. & Ryan, T.J. 1986. Effect of externally applied skin surface forces on tissue vascularization. *Arch. Phys. Med.Rehabil*. 67(11): 807–811.

Chow, W.W. & Odell, E.I. 1978. Deformations and stresses in soft body tissues of a sitting person. *Journal of Biomechanical Engineering* 100: 79–87.

de Looze, M.P., Kuijt-Evers, L.F.M. & van Dieen, J. 2003. Sitting comfort and discomfort and the relationships with objective measures. *Ergonomics* 46(10): 985–997.

Diebschlag, W., Heidinger, F., Kuurz, B. & Heiberger, R. 1988. *Recommendation for ergonomic and physiological climatic vehicle seat design*. Society of Automotive Engineers, T Paper No. 880055.

Ebe, K. & Griffin, M.J. 2001. Factors affecting static seat cushion comfort. *Ergonomics* 44(10): 901–921.

Eklund, J.A.E. & Corlett, E.N. 1987. Evaluation of spinal loads and chair design in seated work tasks. *Clinical Biomechanics* 2: 27–33.

Genaidy, A.M., Barkawi, H. & Christensen, D. 1995. Ranking of static non-neutral postures around the joints of the upper extremity and the spine. *Ergonomics* 38(9): 1851–1858.

Genaidy, A.M. & Karwowski, W. 1993. The effects of neutral posture deviations on perceived joint discomfort ratings in sitting and standing postures. *Ergonomics* 36(7): 785–792.

Gyi, D.E. & Porter, J.M. 1999. Interface pressure and the prediction of car seat discomfort. *Applied Ergonomics* 30: 99–107.

Gyouhyung, K., Maury, A., Nussbauma, M.A. & Kari, B.R., (2008), Driver sitting comfort and discomfort (part I): Use of subjective ratings in discriminating car seats and correspondence among ratings. *International Journal of Industrial Ergonomics* 38 (2008) 516–525.

Hertzberg, H.T.E. 1972. *The human buttocks in sitting: pressures, patterns, and palliatives*. Society of Automotive Engineers, Inc., New York. Technical Paper No. 72005.

Hsu, W.-H. & Wang, M.-J. 2003. Physical discomfort among visual display terminal users in a semiconductor manufacturing company: a study of prevalence and relation to psychosocial and physical/ergonomic factors. *American Industrial Hygiene Association Journal* 64(2): 276–282.

Jung, E.S. & Choe, J. 1996. Human reach posture prediction based on psychophysical discomfort. *International Journal of Industrial Ergonomics* 18(2/3): 173–179.

Kamijo, K., Tsujimura, H., Obara, H. & Katsumata, M. 1982. *Evaluation of seating comfort*. Technical Paper No. 820761. Society of Automotive Engineers, Inc., Warrendale, PA.

Kayis, B. & Hoang, K. 1999. Static three-dimensional modeling of prolonged seated posture. *Applied Ergonomics* 30(3): 255–262.

Kee, D. & Karwowski, W. 2001. The boundaries for joint angles of isocomfort for sitting and standing males based on perceived comfort of static joint postures. *Ergonomics* 44(6): 614–648.

Kee, D. & Karwowski, W. 2003. Ranking system for evaluation of joint and joint motion stressfulness based on perceived discomforts. *Applied Ergonomics* 34(2): 167–176.

Kee, D. & Karwowski, W. 2004. Joint angles of isocomfort for female subjects based on the psychophysical scaling of static standing postures. *Ergonomics* 47(4): 427–445.

Kohara, J. & Sugi, T. 1972. *Development of biomechanical manikins for measuring seat comfort* (Technical Paper No. 720006). Society of Automotive Engineers, Inc., New York, NY.

Kyunga, G., Nussbauma, M.A. & Babski-Reevesb, K. 2008. Driver sitting comfort and discomfort (part I): Use of subjective ratings in discriminating car seats and correspondence among ratings, *International Journal of Industrial Ergonomics* 38: 516–525.

LeBlanc, J., Ducharme, M.B., Pasto, L. & Thompson, M. 2003. Response to thermal stress and personality. *Physiology & Behavior* 80(1): 69–74.

Liao, M.H. & Drury, C.G. 2000. Posture, discomfort and performance in a VDT task. *Ergonomics* 43(3): 345–359.

Nick. C., Brynley, P. & Hull. Glena E. 1998. An empirical study of preferred settings for lumbar support on adjustable office chairs. *Ergonomics* 41(4): 401–409.

Reed, M.P., Schneider, L.W. & Ricci, L.L. 1994. *Survey of auto seat design: Recommendations for improved comfort*. Report No.UMTRI-94–6. University of Michigan Transportation Research Institute, Ann Arbor, MI, USA.

Sauter, S.L., Swanson, N.G., Waters, T.R., Hales, T.R. & Dunkin-Chadwick, R. 2005. Musculoskeletal discomfort surveys used at NIOSH. In: Stanton, N., Hedge, A., Brookhuis, K., Salas, E., Hendrick, H. (Eds.), *Handbook of Human factors and Ergonomics Methods*. CRC Press, Boca Raton/London/New York/Washington, DC.

Seghers, J., Jochen, A. & Spaepen, S. 2003. Posture, muscle activity and muscle fatigue in prolonged VDT work at different screen height settings. *Ergonomics* 46(7): 714–730.

Stevenson, M.G. 1991. Ergonomic factors in the design and selection of chairs, Ergonomics and Human Environments: *Proceedings of the 27th Annual Conference of the Ergonomics Society of Australia* (Adelaide: Ergonomics Society of Australia), 47–54.

Straker, L.M., Pollock, C.M. & Mangharam, J.E. 1997. The effect of shoulder posture on performance, discomfort and muscle fatigue whilst working on a visual display unit. *International Journal of Industrial Ergonomics* 20(1): 1–10.

Van Dieen, J.H., De Looze, M.P. & Hermans, V. 2001. Effects of dynamic chairs on trunk kinematics, trunk extensor EMG and spinal shrinkage. *Ergonomics* 44(7):739–750.

Yamazaki, N. 1992. Analysis of sitting comfortably of driver's seat by contact shape. *Ergonomics* 35(5–6): 677–692.

Zhang, L., Helander, M. & Drury, C. 1996. Identifying factors of comfort and discomfort. *Human Factors* 38(3): 377–389.

Zhao, J.H. & Tang, L. 1994. An evaluation of comfort of a bus seat. *Applied Ergonomics* 25: 386–392.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Applying the Kano model into designing dinnerware for Iranian user with the approach of emotional design

Nasser Koleini Mamaghani & Parisa Izadpanah
Department of Industrial Design, School of Architecture and Environmental Design, Iran University of Science & Technology, Tehran, Iran

ABSTRACT: Living in the modern and industrial society with the high tech designed products in terms of fulfilling the physical properties has led to the satisfaction of psychological need more than ever for the happiness of the human beings. Hence, today the designers focus more attention to emotional factors. In this study, Kano model has been used in order to identify and study the emotion of the users and practicing them in to the designing of dinnerware. The attributes were studied in the two kinds of questionnaires. The first questioner contained functional and dysfunctional questions of Kano model. In the second questionnaire, the subjects were asked from their own points of view to rank these attributes within Kano categorization. According to the all finding and results, the attributes of the product were categorized based on the requirements of the users and also the amount of their participation in designing was determined.

Keywords: Kano model, self-report, emotional design, dinnerware

1 INTRODUCTION

Due to the variety of products which have a high ranking regarding the function and aesthetic factors, today completion is not an easy task in the market. In this line, many companies are trying to find a specific position for themselves. In recent years, the designers focus mainly on two points: experience and emotion. In order to improve the products, many companies state that they use emotional design (Aune, 1999, Chen et al. 2005). In this line, it seems that the functionality and aesthetics of the product are not adequate for its success. On the other hand, it has been suggested that the product quality derives its source from human requirements. The authors integrate Kano's (1982) five quality mechanisms, Maslow's (1954) hierarchy of human requirements, and Herzberg's (Herzberg et al. 1959) theory of dual factors to construct a quality mechanism questionnaire items thinking procedure chart based on the human physiological and psychological concept. There are different methods to evaluate both two basic requirements of the customers. Regarding psychological factors, one of these methods is the Kano model (Kano et al. 1984). With this in mind, the aim of this study is to investigate the Kano classification system and the self-report approach to designing a new product. Dishes of dinnerware are the most used product for all people. Therefore dinnerware was selected as a case study in current research.

1.1 *Emotional design*

Emotional design, based on the role of human emotion as an influential factor, is in a way that humans communicate and interact with objects and the world around them (Norman, 2004). Undoubtedly, emotions are integral part of human life and affect their understanding of the environment and thus they rule their thoughts (Aboulafia & Bannon, 2004). Recent scientific progresses have proved the role of human brain's activities in the human behavior to a great extent. The effects which include the emotions lead to our understanding and assessment of the surrounding world and its facts, a system to judge: good, bad, ugly, pretty, etc. (Trappl et al. 2003). These effects also include the usage and interaction with the objects. Definitely, convenient usage is an important point but today, it is not adequate just to design products with high functionality and the designers are trying to design products which can lead to the satisfaction, joy and happiness of the users while the interaction. The products which create emotions such as joy and happiness in their users can cause an easier interaction and more compatible consequences. Objects can create strong positive emotions in human such as love, happiness, joy and attachment (Norman, 2004). Meanwhile, the products which succeed in creating stronger emotions in the user have powerful and influential designs. The joy of usage is a concept which has been offered in the field of the interaction of the product and the user. This concept, in relation with the design of the functions of the product, is in a way that can cause the satisfaction of the user from the product while using it. The emotional satisfaction of the user through the improvement of the functions of the product has been set out in Kano model.

2 METHODOLOGY

In this study, in order to get the opinion of the users and the requirements of designing dinnerware, Kano model has been used. Kano model was first developed by the Japanese professor and international adviser Dr. Noriaki Kano.

This model describes the complexities of the requirements of the user and their relationship with his satisfaction (Kano, 1984). As it can be seen in Figure 1, there are three main attribute included: threshold, performance, and excitement. Threshold or basic attributes are the expected attributes or "must" of a product, and do not provide an opportunity for product differentiation. Performance attributes are those for which more is generally better, and will improve customer satisfaction. Excitement attributes are unspoken and unexpected by customers but can result in high levels of customer's satisfaction.

On this basis, this model is based on a combination of functional improvement and the emotional satisfaction of the user regarding the product.

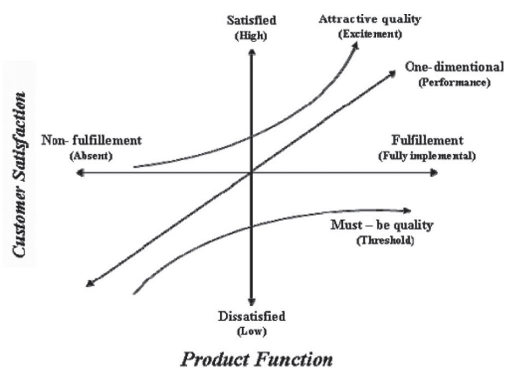


Figure 1. Kano model. This diagram enables to takes three indicated kind of needs (attribute, emotional satisfaction and functional fulfillment) which make it easier to develop a strategy to relations with consumers.

Kano states that the amount of this functional improvement and emotional satisfaction is different in the attributes of a product and the designer must consider a proper combination of these two items while designing the attributes of a product (MacDonald et al. 2006). The amount of functional improvement in the attributes of a product is an index to put it in a categorization which Kano calls quality elements and we call it Kano categorization (Lieberman, 2008). Table 1 includes a description of the attribute in the product's user that the designer should take in order to create adequate quality for each category of attribute.

2.1 *System of kano categorization*

In order to categorize the opinions of the user in respect of the attributes of the product and categorizing them in the form of Kano model, the first step is to present a questionnaire. To receive their opinions for each attribute a pair of questions are set out. The first is called functional question and the second is dysfunctional question which are shown in Table 2.

On the other hand, according to the user's answers to the questions, the relevant attributes will be categorized based on the assessment table in the form of Kano model (Ben-Rejeb et al. 2004). A sample of these pairs of questions which is related to the case study of this article is presented in Table 3.

Table 1. Definition of the Kano categorization.

Category	Description to user
Attractive	This attribute has been designed for better satisfaction of the user but dissatisfaction will not be caused if the product lacks it.
One-dimensional	In the process of designing the product, more attention has been paid to this attribute in order for better satisfaction of the user and better functionality of the product.
Must-be	The lack of this attribute will cause the dissatisfaction of the user and probably can make the product unusable.
Indifferent	The lack of this attribute will not cause dissatisfaction and its existence will not cause satisfaction.
Reverse	The existence of this attribute will cause the dissatisfaction of the user and will cause the product to be unusable.

Table 2. Kano classification table for responses to functional and dysfunctional questions (Kano, 1984).

Response to dysfunctional question						
		I like it that way	It must be that way	I am neutral	I can live with it that way	I dislike it that way
Response to Functional Question	I like it that way	Questionable	Delighting	Delighting	Delighting	One-dimensional
	It must be that way	Reverse	Indifferent	Indifferent	Indifferent	Must-be
	I am neutral	Reverse	Indifferent	Indifferent	Indifferent	Must-be
	I can live with it that way	Reverse	Indifferent	Indifferent	Indifferent	Must-be
	I dislike it that way	Reverse	Reverse	Reverse	Reverse	Questionable

Table 3. Example set of questions used in this study for each attribute.

Functional question	Dysfunctional question
How would you feel if the dinnerware has big dimensions?	How would you feel if the dinnerware has small dimensions?
Responces:	Responces:
I like it that way.	I like it that way.
It must be that way.	It must be that way.
Makes no difference.	Makes no difference.
I can live with it that way.	I can live with it that way.
I don't like it that way.	I don't like it that way.

For each attribute, respondents are first asked how they would feel the product includes the attribute (referred to as the functional question) and then how they feel if the product did not include the attribute (referred to as dysfunctional question).

Table 4. Thirteen investigated attributes of dinnerware for the first questionnaire in four groups included function, form, color and pattern.

Specification of dinnerware design	Attributes used in the design of dinnerware
Function	Type of arrangement Dimension of dishes Having side dishes
Form	Flat form Free form Form inspired from nature Harmonious form
Colors	Bright color Warm color Hot color
Pattern	Curve pattern Patterns in design Simple patterns

2.2 *Research methodology*

In this study, the emotions of the Iranian users in respect of the attributes of the dinnerware were studied by Kano model. Thus, a questionnaire was distributed among 37 university students. In this questionnaire, 13 attributes of dinnerware had been categorized in four groups of function, form, color, and pattern. These attributes are shown in Table 4. This questionnaire has been executed in two sections.

In the first section of the questionnaire, the attributes had been questioned through pairs of functional and dysfunctional questions in the form of Kano model as it was mentioned above.

In the second section, a brief description of Kano categorization and its types including exciting, one-dimensional, indifferent, must-be, and reverse was given together with an example for each one and then the subjects were asked to categorize the attributes according to their own opinion in the form of Kano model.

3 RESULTS

The results of the first questionnaire, after being prepared and steps of responding will be combined in Kano assessment and ultimately the result will be extracted (Au et al. 2006). The results of this questionnaire were presented in the Table 5. For each attribute, initially the percentage of frequency was calculated. Due to weakness in making use of mode in the findings, other formula and methods were used which have been completely explained in the part of following discussion. As it can be seen in Table 5, and accordingly, five attributes of exciting

Table 5. Results of the first questionnaire.

Product Requirement (Dishes of dinnerware)	A	O	M	I	R	Q
Type of arrangement	48.6	13.5	10.8	13.5	13.6	0
Dimensions of dishes	29.7	10.8	18.9	8.1	27	5.46
Having side dishes	37.8	13.5	10.8	24.3	10.3	3.26
Flat form	24.3	8.1	0	29.7	21.6	16.2
Free form	8.1	0	0	10.8	76	5.4
Form inspired from nature	35.1	10.8	0	10.8	40.5	2.74
Harmonious form	27	13.5	8.1	16.2	21.6	13.6
Bright color	13.5	37.8	13.5	8.1	18.9	8.14
Warm color	21.6	16.2	8.15	27	27	0
Hot color	5.4	10.8	2.7	8.1	70.3	2.74
Curve pattern	40.5	13.5	5.4	13.5	16.2	10.9
Patterns in design	10.8	0	0	21.6	62.2	5.44
Simple patterns	27	43.2	10.8	0	13.5	5.44

For each attribute the percentage of frequency, across all the subjects was calculated and presented in this table. A: attractive, O: one-dimensional, M: must-be, I: indifferent, R: reverse, and Q: questionable.

Table 6. Results of the second questionnaire.

Product Requirement (Dishes of dinnerware)	A	O	M	I	R
Type of arrangement	48.6	21.6	13.5	16.2	0
Dimension of dishes	37.8	18.9	10.8	13.5	18.9
Having side dishes	48.6	5.4	27	16.2	2.7
Flat forms	37.8	21.6	5.4	35.1	0
Free form	45.9	32.4	0	21.6	0
Form inspired from nature	37.8	18.9	2.7	35.1	5.4
Harmonious form	45.9	21.6	24.3	5.4	2.7
Bright color	32.4	24.3	16.2	27	0
Warm color	48.6	16.2	2.7	24.3	8.1
Hot color	13.5	8.1	10.8	27	40.5
Curve patterns	32.4	24.3	8.1	27	8.1
Patterns in design	29.7	27	27	16.2	0
Simple patterns	54.1	24.3	13.5	8.1	0
Modern design	67.6	18.9	10.8	2.7	0
Classic design	16.7	16.7	2.7	47.2	16.7

For each attribute the percentage of frequency, across all the subjects was calculated and presented in this table. A: attractive, O: one-dimensional, M: must-be, I: indifferent, and R: reverse.

requirement, six attributes of reverse requirement, and two attributes of one-dimensional attribute have been recognized.

The results gained from the second questionnaire were arranged as below after the completion of answers. The analysis of these data (same as the first questionnaire) has been calculated as a percentage of frequency which is shown in Table 6.

4 DISCUSSION

The results of the first questionnaire were reviewed based on the percentage for each question. Thus, the attribute of "type of arranging" the dishes and also "curve pattern" in the design of dinnerware are considered as attractive factors of designing which have the most percentages in comparison with other requirements. This shows the willingness of Iranian users for the design of the dinnerware according to their requirements and their food diet. Meanwhile, using "hot color" and "pattern in the design of dinnerware" have a high percentage of reverse attributes which shows the Iranian user's willingness for more simplicity and usage of mild colors. Also, "bright color" is a must-be requirement with 37.8% in this primary analysis. Regarding the "dimensions of dinnerware", we have two close percentages, the exciting requirement with 29.7% and the reverse requirement with 27%.

Due to this trend, we can say that the dimensions of the dinnerware are a factor variety and questionable as it has been described as a factor for satisfaction and dissatisfaction with a difference of 2%. This factor can be dependent on other conditions and factors. If the data gained from the table cannot be directly and definitely attributed to a certain requirement, the following two rules can be used (Sauerwein, 1999, Jacobs, 1999):

1. $(M + A + O) > (I + Q + R)$ Maximum of M or A or O
2. $(M + A + O) < (I + Q + R)$ Maximum of I or Q or R

where A = attractive; O = one-dimensional; M = must-be; I = indifferent; R = reverse; and Q = questionable. These rules can be applied for determining the requirement of each attribute. According to the data presented in Table 5, in respect of the "dimensions of dishes" we can say:

$$(M = 18.9 + A = 29.7 + O = 10.8) > (I = 8.1 + Q = 5.46 + R = 27)$$

As a result in this condition; A = 29.7 is the maximum of these data therefore the requirement of the product in the attribute of the dimensions is attractive. The factor of lack of meaningful difference among the data is more distinctive in the factors of warm colors. In this regard, according to the data presented in Table 5, we can use the following formula:

$$(M = 8.15 + A = 21.6 + O = 16.2) < (I = 27 + Q = 0 + R = 27)$$

Thus the requirement of the product is reverse. As the difference between reverse and indifferent is very little, it cannot be attributed to a certain category (Sauerwein et al. 1996).

In the results gained from the second questionnaire, the "form inspired from nature" and "classic design" are considered as reverse requirement of designing with a great difference. As it can be seen in Table 6, the items such as using "flat forms" and "forms inspired from the nature" are varied between the two requirements of indifferent and exciting. Also, use of pattern has percentages very close to each other. Then, we compare these percentages with the percentages gained from the first questionnaire. As it can be seen in Tables 5 and 6, the percentages gained from the two questionnaires are very close. The same also applies to other attributes. For instance, regarding the use of "pattern in design" of dinnerware and the "simple patterns", the results from the second questionnaire can be used to analyze the results of the first questionnaire. As it was discussed earlier, the attribute of using "pattern in the design" had been in the reverse requirement rank and this shows the willingness of Iranian users to use simple dinnerware without patterns. But the results of the second questionnaire indicate that the users consider the pattern in the design as an exciting or must-be requirement. But in the data which are for "simple patterns", it can be seen that from the point of view of the subjects, this attribute is an attractive requirement of design. We can conclude from these two items that Iranian users enjoy the use of "simple patterns" in their dinnerware and consider it as attractive.

5 CONCLUSION

Designing based on the emotional method will lead to the supply of products which create joy and happiness in their users. This emotion of joy and happiness can be transferred through different ways; function of the product, ease of usage, and appearance of the product. In this study, the emotions of Iranian users toward different functions of dinnerware were received and analyzed by using Kano model. This study was conducted in two parts. The first part was a questionnaire composed of functional and dysfunctional pairs of questions of Kano. The second part was the individual opinion and report questionnaire which had been prepared to complete the data gained from the first questionnaire and was influential in the analysis of the data.

The results gained from these two parts indicated the fact that Iranian users consider the functions of type of arrangement of the dishes for each person and usage of curve and free forms as attrative requirements in the design of the dinnerware. Meanwhile, the use of hot colors is considered as the reverse requirement. Iranian users are interested in using simple patterns in the dinnerware and believe that more attention must be paid to the design of patterns in the dinnerware. In general, they are more inclined to use dinnerware with modern design and this shows the change in the taste of purchase and use of classic dinnerware in Iranian users. The use of color is considered as an indifferent factor which shows the vast range of the opinions of the users for their choice of color between warm and cold colors. In this research, the emotions of the users toward the dinnerware were received and the gained results were analyzed and practiced in the design of the dinnerware.

REFERENCES

Aboulafia, A. & Bannon, L.J. 2004. Understanding affect in design: an outline conceptual framework. *Theoretical Issues in Ergonomics Science* 5(1): 4–15.

Au, E.Y.L., Li, W.Y. & Goonetilleke, R.S. 2006. The challenges of applying Kano s method to footwear design. *Conference of Biomedical Engineering (BME 2006)*.

Aune, A. 1999. *Leadership and Management for* Quality*–a Discipline called Quality Technology. (Eds Edvardson B. And Gustafsson A.),* The Nordic School of Quality Management.

Ben-Rejeb, H., Boly, V. & Morel-Guimraes, L. 2004. A new methodology based on Kano model for evaluation of a new product acceptability during the front-end phases. *Annual IEEE International Computer Software and application Conference*.

Chen, S.C., Chen, K.S. & Hsia, T.C. 2005. Promoting customer satisfactions by applying six sigma: an example from the automobile industry. *Quality Management Journal* 12(4): 21–33.

Herzberg, F., Mausner, B. & Snyderman, B. 1959. *The motivation to work*. New York: John Wiely and on, Inc.

Jacobs, R., 1999. *Evaluating Satisfaction with Media Products and Services: An Attribute Based Approach*, European Media Management Review.

Kano, N. 1984. The charming quality and the should-be quality. *Quality Control Journal* 21(5): 33–41.

Kano, N., Seraku, N. & Takahashi, F. 1984. Attractive quality and must be quality. *Quality* 14(2): 39–44.

Kano, N., Tsuji, Sh. & Seraku. 1982. *Attractive quality and obligatory quality*. In proceeding of: The annual meetings of Japanese Quality Control Society.

Lieberman, M.D. 2008. *Design Performance: The Kano Model*, OURK'S Marketing Research Review.
MacDonald, E., Backsell, M., Gonzalez, R. & Papalambors, P. 2006. The Kano Method's Imperfections and Implications in Product decision Theory. *International Design Research Symposium*.
Maslow, A.H. 1954. *Motivation and personality*. New York: Harper Row Inc.
Norman, D.A. 2004. *Emotional Design: Why We Love (or Hate) Every Things*, New York. NY: Basic Books.
Sauerwein, E. 1999. Experience with reliability and validity of the Kano-method: comparison to alternate forms of classification of product Requirements, Quality Function Deployment Institute.
Sauerwein, E., Bailom, F., Matzler, K. & Hinterhuber, H. 1996. The Kano Model: How to delight Your Customer. *International Working Seminar on Production Economics*, Austria.
Trappl, L., Petta, P. & Payr. S. (eds) 2003. *Emotions in Human and Artifacts*. Cambridge, MA: The MIT Press.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Investigation of design and implementation guidance for computerized procedures and associated automation

Chih-Wei Yang, Tzu-Chung Yenn, Tsung-Chieh Cheng & Ming-Huei Chen
Nuclear Instrumentation Division, Institute of Nuclear Energy Research, Taoyuan, Taiwan, ROC

Tung-Ming Wu
Department of Nuclear Engineering, Taiwan Power Company, Taipei City, Taiwan, ROC

Chiuhsiang Joe Lin
Department of Industrial and Systems Engineering, Chung-Yuan Christian University, Taoyuan, Taiwan, ROC

ABSTRACT: This paper provides general Human Factors Engineering (HFE) design implementation guidelines for computerized procedures. Paper-based procedures have issues that impact human reliability: (a) Complexity. Procedures often force the operator to think at several levels concurrently; (b) Data accessibility. Procedures usually require collection of plant data from different media around the control room; (c) Mental workload. Procedures often require the operator to remember items in parallel with the execution of a serial action path. Computer-based procedures provided at the workstations integrate process and equipment information and alarms with procedure steps, and are integrated with the automation features to provide efficient execution of tasks and ready access to information and controls. Text-document forms of the procedures are generated automatically and available as a consistent backup to computer-based procedures. The Advanced Boiling Water Reactor (ABWR) provides for automation of the integrated operating procedures for plant startup, shutdown and power maneuvers. A Main Control Room (MCR) with expanded functionality to allow elimination of local operations and local control rooms. This MCR Human Systems Interface (HSI) is based largely on soft video controls, electronic procedures and procedure based automation, and paperless records and work order management. This paper investigates the HFE design and implementation guidance for computerized procedures and associated automation. The contribution of this study could be contributed on considering the design of computerized based procedures and levels of procedure-based automation to the HSI designers.

Keywords: human factors engineering, computerized procedures, advanced boiling water reactor, main control room, human systems interface

1 INTRODUCTION

This paper provides general Human Factors Engineering (HFE) design implementation guidelines for Computerized Procedures (CPs). Plant procedures are instructions to guide operators in monitoring display; generating processing options; selecting an 'optimal' option, and implementing this option in the Main Control Room (MCR). (O'Hara, Higgins, Stubler & Kramer, 2000).

Historically, plant procedures have been paper-based and were not considered part of the Human-System Interface (HSI). Following the accident at Three Mile Island, the nuclear power industry recognized the importance of having technologically sound and easy-to-use procedures to handle major plant disturbances.

For emergency operations, symptom-based procedures were established that enabled operating crews to restore and maintain the plant's safety functions without having to diagnose events or the specific causes of process disturbances.

Paper-Based Procedures (PBPs) have characteristics that limit how information can be presented to the operators. These limitations include presenting information in sequential form, requiring numerous iterations through steps, and cautions or warnings that may not be applicable for all system states (Wourms & Rankin, 1994; Mampaey et al., 1988).

PBPs also impose tasks on the operator that are not directly related to controlling the plant. To make transitions between procedure steps and documents, and maintain awareness of the status of procedures that are in progress, operators must handle, arrange, scan, and read PBPs in parallel with monitoring and control tasks.

This study covers three different categories of CPs that are defined according to the functionality provided (O'Hara, Higgins, Stubler & Kramer, 2000):

- Electronic Procedures (EPs): CPs presented on a computer-driven Video Display Units (VDUs) in text or graphical form that are essentially replicas of paper based procedures but may include the ability to call up other EPs and/or include links to relevant indications or controls.
- Computer-Based Procedures (CBPs): CPs that include additional functionality such as: automatic retrieval and display of specific information and/or controls to perform a step; automatic processing of step logic and display of the results; automatic checking of prerequisites or preconditions; cautions or warnings based on current plant conditions.
- CBPs with procedure-based automation: CBPs that include the ability for the system to automatically carry out multiple procedure steps when directed to by the operator. Once a sequence of automated steps has been authorized/commanded by the operator, the PBA system can make decisions as to whether and when to carry out each succeeding step within the sequence based on plant conditions that are changing in real time.

Computer Based Procedures are being developed to support procedure management. CBPs have a range of capabilities that may support operators in controlling the plant and reduce the demands associated with PBPs. In their simplest form, CBPs show the same information via computer-driven VDUs. More advanced CBPs may include features to support managing procedures (e.g., making transitions between steps and documents, and maintaining awareness of procedures in progress), detecting and monitoring the plant's state and parameters, interpreting its status, and selecting actions and executing them.

According to O'Hara et al. (2000), CBP systems were narrowly defined to encompass computer systems that support procedure presentation and use. CBPs were characterized along the following dimensions: representation of procedure elements, procedure functionality, interface management and support, CBP hardware, backup systems for procedures, integration of CBP system with the HSI.

Several human performance issues associated with CBPs were identified, they are, methodological and criterion requirements for evaluating CBP effects, role of plant personnel in procedure management, team performance, situation awareness, response planning, operator error, level of automation of procedure functions, key effects and use of multiple CBP procedures, CBP failure in complex situations, hybrid procedure systems, and specific CBP design features.

The human performance research was organized into three categories: comparisons of CBP and PBP systems, observations of operators' use of CBPs, and comparisons of design characteristics of procedures. Several conclusions were made from comparing CBPs with PBPs: operators perform tasks more quickly, operators' overall cognitive workload is reduced, operators may make fewer errors in transitioning through procedures, operators may accept CBPs readily and find them easier to use. However, much of the human performance research had insufficient detail to evaluate its generalizability. Studies that were sufficiently documented had potential methodological weaknesses which limited their conclusiveness and generalization.

Lin, Yenn and Yang (2010a) concludes that to implement high level of automation design would be an efficient way to decrease the operators' mental demand. However, the high level of automation usually causes insufficient operating information. To provide sufficient operating information, system designers could consider rearranging the function allocation for human and automation. Then, the interfaces design will be enhanced by the results of reallocation of functions between human and automation.

Many studies (Wiener & Curry, 1980; Billings, 1991; Endsley and Kiris, 1995; Parasuraman, 2000) determined that automation does not exist as an all-or-none solution but can be based on different Types Of Automation (TOAs), each of which can be automated using different Levels Of Automation (LOAs). Generally, TOA refers to the particular human information processing functions that are considered for automation. On the other hand, LOAs inform designers of the extent to which each function should be automated.

Lin, Yenn and Yang (2010b) evaluated the effects of LOAs decisions in advanced control rooms of the modernized nuclear power plants. Experimental results and the discussion suggest that LOAs distributing the roles of option generation and generation between humans and/or computer servers which impact automated system performance significantly. Specifically, LOAs that combine option generations and planning with computer implementation (blended decision-making) produce higher situation awareness and lower mental workload, as compared to low automated control (action support and shared control) and to high LOA involving computer generation and option selection (supervisory control).

Lin, Yenn and Yang (2010c) considered the technological change that has occurred in complex systems within the past thirty years. In their study, they proposed a systematic framework to help in making an appropriate decision towards TOAs and LOAs based on a 'Skill-Rule-Knowledge' (SRK) model. For preventing the occurrences of human errors and ensuring the safety in MCR, the proposed framework can be valuable for making decisions in human-automation allocation.

2 COMPUTERIZED PROCEDURES AND ASSOCIATED AUTOMATION

2.1 *Level of automation*

Parasuraman (2000) and Parasuraman, Sheridan, and Wickens (2000) proposed a model for types and levels of human interaction with automated systems based on the Human Information Processing (HIP) model. Parasuraman (2000) addressed that automation need not be viewed as all-or-none but can be distinguished according to different types or stages of information processing, each of which can be automated at different levels.

Further, Parasuraman et al. (2000) considered that the automation can differ in type and complexity, from simply organizing the information sources, to integrating them in some summary fashion, to suggesting decision options that best match the incoming information, or even to carry out the necessary action. In Parasuraman's study (2000), types of automation were suggested to be decomposed by functional dimensions.

For example, a fault-management system in process control might sense the values of several critical system variables, integrate them in some manner, and provide a predictive display of system state to the human operator. This type of automation is qualified as information automation by Parasuraman (2000).

There are four steps in the framework for automation design. First, their model applied the four stages of HIP to classify automation types. Levels of automation were then identified based on each automation type. Primary evaluative criteria were then applied to assess human performance, mental workload, and Situation Awareness (SA). Finally, secondary evaluative criteria, such as automation reliability and costs of action outcomes, were used to confirm automation types and levels.

This model can be used as a starting point when determining the automation types and levels that should be implemented in a particular system (Parasuraman et al., 2000). This model also provides a framework for evaluating issues relevant to automation design, including human performance, mental workload, SA, automation reliability and costs.

Lin, Yenn, Jou, Yang and Cheng (2008) developed another model for automation types and levels based on a classification framework for skill, rule, and knowledge human behavior (Figure 1). In the first phase of this model, human errors associated with a target system were analyzed by the SRK model. Skill-based slips occur when situation understanding and intention are correct, but execution of an action is erroneous. In these circumstances, the so-called 'implementing automation' (Table 1) can be applied to prevent poor skill execution.

Skill-based lapses are usually due to memory failure (i.e., forgetfulness); in these circumstances, 'monitoring automation' (Table 1) can be used to assist with perceptual errors. Rule-based mistakes occur when actions are in accordance with a plan, but the plan is inadequate to achieve a desired outcome; in other words, the wrong rules were selected for an action. In these circumstances, the so-called 'selecting automation' (Table 1) can be applied to improve the development of appropriate rules.

Knowledge-based mistakes occur when decision-making is based on an inadequate understanding of a situation. Such mistakes can result from insufficient knowledge ('perceptual error') or from insufficient expertise in interpreting complex information ('decision-making error'). In these circumstances, the so-called 'planning (or generating) automation' (Table 1) can be used to assist or replace human decision-making.

In recent years there has been renewal of interest in discussions of LOAs within the context of the automation design in variety human-system interfaces. Although a large number of studies have

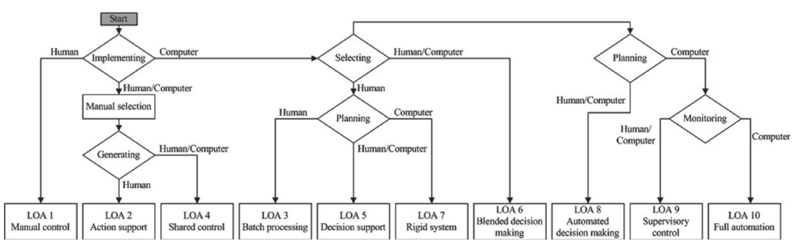


Figure 1. LOAs decisions framework (adapted from Lin et al. 2008).

Table 1. LOAs taxonomy (Endsley and Kaber, 1999).

	Roles			
Levels of automation	Monitoring	Generating	Selecting	Implementing
(1) Manual control	Human	Human	Human	Human
(2) Action support	Human/Computer	Human	Human	Human/Computer
(3) Batch processing	Human/Computer	Human	Human	Computer
(4) Shared control	Human/Computer	Human/Computer	Human	Human/Computer
(5) Decision support	Human/Computer	Human/Computer	Human	Computer
(6) Blended decision making	Human/Computer	Human/Computer	Human/Computer	Computer
(7) Rigid system	Human/Computer	Computer	Human	Computer
(8) Automated decision making	Human/Computer	Human/Computer	Computer	Computer
(9) Supervisory control	Human/Computer	Computer	Computer	Computer
(10) Full automation	Computer	Computer	Computer	Computer

been made on comparing the effects of variety LOAs, little is known about the effects of LOAs in the main control room of advanced NPPs and their impact on human factors in response to recent increases in digitalization of control rooms.

2.2 *Computerized procedures classified by LOAs*

This study is intended as an investigation of design and implementation guidance for computerized procedures and associated automation. A classification for the computerized is developed in this study.

- Electronic Procedures (EPs): The operator is required to: (a) generate a target processing order; (b) select one of the control modes and then press the withdraw/insert button to handle the selected control rods. Finally, the VDU of the Flow-Power Map illustrates the percentage power, percentage core flow, and relationship between power and core flow. In conclusion, the LOA of EPs can be assigned to LOA 2: action support.
- Computer-Based Procedures (CBPs): The operator is required to select one group of the control modes in preset and then press the withdraw/insert button to handle the selected control rods. Then, the VDU of the Flow-Power Map illustrates the percentage power, percentage core flow, and relationship between power and core flow. In conclusion, the LOA of CBPs can be assigned to LOA 6: blended decision making.
- CBPs with procedure-based automation: This mode represents automation of all functions via human override capability. The computer: (a) generates a processing strategy that considers all target variables; (b) selects and handles the selected control rods. Finally, the VDU of the Flow-Power Map displays the percentage power, percentage core flow, and relationship between power and core flow. In the automatic mode, the system is mostly automated, but the design assumes human monitoring and intervention. In conclusion, the LOA of CBPs with procedure-based automation can be assigned to LOA 9: supervisory control.

2.3 *Empirical evaluations of CPs based on personnel performance*

Plant procedures provide instructions to guide operators in monitoring, decision making, and controlling the plant. More advanced CPs may include features to support managing procedures (e.g., making transitions between steps and documents, and maintaining awareness of procedures in progress), detecting and monitoring the plant's state and parameters, interpreting its status, and selecting actions and executing them.

The human performance research within computerized procedures was organized into three categories: comparisons of CBP and PBP systems, observations of operators' use of CBPs, and comparisons of design characteristics of procedures. Several conclusions were made from comparing CBPs with PBPs: operators perform tasks more quickly; operators' overall cognitive workload is reduced; operators may make fewer errors in transitioning through procedures; operators may accept CBPs readily and find them easier to use.

The operators' impact on the plant's functions, processes, systems, and components is mediated by a causal chain from their physiological and cognitive processes, to task performance, and ultimately, to the plant's performance through the operators' manipulation of the HSI. HSI design, including its procedures, affects the plant's performance through personnel tasks that support operations.

Lin, Yenn and Yang (2010a) found that human operators using an automatic control mode (LOA 9: supervisory control) had significantly lower mental demand than those using a semiautomatic mode

(LOA 2: action support). The measurement results for secondary task demonstrate that participants were able to respond to red dots on a white board more quickly under the automatic mode than when under semi-automatic, indicating that the mental demand placed on participants under automatic mode (LOA 9: supervisory control) was less than that under semiautomatic mode (LOA 2: action support). They therefore conclude that mental demand under a high LOA was less than that under a low LOA.

Unlike the studies by Kaber and Endsley, which manipulated many LOAs, this study compared only two LOAs due to the limitation of the real-world simulation platform. However, these experimental results show that operator workload was significantly decreased by a higher LOA (LOA 9: supervisory control) than lower LOA (LOA 2: action support) within various systems.

Early studies that measured SA with LOAs obtained different experimental results. Endsley and Kaber (1999) found that LOA was not a significant factor influencing the percentage of correct responses to Situation Awareness Global Assessment Technique (SAGAT) queries involving operator awareness of target features (level 1 SA questions) and operator projections of the near future (level 3 SA questions).

Endsley and Kaber (1999) found a high SA (level 2: understanding the impacts upon specific goals and objectives) in high LOAs (LOA 6, 8, 9, 10), as compared with the OOTL notion that a high degree of automation reduces SA. Kaber, Onal and Endsley (2000) found that only level 3 SA was influenced by LOAs; that is, level 3 SA in LOA 2 was significantly higher than that in LOAs 9 and 10.

Kaber and Endsley (2004) again identified that the LOA had a significant effect on the average percentage of correct responses to task comprehension queries (level 2 SA). Their experimental results identified a low SA (level 2 SA: understanding target characteristics relative to task goals) in LOAs 3 and 9. Based on these studies, a complex association exists between LOA and SA. Whether LOA affects SA is likely dependent on more factors than expected.

3 DISCUSSION

Experimental results indicate that current automation design at the NPP was imperfect, and could be improved; this finding is similar to that obtained by Huang et al. (2006, 2007). They evaluated human operator performance using an alarm-reset system under automatic and manual modes, which is a subsystem in the same platform as PCTRAN in the fully digital control room. Although Huang et al. (2006, 2007) found that different automation modes influence human performance. They did not provide detailed analyses using LOA taxonomy. They determined that participants perform well in manual selection tasks.

The major cause is that automation of alarm information is still insufficient for participants to understand and operate alarm states. The experimental results acquired by this study about using an alarm-reset system under automatic and manual mode are consistent with those obtained by Huang et al. (2006, 2007).

This study concludes that to implement high level of automation design would be an efficient way to decrease the operators' mental demand. However, the high level of automation usually causes insufficient operating information. To provide sufficient operating information, system designers could consider rearranging the function allocation for human and automation. Then, the interfaces design will be enhanced by the results of reallocation of functions between human and automation.

4 CONCLUSIONS

As noted above, several human performance issues associated with CBPs were identified. They represent topics for which research is necessary before developing guidance. From a regulatory review perspective, O'Hara, Higgins, Stubler and Kramer (2000) pointed out many of them can be dealt with on a case-by-case basis during the design process review. Briefly, the issues included the following:

1. Methodological and Criterion Requirements for Evaluating CBP Effects: Most of the studies reviewed had methodological weaknesses which limited their conclusiveness and generalizability. This issue addresses the need to evaluate CBPs and their effects on crew performance comprehensively, to better understand them under a wide range of scenarios and complex situations, using varied measures of personnel and system performance.
2. Role of Plant Personnel in Procedure Management: This issue addresses the need to determine how to design and review CBP systems (a) to allow operators to maintain an independent perspective and to recognize the procedure's contribution to higher-level safety goals, (b) to automate distracting and lower-level error-prone tasks, and (c) to monitor the crew's performance, especially when the crew and CBPs disagree.
3. Team Performance: This issue addresses the requirement to explore the effect of CBPs on crew member's roles, teamwork, and communication. How CBPs can be designed to effectively promote both is considered as well.

4. Situation Awareness, Response Planning, and Operator Error: This issue addresses the need to assess the effect of CBPs on situation awareness including: procedure management, such as status of procedure steps, how procedures are structured, and the current location within a procedure or between a set of procedures; the appropriateness of procedures for achieving high-level procedure goals, and the plant's status.
5. Level of Automation of Procedure Functions: This issue addresses the need to evaluate the tradeoffs between automating procedure functions, e.g., the analysis of procedure step logic, and the operator's involvement, independence, and supervisory control.
6. Keyhole Effects and Use of Multiple CBP Procedures: This issue concerns the requirement to evaluate the significance of the keyhole effect in situations where operators are required to be in multiple procedures and must access information in parallel.
7. CBP Failure in Complex Situations: This issue involves the need to evaluate operator's management of the transition from CBPs to PBPs and back to CBPs under complex conditions, e.g., in a situation where operators are deep into the procedures, multiple procedures are open, many steps are completed, many are continuously applicable, and time and parameter steps are being monitored by the CBPs.
8. Hybrid Procedure Systems: This issue addresses the need to evaluate any differential effects of having all plant procedures presented in a CBP system versus a hybrid system, e.g., EOPs presented using CBPs and all other procedures are paper-based.
9. Specific CBP Design Features: This issue addresses the need to evaluate the relative effects of specific CBP design features on performance.

This study put emphasis on the issue (5) and investigated several present studies to describe findings from them. Besides issue (5), all the other issues also needs to be deeply investigated in the future study.

REFERENCES

Billings, C.E. 1991. *Human-centered aircraft automation: a concept and guidelines* (NASA Tech. Memo. No. 103885), Moffet Field, CA: NASA-Ames Research Center.

Endsley, M.R. & Kiris, E.O. 1995. The out-of-the-loop performance problem and level of control in automation. *Human Factors* 37(2): 381–394.

Endsley, M.R. & Kaber, D.B. 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics* 42: 462–492.

Huang, F.-H., Hwang, S.-L., Yenn, T.-C., Yu, Y.-C., Hsu, C.-C. & Huang, H.-W. 2006. Evaluation and comparison of alarm reset modes in advanced control room of nuclear power plants. *Safety Science* 44: 935–946.

Huang, F.-H., Lee, Y.-L., Hwang, S.-L., Yenn, T.-C., Yu, Y.-C., Hsu, C.-C. & Huang, H.-W. 2007. Experimental evaluation of human–system interaction on alarm design. *Nuclear Engineering and Design* 237(3): 308–315.

Kaber, D.B., & Endsley, M.R. 2004. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science* 5(2): 113–153.

Kaber, D.B., Onal, E. & Endsley, M.R. 2000. Design of automation for telerobots and the effect on performance, operator situation awareness and subjective workload. *Human Factors and Ergonomics in Manufacturing* 10(4): 409–430.

Lin, C.J., Yenn, T.C. & Yang, C.W. 2010a. Evaluation of operators' performance for automation design in the fully digital control room of nuclear power plants. *Human Factors and Ergonomics in Manufacturing & Service Industries* 20(1): 10–23.

Lin, C.J., Yenn, T.C. & Yang, C.W. 2010b. Automation design in advanced control rooms of the modernized nuclear power plants. *Safety Science* 48(1): 63–71.

Lin, C.J., Yenn, T.C. & Yang, C.W. 2010c. Optimizing human-system interface automation design based on a skill-rule-knowledge framework. *Nuclear Engineering & Design* 240(7): 1897–1905.

Lin, C.J., Yenn, T.-C., Jou, Y.-T., Yang, C.-W. & Cheng, L.-Y. 2008. A model for ergonomic automation design of digitalized human-system interface in nuclear power plants. In *2nd International Conference on Applied Ergonomics*, July 14–17, 2008, Las Vegas, USA.

Mampaey, L., Moeyaert, P., Bastenaire, F., Casier, F. & Chi, N. 1988. Operator advisor: An expert system to reduce human error in emergency response. In *Proceedings of the International ENS/ANS Conference on Thermal Reactor Safety, NUCSAFE 88.* Avignon, France: Societe Francaise d'Energie Nucleaire.

O'Hara, J.M., Higgins, J.C., Stubler, W.F. & Kramer, J. 2000. *Computer-based procedure systems: Technical basis and human factors review guidance*. (NUREG/CR-6634) Washington, DC: U.S. Nuclear Regulatory Commission.

Parasuraman, R. 2000. Design automation for human use: empirical studies and quantitative models. *Ergonomics* 43(7): 931–951.

Parasuraman, R., Sheridan, T.B. & Wickens, C.D. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans* 30(3): 286–297.

Wiener, E.L. & Curry, R.E. 1980. *Flight deck automation: Promises and problems. Ergonomics* 23: 995–1012.

Wourms, D. & Rankin, W. 1994. *Computer-based procedures* (Tech Report CSERIADC-RA-94-002). Wright-Patterson Air Force Base, OH: Crew Systems Ergonomics Information Analysis Center.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Inventor motivations to go from idea to innovation?

Gerth Öhman & Karl W. Sandberg
Mid Sweden University, Sundsvall, Sweden

ABSTRACT: The objective of present paper is to describe the motivations that underlie the choices made by inventors in the development of their technologies. These classifications can help managers of innovation programs direct sparse resources to inventors with the greatest probability of commercial success. The categorisation of motivation to go from idea to innovation and commercialisation of new technologies is presented for inventors for the small business in rural area of Sweden. The main results showed least eight different inventor motivations.

Keywords: inventor, motivation, success factors

1 INTRODUCTION

Many writings on innovation emphasise the importance of flexibility and the need to think broadly. The more robust research concerning product innovations has centred on success or failure by absorbed outside the development process itself (Cooper & Kleinschmidt 1990, Maidique & Zirger 1984).

Successful revolutionary innovation is rare. In any given arena, it happens only a few times per decade. The course of diffusion of innovation has been well studied, but most radical innovations fail. Those that succeed can take decades before they are successful. After many studies into new product and service performance, almost every product and service developer should be able to list off the 10 or 15 critical success factors that make the difference between winning and losing.

Most new-product development is innovative, but at a very small growth stage and costs are slimmed (Norman 2010).

We make still the same mistakes. The success factors are invisible and not found in typical business practices. Studies make known that the art of product development has not improved all that much and many products enter the development phase missing clear description (Cooper 1998, Cooper 1995, Montoya-Weiss & Calantone 1994). We have, for instance, learned that while some small businesses may innovate far more effectively than large ones, only a fraction of small firms are in fact innovative (Livesay et al. Brown 1996).

Similarly, while small business may power the engine of job creation and economic growth, many small business entrepreneurs do not want their companies to grow at all, and few aspire to large-scale expansion (Davidsson 1989). The key to understanding these varied attitudes lies in the recognition that attitudes towards growth reflect a range of motivations, which in turn root in definitions of success, the stars by which the captains of small business steer their craft.

What are these critical success factors that are so noticeably absent in most businesses' new product projects? Research has uncovered two types or classes of success factors. The first deals with doing the right projects; the second with doing projects right. The right projects are captured by a number of external or environmental success factors, over which the project team has little control (Montoya-Weiss & Calantone 1994). These include characteristics of the new product's market, technologies, and competitive situation, along with the ability to leverage internal competencies. While not within the control of the project team, these are nonetheless useful factors to consider when selecting and prioritizing projects. These success factors emphasize doing projects right, and focus on process factors or action items which are things the project team does and they are the invisible ones. But these actions are controllable and discretionary, so they are indeed seen from time to time.

This article is a continuation of previous studies of the science of innovation processes and success factors in a small business (Sandberg & Öhman 2009). We attempt a focus on what factors motivate inventors to go from idea to innovation in small firms in rural areas.

2 THEORETICAL FRAMEWORK

In the late twentieth century, innovation, the process of transforming an idea into a new product or service, appears to be playing a more vital role than ever before. Innovation and those who carry it out inventors among them has thus become the subject of extensive research aimed at achieving greater understanding of technical entrepreneurship (Livesay et al. 1996).

Major shifts in industrial processes have typically been the result of radical innovations by a small business and/or inventor. The inventor and innovator who bring inventions to the marketplace are essential in the development of society. Some research has shown that inventive and entrepreneurial activity can be integrated, and the inventor plays an important role in our economy. To realise a financial gain from inventor, the following serious problems are typically encountered; a lack of information about the innovation process; the difficulty in obtaining legitimate, unbiased analysis of the product or idea which could assist in a decision to commercialise the product; and the financial community's uncertainty to fund inventors. Without outside funding, inventors are often forced to act as entrepreneurs and market their products without professional assistance (Bowman-Upton et al. 1989).

The inventors will invent, for that is what inventors do. The technology will come first, the products second, and then the needs will slowly appear, as new applications become luxuries, then needs, and finally, essentials. Once a product direction has been established, research with customers can enhance and improve it (Norman 2010).

In practice, a positive attitude will drive an inventor toward the market; a 'hostile' attitude will prevent the marketing of the invention by the inventor him/herself; and indifference will provide no impetus, nor will it present an insurmountable barrier to commercialization.

The dramatic increase in numbers of innovation centres, incubators, small business development centres, and other government sponsored programs highlights the awareness of and importance of new business enterprise development. Critical to the evolution of these programs is an accurate evaluation of their efficacy in nurturing innovations. Efforts to measure the impact and/or effectiveness of innovation centres, incubators, entrepreneurship curricula, and various government programs have served a vital role in the fine-tuning of these interventions (Bowman-Upton et al. 1989).

2.1 *Innovation and digital technology*

Consider the cycle. First comes a new technology; perhaps it is a new idea, or perhaps an old idea that has finally reached a commercially viable state where inventors can consider it. Multiple-touch interaction with displays took roughly two decades to move from the research laboratory to its appearance in everyday products, and even so, it is not yet common outside of a few limited product categories.

We would prefer to believe conceptual breakthroughs occur because of a detailed consideration of human needs, especially fundamental but unspoken hidden needs so beloved by the design research community; the fact is that it simply doesn't happen (Norman 2010). New conceptual breakthroughs are invariably driven by the development of new technologies. If we examine the business impact of innovation, we find that the most frequent gains come from the small, incremental innovations-changes that lower costs, add some simple features, and smooth out the rough edges of a product. Most innovations are small and relatively simple (Norman 2010).

3 INVENTOR MOTIVATIONS

The purpose of this paper is to describe categories of human factors that try to explain the technical innovation process using data from the small business and independent inventors in rural area of northern Sweden. We should identify and understand motivation factors behind innovators drivers to go from idea to innovation.

The motivational categories derive from attempts to formalise understanding of non-standard approaches to innovation and commercialisation planning and are meant to capture the most common motivational problems consultants encounter in dealing with inventors involved in the commercialization of technology.

3.1 *Finding inventor motivations*

This part present the qualitative results from an interview study of ten successfully inventors in small business in rural areas to recognize their motivations to go from idea to innovation. Rather than seeking to capture best practices, we seek to set up supposed categories that can serve to identify patterns of failure and predict poor innovation performance. It is possible to identify following types of inventor motivations based on interview:

- The way of life of the corporate executive as a primary motivation for pursuing the commercialisation of technology.
- Lies in the innovation process activity in itself.
- Use the innovation process as an instrument for social reform.

- Public recognition of unique skills and knowledge through the innovation process.
- Held within the borders of a business relationship.
- Stay away from commitment of personal effort and resources to the innovation process.
- Enter the innovation process primarily to keep away from an unlikeable option.
- Main reward from the commercialisation of technology.
- Create a family business as an ethical extension of private life.

3.2 *Implications*

The analysis of underlying inventor motivations presented in this paper has implications both for the business and for the larger arena of technology transfer.

One of the primary strengths of this categorisation as a management tool lays in the fact that more than one nominal category may quite frequently be used to describe a particular inventor. In such cases, the key to innovation and commercialisation preparation is found in an understanding of motivations connecting of the motivational categories.

Even when not directly related to inventors, these findings help explain the lack of enthusiasm of many businesses to interact with inventors, and the lack of success meet by many inventors who have attempted to initiate a relationship themselves.

Further, as some researchers have specifically argued, public policy efforts to make possible technology transfer and innovation need to take variety into account, especially diversity of motivation, and then adapt programs to get used to those realities (Cooper et al. 1989, Dubini 1988).

In addition, evaluation findings, such a typology might contribute to a more effective integration of small business and independent inventors to create new and improved products and services. Micro businesses and especially independent inventors are in part underutilised, because of common misunderstanding about their attitudes, values, and motivations.

Innovation can be obstructed by such knowledge barriers because it requires the effective integration of research, development, manufacturing, and marketing.

Motivational factors can play a critical role in innovation and commercialisation planning and ultimately in market success, a point that can be made clearly through brief commentary on motivational factors as they appear in present study of inventors in rural area of Sweden.

The value of exploring the particular and personal in inventors' motivations is found in the recognition that while there are commonly recognizable approaches to technology development and the innovation process, there are no universally accepted formulas for success, but inventors and entrepreneurs must make choices at each step in the innovation process: choices of product and service definition, technical development, market strategy, and organisation.

Application-oriented, descriptive research can serve as a bridge between empirical testing and the pressing issues confronting today's manager. Hopefully, results such as presented here will entice others to overcome our shortcomings and provide even more concrete rationale for the theories and methods of new product/service development. A contribution to increased success comes from increased managerial and other employee involvement. While these general findings seem quite obvious, they should help put to rest the notion that a few specialists, working in isolation, can produce significant numbers of successful service innovations.

4 CONCLUSIONS

The value of exploring the particular and personal in inventors' motivations is found in the recognition that while there are commonly recognizable approaches to technology development and the innovation process, there are no universally accepted formulas for success, but inventors and entrepreneurs must make choices at each step in the innovation process: choices of product and service definition, technical development, market strategy, and organisation.

REFERENCES

Bowman-Upton, N., Seaman, S.L. & Donald, L. Sexton, D.L. 1989. Innovation evaluation programs: do they help the inventors? *Journal of Small Business Management* 27(1): 23–30.

Brown, M.A. & Major, C.H. 1990. Technology transfer strategies of DOE's conservation programs. *Journal of Technology Transfer* 15: 33–40.

Cooper, A.C., Woo, C.Y. & Dunkleberg, W.C. 1989. Entrepreneurship and the initial size of firms. *Journal of Business Venturing* 4: 317–332.

Cooper, R.G. & Kleinschmidt, E.J. 1990. New product success factors: A comparison of kills versus successes and failures. *R&D Management* 20(1): 1–7.

Cooper, R.G. & Kleinschmidt, E.J. 1995. Benchmarking the firm's critical success factors in new product development. *Journal of Product Innovation Management* 12(5): 374–391.

Cooper, R.G. 1998. Benchmarking new product performance: results of the best practices study. *European Management Journal* 16(1): 1–17.

Davidsson, P. 1989. Entrepreneurship and after? A study of growth willingness in small firms. *Journal of Business Venturing* 4: 211–226.
Dubini, P. 1988. The influence of motivations on business startups: some hints for public policies. *Journal of Business Venturing* 4: 11–26.
Livesay, H.C., Lux, D.S. & Brown, M.A. 1996. Human factors and the innovation process. *Technovation* 16(4): 173–186.
Maidique, M.A. & Zirger, B.J. 1984. A study of success and failure in product innovation: the case of the US electronics industry. *IEEE Transactions on Engineering Management* 4: 192–203.
Montoya-Weiss, M.M. & Calantone, R.J. 1994. Determinants of new product performance: a review and meta analysis. *Journal of Product Innovation Management* 11(5) 397–417.
Norman, D. 2010. Technology First, Needs Last: The research-product gulf. http://interactions.acm.org/content/?p = 1343
Sandberg, K.W. & Öhman, G. 2009. The Science Inside Innovation Process Factors. *The 3rd Symposium on the Entrepreneurship-Innovation-Marketing Interface and 2nd BIEM-Symposium*, June 11–12, 2009, Cottbus, Germany.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Study of ergonomics improvement on disaster prevention for ancient buildings: An example of Wufeng building clusters

Pei-Chi Su
Graduate School of Design Doctoral Program, National Yunlin University of Science and Technology, Yunlin, Taiwan

Shang-Chia Chiou
Department of Architecture and Interior Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

ABSTRACT: This study aims at environmental issues in consideration of ergonomics under survey in actual space when encountering impacts on environment and relief in use and disaster prevention of ancient buildings, focusing on all kinds of conditions of hardware inside ancient buildings of cultural heritage. Evaluation is made on standards of laws and regulations of facilities inside the buildings in comparison with those in Japan. On one hand, we try to understand how hardware facilities of buildings in Taiwan are used through site investigation and tests. On the other hand, universal issues of use of ancient buildings in Taiwan are proposed in the hope to find possible changes and reasonable modification from site investigation to give major suggestions on hardware facilities for the reference of relevant authorities in future disaster relief of ancient buildings on spatial dimensions.

Keywords: traditional, buildings, ergonomics

1 FOREWORD AND PURPOSES

In restoration of historic buildings and historic spots, emphasis is on the original appearance in the purpose of preserving the cultural heritage of a country and promotion of preservation of cultural historic spots. Number of historic spots and buildings requiring restoration in Taiwan is increasing. Each year, Council of Culture Affairs proposed subsidy on restoration of original appearance of buildings from each city/county. The main goal is to preserve ancient buildings and keep the values of cultural heritage. Nevertheless, current restoration of ancient buildings in Taiwan focuses on recovery, without consideration of assessment of environment use. A lot of ancient buildings are not access free and make disaster prevention difficult. It is indeed unlikely to prevent disasters. In real operation, and no way to give real protection to the role in consideration of disaster prevention, the government is required to evaluate ancient buildings relief issues and improvement on the environment.

Under preceding considerations, this study aims at evaluation of use and necessity of modification of hardware facilities of ancient buildings in Taiwan in light of ergonomics and exploration of people's satisfaction on use of hardware facilities. Other than analysis of current conditions, the findings shall help relevant authorities on reviewing hardware facilities of ancient buildings.

2 STUDY METHODS

2.1 *Target*

This study aims at ancient building cluster of the Lin's in Wufeng in operative survey with site investigation for problem analysis and discussion. The target is currently the largest designated ancient building cluster in Taiwan. Analysis and explanations are given after investigation and research results on operation of disaster prevention and relief to understand the work of prevention of disasters of ancient buildings and subsidy of modern relief or prevention of disasters.

2.2 *Analysis methods*

Analysis of materials after investigation helps understand the association of use and dimensions of hardware facilities of ancient buildings in current space and further explore differences of dimensions of hardware facilities in the buildings. Discussions are made on the obstacles on the ground and explanations are given on areas that easily make relief difficult.

3 DETAILS

In light of ground space of ancient buildings that affect relief, the following illustrates the dimensions and potential subsidy in the example of Dahua Hall:

3.1 *Ground stiles*

The general height of stile height in the Lin's in Wufeng is 18 cm. When entering from outside, it takes 2–3 stepping on the stiles. Threshold is also 6 cm tall. This causes great inconvenience when carrying stretchers and may even hinder relief. Yet, such stiles have been in historic buildings. In restoration, one shall add auxiliary slope plate on stile to help relief. This is the first priority in modification.

3.2 *Ground height*

Regulations on ground height in restoration of historic spots are different from those on regular buildings. The former are in stacking construction, which is different from modern construction. Modification is required and original appearance of ancient buildings in relief.

3.3 *Threshold*

Item	Dimension	Modification
Height	45 cm	Making movable threshold, offering slope design for access free environment

The biggest problem in disaster relief of historic buildings is the height of threshold. That of most traditional buildings is over 45 cm. This causes great inconvenience when moving sickbed to ambulances. Patients can only be lifted manually, resulting in slow relief and safety concern (in seeking refuge in emergencies).

3.4 *Sidewalks, stairs*

Item	Dimension	Satisfactory dimensions
Sidewalk width	1.8 meters max.	Minimizing problems of height among sidewalks
Stair brick height (level height)	20 cm	17.24 cm*
Stair brick width	90 cm	132.84 cm*
Handle height	n/a	89.44 cm*
Number of steps	17	Only few have platforms

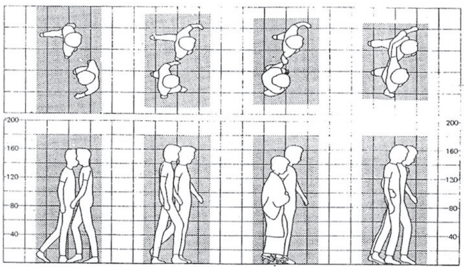
a. Sidewalks

Width: under laws and regulations of construction in Taiwan, sidewalk width shall be no less than 1.1 meters. Where there are rooms on two sides, the width shall exceed 1.6 meters. For ancient buildings, dimensional modification was based on then construction space without consideration of width issue. Space is insufficient when visitor crisscross. There is no sufficient space for normal activities. Spatial dimensions of each building under survey are different. The widest is around 1.8 meters. It is the corridor with larger space and meets regulations on walking.

Table 1. Explanations of internal activities in the space.

	Main building Dahua Hall	Front courtyard	Display of ancient objects	Drama platform	Viewing planning on the 2nd floor	Central courtyard
Sedentary	Rest			Rest	Rest	Rest (sitting high, platform surface height)
	Meeting					Dining
Standing	Tour guide exploration	Explanation environment	Tour guide explanation	Explanation environment	Appreciation	Chatting
			Visiting ancient objects	Appreciation	Interaction	Playing games, activities
Bending down	Unloading/ loading goods (bending down, arm diameter)	Unloading/ loading goods (bending down, arm diameter)				
	Visiting					

Inclination: the best inclination as Architectural Society of Japan is 1:12 or 4.76°. Architecture and Building Research Institute, Ministry of Interior suggests gentle gradient as much as possible. In general, sidewalk gradient is recommended at less than 7°. Slippery proof is recommended to avoid accidents. AIA and DIN suggest gradient below 8° for safety concern. Most of historic building cluster has no gradient, which is inconvenient to the disabled and makes relief difficult.

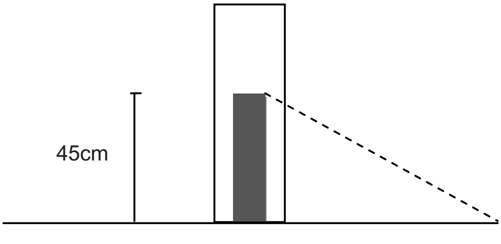


A. Walk together
B. Walking side by side
C. Holding hands walking
D. Hands on shoulders walking

b. Stairs
Handle: under laws and regulations of construction in Taiwan, handle height shall exceed 75 cm; Architectural Society of Japan suggests 84 to 88 cm. There is on handle at Dahua Hall that is the only one with stairs in the survey in this study.

Ladders: under laws and regulations of construction in Taiwan, width shall exceed 1.2 meters and level height is 13–18 cm while level depth (ladder depth) shall be over 25 cm. Architectural Society of Japan recommends level height at 12–20 cm and depth between 20 and 36 cm. Ladder gradient is 25° to 35°. Slippery proof shall be added on the edges of ladders to avoid accidents. For ladders with over stairs, platform is required where senior people can take a rest.

Ladders in buildings in this study are wooden, which may make people fall down or have unstable gravity; steep ladders makes walking on unsafely. In light of ergonomics, sidewalks and ladders in most buildings are not satisfactory.



High threshold

4 OTHER FACILITIES

a. Fire equipment: validity of some fire distinguishers already expired; most people do not know how to use them. In fires, the function of fire distinguishers is in question.
b. Lighting equipment: visited home care centers have lighting equipment but the illumination is around 10–20 candela. The locations are not in uniform. Potential danger is foreseeable.
c. Threshold: the height better under 3 cm in construction regulations. In historic buildings, it is over 45 cm, which is bad to relief. Immediate alternative is required; movable threshold is suggested for alternative ways to accomplish this difficult.
d. Identification system in the courtyard: the senior people have weaker eyesight. They are unable to recognize small characters or little contrast between background colors and target characters. Identification shall be in large wording and strong contrast; in dark places, fluorescent lights are recommended besides identification boards. Also, high colorfulness tends to cause fatigue; warm colors cause more fatigue than cold colors. One shall pay attention to this when making identification system in colors and fonts. Only signs and protection are poor in our surveyed buildings. This may cause another problem in future disaster prevention.

5 CONCLUSIONS

Whether prevention of disaster and relief issues shall be taken into consideration in restoration of historic buildings still remain to be discussed. Other than survey on historic building cluster of the Lin's in Wufeng, further investigation and analysis on dimensions of hardware facilities of other historic buildings in Taiwan are also made. Comparison is made with regulations in Japan. Suggestions are given on laws and regulations that are out of date and principles on establishment of historic building under modification. Restoration of original appearance of historic buildings is important. In disasters and accidents, the authorities shall overcome the obstacles in the buildings. Under current living styles, one shall consider countermeasures in disasters. Modification of cultural buildings not only covers restoration of original appearance. We shall also take safety of the senior people and children into consideration.

The main purpose of ergonomics of environment is *to have objects meet the demand of people*. The nature of ergonomics, simply put, is on *how to make objects meet the demand of people*. We see systematic studies in progress but there is rare result and goal. It is our hope to build common census and establish professional research units to meet needs of all people and improve our quality of life.

REFERENCES

Architectural Society of Japan 1978. *Collection of architecture design materials*. Liu He Publication Company.

Architectural Society of Japan, 1986. *Environment planning for the physically disabled*. Japan Health Environment System Research Association.

AIA, *Architectural Graphic Standards* 1992. John Wiley & Sons Inc., 8th edition.

Bednar, M.J. 1977. *Barrier-free environment*. Dowden Hutchinson & Stroudsburg.

Diane R. Constens, *Site planning and design for the elderly, issues, guidelines, and alternatives*. Nonstrand Reinhold Company Inc.

Lawton, N.P. 1977. Current issues in planned housing for the elderly. In. *Annual meeting of the Western Gerontological Society*, Denver.

Riley, J.M. 1988. *Meeting the needs of the elderly in nursing home design*, Mechamical/Electrical Design, 42–50, Aug.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Seeing is believing: To investigate the traditional building façade colors based on human vision

Chi-Shoung Tseng
Department of Visual Communication Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

Chu-Heng Lee
Graduate school of Design Doctoral Program, National Yunlin University of Science and Technology, Yunlin, Taiwan

ABSTRACT: The existing color survey tools, color chips or color measurement instruments, are used to collect the partial color on the objects. Therefore, the color information collected is partial, which can not represent the whole color reflecting the material differences of the object. In view of this, this study used multiple human eyes, standard color chips and the distance to evaluate color as well as to explore the entire color composition of the Kinmen Jincheng Ou-Yang family temple. The result shows that it is easy to measure the object's entire color composition from evaluating the whole color in the distance. Moreover, under different conditions of distance, it shows different results from color chips. The color of granite surface in long distance (Zoom-Out) is lighter and brighter than the one in short distance (Zoom-In). The black on the door in the distance shows an opposite result of deeper darkness and a decline in brightness. As a result, the lightness of the whole color detected by human eyes in long distance shows greater contrast than the one in short distance. As for the whole, this study adopts the Zoom-In color data as a basis to compare the results with the Zoom-Out one. We can evaluate the color composition of a building façade and the visual differences phenomenon, and finally complete the color simulation results.

Keywords: color investigation, color chips, visual color, color simulation

1 INSTRUCTIONS

Color investigation has been mostly been applied in nature, culture and design. The research for environment color projects and investigation methods is taken as a case study in which environment color investigation discussion is conducted on multiple projects, district or design procedures as to submit the opinions for color investigation and application and development. Such a research uses Color Chips as the instrument, adopts Direct Visual Evaluation to conduct color investigation. (Ming-Shih Chen 1995, Jing-Hua Bu et Yue Wang 2005, C. Sáenz, B. Hernández, C. Alberdi, S. Alfonso & J.M. Diñeiro 2008). Composed of macro views, an environment color project collects natural and human colors extensively and then collected objective data are being mused for such demands as environment development. The designer provides personally-preferred color planning and suggestion to form color application and regulation. Overall, environment color project is the color project and design creation made by subjective notions.

In addition to environment color projects, E. Tantcheva, V. Cheung & S. Westland (2008) et al. applied spectrophotometer and colorimetric technology on church dome color painting to make Color Collection and Color Chips System Nomination Research trying to compare color expression of different churches by testing color observations of relic colors. The comparison observations are converted into Munsell Color System to provide 2-dimensional material vision ad communication possibilities. The procedures of such a research can be referenced by this research that the corresponding relationship between shapes and architecture colors is tried to be collected. Color application appearance can be mostly controlled according to color information collected from architecture colors. The design composed of application of certain ranges is more controllable.

"Seeing is believing" mentioned in this essay refers to discussing possible problems and difficulties in architecture color investigation and submitting one's own research design and instrument. The key points in this essay are research procedures and instrument design and research observations, regardless of interviewees' cultural problems. In the cases of architecture color investigation, Color Chips, human-vision comparison and color measurement equipments and spectrophotometers are used in Color Investigation. The frequently-seen problems of using Color Chips include: comparison standard samples' failure to cover compared objects' colors, comparison difficulties for highly reflection on the measured surface of glasses, porcelains, comparison observations representing partial colors instead of all colors. The frequently-seen problems of using color photometer: standardization problems for measured environment, deviations from instruments. The timing without human-vision investigation and confirmation usually leads to digital errors which are usually not detected until statistics for color testing data and color conversions are made.

Therefore it is affirmed in this essay that: color investigation shall be effectively and directly color to that of human-vision detection. Color photometer can not make adjustment as precisely as human-vision's adjustment against environment colors "Seeing is believing". Therefore, Color Chips Comparison Method is chosen as the major research instrument. The frequently-seen problems of aforesaid shall thus be solved accordingly. Therefore, the concrete research purpose for this research includes:

1. Using the Color Chips currently available to construct architecture overall color investigation procedures complying with human-vision.
2. Using Color Investigation data to provide possible overall color simulation observations.

2 DISCUSSIONS ON RELEVANT RESEARCHES

2.1 *Colors in term of human vision*

Human vision detection observations from eyes receive various light stimulation radiating from reflective light of photo-reception objects or from the radiation of light bulbs or TV, etc. (Richard, J.D. Tilley, 2000). The optical radiations able to stimulate human eyes are visible light of which the range is difficult to be decided. Usually, the wave length of between 380–780 nm is the visible light range (Wei-Chei Hu et al., 2007). Humans can detect colors when the energy of visible light wave enters into eyes. Color vision observations from energy are imbalanced therefore due to energy distributions from various wave lengths. Therefore, the object colors seen by human eyes are the physical object reflection color. The color decision factors before eyes receive color lights include environment lights, light-reflective materials and environment background colors, etc. The white light is the best color identification light source, followed by mercury, incandescent lamps. Yellow light sources have the highest error rate for color identification in average. (Chi-Yun You, 2001).

The aforesaid research indicates that when the viewer tries to identify colors, the light sources provided in space will affect the identification observation of color identification and cloud possibly influence the observation of color perception. Therefore, environment light source is an important controlling factor for light identification. In this research, open space is used as the research object. Colorimetric environment depends fully on natural lights and human vision color adjustment mechanism. When the detected colors are compared with the colors of standard Color Chips, the deviation made during rainy or sunny weather shall be smaller than that caused by artificial light sources. The deviation incurring in rainy or sunny weather can represent color perception observations of various local weather patterns. This research thus put emphasis on weather records as assistant reference.

2.2 *Colorimetric mechanism*

Color testing and colorimetric mechanism are conducted with planned investigation instrument. Color investigation instrument currently includes only Color Chips Investigation and Color Photometer. Being a color standardization instrument, printed or manually-made Color Chips can compare the color status of an object by zoom-in color investigation procedure. Color photometer can use specific geometric measurement to conduct the light-splitting of reflective flight, diffused light or wearing shade as to check the color status quickly. The color investigation observations of either one color investigation instrument completely depend on color environment and color surface texture. Especially, due to the demand color standards strictly used in laboratory, color photometers are not suitable for outdoor field color investigation.

However, applying the Color Chips' investigation on a sample on any practical object will create a deviation (T.R. Lee, 2000). Therefore, when Color Chips are used for making an investigation, the viewer's vision detection can't be fully reflected. Color Chips therefore pose potential problems for making color investigation. It is assumed in this essay that the reason is partial colors can't constitute the whole colors that the zoom-in colorimetric can't reflect the color appearance in term of long distance. The zoom-out colorimetric problems shall thus be solved.

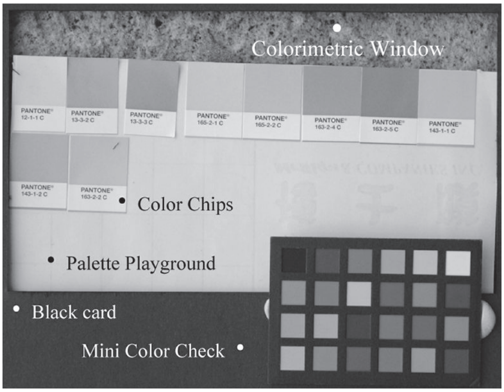


Figure 1. Zoom-in colorimetry.

2.3 *Color assimilation*

Observing how color assimilation stimulates a sample detects that mostly color spectrum create deviation and mostly diffusion are assimilation (Tsao-Huang Wei, 1993). This type of vision perception assimilation indicates that the vision perception of multiple color distribution objects can be influenced by neighboring colors and that is the observation for color physical and psychological perception. The color investigation procedures of color testing or colorimetric conduct physical analysis with the color testing condition of a single color but ignore how the assimilation or vision perception among neighbor color will influence. The mental status of influenced by the physiology of vision perception shall be investigated objectively and amended with proper statistics.

The relevant researches of color investigation conclude that those research cases record zoom-in colorimetric and color testing observations. However, when the human-eye observation mechanism is factored in, all those statistics can reflect only partial colors and still lack comprehensive color perception status. Such an issue has been rarely discussed: as distance causes deviation on human vision's perception on colors, this research has thus factored in how the viewer's distance to the object influences color perception to deduce a research observation that can be used as a factor for color planning and design.

3 RESEARCH OBJECTS, PROCEDURES AND METHODS

3.1 *Architecture and façade features in Ou-Yang family temple in Kimmen county*

Social structure in Kimmen county is constituted of caste-family cluster. Kinship developed from a single surname makes the family temple and ancestral hall the important spiritual venue. As such an emphasis is expressed directly by architecture design, the family temple and ancestral hall represent traditional social and cultural essence and features in this region. Ou-Yang family temple in Kimmen county is taken as a case study in this research. The façade structure is shown as in the Figure 2 and Figure 3 below. The façade is composed mostly of black gate, granite wall, red and black brick belt and rarely of yellow, green and white.

3.2 *Color chips selection*

A.G. Werner, the father of mineralogy in Germany, edited mineralogy sample instrument with samples for checking color names and colors in early 19th century, which has become an objective color expression method (Lu, Chin-Fu, 1988). The Color Chips have thus bee developed into such Color-order systems as Munsell (HVC), Natural Color System (NCS), and Ostwald demonstrating color specification range and order possibility. For better control in color combination and color control, some effective systems for industrial purposes have thus been developed later, such as Pantone Matching System (PMS), a standards providing an application basis for color sample selection, confirmation, paring, controlling reprinted ink colors, etc. For being a color investigation basis, Pantone Color Chips is chosen in this research. Being able

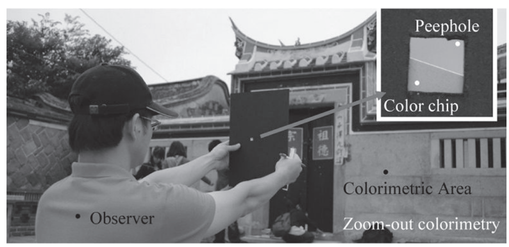


Figure 2. Zoom-out colorimetry.

Figure 3. Color chips group.

to be pasted repetitively to make colorimetric, this back-gummed Color Chips system offers pantone palette playground to expedite colorimetric and collection.

3.3 *Investigation procedures for zoom-in and zoom-out architecture façade colorimetric and color ratio*

3.3.1 *Zoom-in colorimetric*

Face colors are the major investigation object in this research. Block analysis is conducted according to architecture feature. The Ou-Yang family temple is divided into such 5 areas as wall, door, red brick belt, black brick belt decoration and ornament colors. Being too small and located in the high area, the ornament color areas are thus not considered in this research. Color Chips colorimetric is thus conducted on the colors in other 4 major areas. The matched color chips were pasted on pantone palette playground. The colorimetry observations were also videoed.

When video image is being recorded, playground shall be pasted behind the black frame and then be pasted on colorimetry architecture block. The camera can conduct white balance by using White Balance Card and also record with "mini color check" simultaneously (as shown in the figure 1). According to this standard video image recording procedures, multiple image recordings can be made on a single area. The zoom-in colorimetry for 4 architecture areas can thus be completely step by step. The color chip serial number for the zoom-in colorimetry for an architecture area can thus be recorded. The color chip serial number can be used to get the RGB standard colorimetry value to help simulate color chip feature in the monitor with software and also be the basic data for subsequent design.

3.3.2 *Zoom-out colorimetric*

It has not yet been mentioned in color investigation researches the human vision's color perception deviation in zoom-out colorimetry. The zoom-in colorimetry can only detect partial colors of the object but there is still deviation for overall colors. Therefore, in this research a peephole 1×1 cm^2 was made on a black card to create tan 45° trigonometric function distance for observing color chips and making colorimetry again. If any difference, it shall try to find the color chip the most suitable for the zoom-out color of the architecture from other color chips. It shall record the color chip serial number measured in term of long distance for this architecture area (as shown in the figure 2).

The peephole can improve the neighboring color assimilation. By making distance observation, the color values representing various color chips in the block can be vision-equalized under the status of zoom-out vague focus. Calculating such an equalized color chip value can collect the appearance of the architecture surface in term of the human-vision perception. The reason is that the human-vision still concentrates on the overall colors and the overall colors observed under a certain distance can better comply with the observer's experience. That is the breakthrough tried to be made in this research. However, when it is observed with a peephole under zoom-out, the measurable architecture area will become larger and many ornament colors will become less identifiable and could even be assimilated possibly. The investigation is made toward on large area mostly in this procedure.

3.3.3 *Human-vision's perception for color ratios*

Human vision can make measurement on overall color ratios. Such a mechanism is usually used as a method for selecting colors in arts and design. After partial colorimetry color chip groups were collected, this research team conduct color ratio appraisal according to architecture areas. The ratio combination made by this research team includes partial and overall ratios. In the case of Ou-Yang family temple, most of the area is composed of granite, among which the color chip values measured from black door, red brick and black-belt ornament are highly similar. It is therefore necessary to gain the collective human vision color ratios for the color chip groups of Ou-Yang family temple's granite. After color chip groups were checked, this investigation team decided to make appraisal by using such 4 groups of color chips as LY (light yellow), DY (deep yellow), WG (warm grey) and CG (cold grey) (as shown in the figure 3). 7 team members conducted ratio appraisal on granite color compositions in sequence and gain the ratio values ad average values as shown in the Table 1 below.

It is thus can deduce from the list above that the accumulative average value for the LY is close to one half of the average value and the overall color sensation of warm tones exceeds 90%.

Table 1. List for investigation team members' records for granite color compositions.

Chips	LY	DY	WG	CG
	40.0%	40.0%	10.0%	10.0%
	40.0%	40.0%	10.0%	10.0%
	50.0%	20.0%	20.0%	10.0%
	60.0%	30.0%	20.0%	10.0%
	50.0%	20.0%	25.0%	5.0%
	45.0%	20.0%	20.0%	15.0%
	55.0%	20.0%	20.0%	5.0%
Average	48.6%	27.1%	17.9%	9.3%

CG is the color orientation for newly-built granite. It is thus able to decide the color simulation priority sequence with the aforesaid reference digits to be used by this research for making simulation.

3.4 *Weather pattern records*

Date	Time	Temp.	Humidity	Visibility	Ground
29th, May	17~18	21.7°C	91.0%	7 Km	Dry

* Data from Kinmen weather station (2010).

4 COLORIMETRY OBSERVATIONS

The zoom-in color chip groups were thus gained. The vision appraisal of color chip group of LY (light yellow) was estimated by this team to have the highest ratio. The RGB values deduced from zoom-in and zoom-out colorimetry were thus converted into CIELab for making comparisons. The deduced digits are shown in Table 2 below and dotted into a 3-D point chart (as shown in Figure 4).

It is able to see from the figure that among the corresponding color chips made under LY-OUT zoom-out colorimetry, the brightness of L increases

Table 2. List for values of CIELab from zoom-in and zoom-out colorimetry.

Group	Chip no.	L*	a*	b*
LY-IN	12-1-1C	89.15	2.18	25.07
	143-1-1C	82.18	5.59	26.36
	143-1-2C	75.72	8.11	36.87
LY-OUT	13-2-1C	88.03	5.75	16.74
	13-2-2C	85.09	10.43	25.81
DK-IN	162-1-6C	16.16	1.15	–1.3
	162-1-7C	20.44	1.43	–1.07
DK-OUT	164-1-5C	12.92	–0.56	–6.25

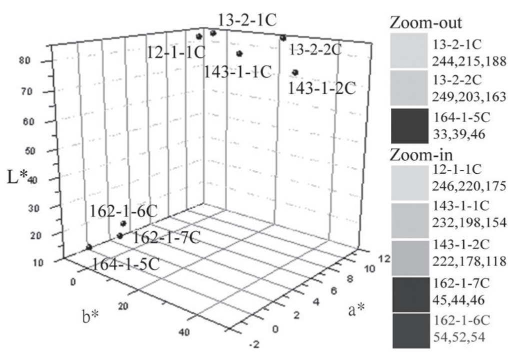


Figure 4. Comparison 3D chart for zoom-out and zoom-in colors.

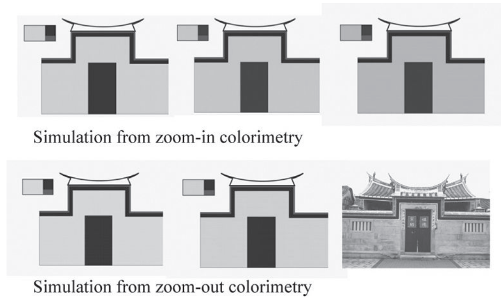


Figure 5. Simulation from zoom-in colorimetry and zoom-out colorimetry.

substantially and it means that human vision can detect more brightness toward colors under zoom-out. When the color chips gained from DK-OUT zoom-out colorimetry is observed, the brightness decreases slightly. It means that human vision can sense darker under Zoom-out than under Zoom-in.

Before color simulation is conducted, each colorimetry composition ratio in this team was checked respectively. LY Group color chip value is used for partial red bricks and black door color chip values of zoom-in colorimetry respectively to gain the color combination as shown in Figure 5. Due to human vision perception on color simulation, it is recommend by this team that LY Group Color Chip Colorimetry shall be selected for making simulation and the simulation results and corresponding photos are shown as in the Figure 5.

5 DISCUSSION AND CONCLUSION

Colorimetry was conducted under zoom-out peephole in this research to avoid colorimetry objects assimilated by environment colors and also to compare the color chip group observations of partial colors with color chip observations of overall colors and also to gain zoom-in and zoom-out architecture color records. However, the light observed via a peephole can better comply with human vision on overall architecture appearance.

It is able to find differences exist among colors after making zoom-in and zoom-out colorimetry observation. If zoom-in colorimetry observation is treated as partial vision color investigation, the zoom-out colorimetry observation could be treated as overall color investigation. It is able to conclude from the observation on Ou-Yang family temple granite colors that overall vision colors is brighter than partial vision colors because human vision gets less light intensity in zoom-in than in zoom-out.

If human vision impression on architecture façade is mostly constituted by overall vision colors, it would be able to deduce that the zoom-in colorimetry observation made in many color investigations in the past can't correspond to human vision's overall impression and there is thus a rectification for zoom-in colors and zoom-out colors is thus in demand. There thus demands more experiments for proving the rectification quantity. In the observation on the black of Ou-Yang family temple, the zoom-out colors are darker than zoom-in colors instead of being brighter. For making rectification, it is not only to adjust the light intensity of all zoom-in colors but also to consider practical colorimetry observation of different colors.

Such a conclusion doesn't mean that zoom-in colorimetry observation is unimportant because it expresses the original color appearance of architecture and can provide the application possibilities of original colors on architecture maintenance or the printed matters of design. The zoom-out colors can provide a sensation more complying with human vision's overall perception and is easier to conduct color appraisal and is suitable for documents and simulation design for video data. It is highly suggested that the comparative experiments for the relationship between those two shall be conducted as to gain standard conversion mechanism helping designers making simulation design and physical object output conversions.

REFERENCES

Bu, J.H. & Wang, Y. 2005. Objective and methodology of color landscape design. *Huazhong Architecture* 23(3): 117–120. China: Huazhong Architecture Editorial Department.

Chen, M.S. 1995. Research on environment color planning and investigation. *Arts Academic Periodical* (56): 113–127. Taiwan: Taiwan College of Arts.

Hu, W.J., Tang, S.C. & Zhu, C.F. (Eds.) 2007. *Modern color technology principles and application*. China Beijing: Publication Agency of Beijing Institute of Technology.

Lee, T.R. 1990. *Comparisons of psychological meaning of colors on samples and objects with semantic ratings*. Taiwan National Science Council Project Report NSC89–2411-H-034-005.

Lu, C.F. 1988. *Color name and colorimetry—A research for color standardization*. Taiwan National Science Council, Project Report NSC76–0301-H003-10.

Richard, J. & Tilley, D. 2000. *Colour and optical properties of materials: An exploration of the relationship between light, the optical properties of materials and colour.* New York: John Wiley & Sons.

Sáenz, C., Hernández, B., Alberdi, C., Alfonso, S. & Diñeiro, J.M. 2008. The colour of the city: Old colours and new colours. B. Bergström (eds.). *AIC 2008 Interim Meeting: Colour–Effects & Affects* (paper on 079).

Tantcheva, E., Cheung, V. & Westland, S. 2008. Analysis of 17th century church interiors using the Munsell system. B. Bergström (eds.). *AIC 2008 Interim Meeting: Colour–Effects & Affects* (paper on 139).

Wei, T.H. 1993. *Research color assimilation*. Taiwan National Science Council Project Report NSC82–0415-E036-038.

You, C.Y. 1991. *Colour identification under various light sources*. Taiwan National Science Council Project Report NSC89–2213-E-007-121.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Sick Product Syndrome (SPS): A new ergonomic approach on product design

Hassan S. Naeini
A Member of Academy, Iran University of Science & Technology (IUST), Tehran, Iran

ABSTRACT: Human demands evaluation and management of customers' needs are the main problems and dilemma for manufacturing companies and product designers. Customer behavior involves ergonomic and safe products; moreover users' satisfaction has a prominent link to marketing competition. Consequently, attention to basic human needs is crucial for marketing. Ergonomics has been making a vital role in product design, user center approach, and satisfaction. Notwithstanding importance of ergonomic design, some of the products suffer ergonomic problems which create trouble for users. I selected the term of "Sick Product Syndrome-SPS" for the mentioned problems. In this article, SPS is introduced as a new approach based on ergonomics considerations in product design. SPS outlines help to designers to explore product's deficiency. For product diagnosis in terms of ergonomics, several sorts of related standards and checklists were assessed, and then according to them, I chose a cooker as a home appliance, and modified an ergonomics check point for its assessment. After assessment of sample cooker and exploration of ergonomics problems, some recommended for its redesigning was introduced. Then a new design was presented according to gathered data. Sick Product Syndrome, as an ergonomic approach, seeks to explore products' pros and cons, also this syndrome shows ergonomics problems of product, and under this condition product's problem will be highlighted for any revision or redesign process.

Keywords: ergonomics, sick product, design, customer, product, safety

1 INTRODUCTION

Nowadays, both of products manufacturer and customers face with some dissatisfaction. Producers have some problems in terms of appropriated efficiency in marketing, and costumers suffer some awful products which are not fit to human needs. Uder this condition products designers roles are crucial, as they are make a good linkage between users needs and producers expectancies; and in this circumstance, ergonomics as an human based art and science plays a vital role, as well. Despite of well known effects of ergonomics consideration in terms of products design according to human needs, we are still facing several sorts of problems with communication with tools and different products, as well. Therefore, writer believe that products mostly suffer some degrees of illness and syndrome; and as diseases treatment need to diagnosis, so I believe that Ergonomics might be good at diagnosis and treatment of the mentioned products problems.

The writer has already selected the term of "Sick Product Syndrome-SPS" for the mentioned observable fact. In this article, SPS is introduced as a new approach based on ergonomics considerations in product design.

2 SICK PRODUCT SYNDROME (SPS)

In writer's opinion, most of products need rearrangement or redesigning for changing the products to better condition. In this context, both of physical and mental demand should be considered, not only for customers but also for employees. Also, I chose the term of *syndrome* to show products' illness, because products related problems in terms of health, safety, and ergonomics include a mixture of the mentioned problems. For instance a TV remote control may have a nice shape but unfair dimension or have some problems in terms of knobs layout; a toy for 3–5 years old kids may have a good feature with nice colors, but has a sharp edges; and several similar other examples. These samples suffer some degrees of SPS. Undoubtedly, some time these problems might be danger and create risky conditions for users, as well.

3 ERGONOMICS

Ergonomics is a multidisciplinary science which has been changed to beyond the work station and Industrial sectors environments. Ergonomics applications in product design have created a new approach in this field, also consumer behavior, user center design, and macro-ergonomics have been made some new views in ergonomics scope (Naeini 2008).

One of the related fields in ergonomics and human factors design is product design. Under this condition some other disciplines such as industrial design are involved. Furthermore, as the most of products have a close connection to users, so every bad design in terms of ergonomics might be affect on users. Ergonomics problems might be tangible for users more than other products' problems. Ergonomics introduces some practical methods for products evaluation, also work stations and environmental factors such as lighting and noise can be done with some ergonomics methods, i.e. RULA, OWAS.

4 METHOD

In this study, as a first step, most of ergonomics check points and ergonomics ISO were studied. According to the main themes of read questionnaires and checklists, some fields were extracted, i.e. display, controls, anthropometric based dimension, environmental ergonomics, handling, reaches, labeling, coding, and so on. In the second phase, some meeting with some experienced designers and ergonomics experts as brainstorming workshop were held in which participants were involved the evaluation of ergonomics products characteristics. For categorizing of experts opinions and making a fair priority for the main ergonomics features of products, the method of *analytical hierarchy process AHP* was used. Consequently, the first ergonomic checklist includes weighting scores was introduced as a main result of the mentioned brain storming. Then a home cooker included 5 flames and an oven was selected as a product sample (Figure 1).

In the third phase of study, according to preliminary assessment of the cooker, the first ergonomic checklist included ergonomics, safety, and health checkpoints, was re-evaluated. In the ergonomics section some themes were followed by related checkpoints i.e. display and controls layout, dimensions, signs, handling, flame controls; in the safety sections, the themes were some subjects such as, risk condition, gas leakage, and so on. In the health section some items were followed, for instance, wash ability and cleaning procedure.

Figure 1. A general cooker with an oven and gas flames.

Then the cooker was assessed by the final checklist and its total scores which shows degree of its problems as a sick products, was extracted. in addition, the mentioned checklist included three parts (ergonomics, safety, and health).

Considering with cooker ergonomic problems, some recommendations were introduced. In the forth step, another meeting with some industrial designers were hold, and the first recommendation changed to feasible and practical ones. Finally a new model was design, according to introduced recommendations.

5 RESULTS

This study was ended to an ergonomics checklist for evaluation of ergonomics problems of product according to related standards. In this checklist every box has a weighted score which shows value of related item, so summation of scores shows the overall score of evaluated product. The maximum score of total parts of checklist is 1000, 700 and more got score means the sample product has a fair condition in terms of ergonomics, and score up to 300 shows that the sample score suffers some ergonomic problem, as well.

REFERENCE

Naeini, S.H. 2008. Ergonmics principle on manual materla handling systems. *Tehran. Fanavaran Pub.*: 201.

Part VII: Ergonomics for people with special needs

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Evaluation of the cognitive ability among aging groups using mobile phone

Jung-Yong Kim & Young-Jin Cho
Hanyang university, Ansan-si, Korea

ABSTRACT: The aim of this study is to evaluate the cognitive ability with aging groups who use mobile phone. This study consisted of two parts: development of questionnaire and assessment of cognitive ability. Questionnaire included discomfort questions covering four cognitive areas and sixteen cognitive design elements. Forty subjects participated and were divided into four age groups: elderly group (65 and above), high-middle aged group (55–64), middle aged group (45–54) and young group (20–44). Elderly users' performance score was lower than young users in comprehension tests. High-middle aged and middle aged users showed similar results like elderly users. In vocabulary test, there was no difference in incorrect rate among age groups. The perceived level of icon and text size increased as age increased. It is expected that the methodology and the result of this study could be used to design future mobile phone for elderly and aging population.

Keywords: elderly people, cognitive ability, mobile phone

1 INTRODUCTION

Recently, there have been many advances in Information Technology (IT). Most of all, digital convergence is the most significant trend in designing IT products. A mobile phone is the most frequently used among IT products, which is integrated with other functions such as camera, music playing, navigation, Digital Multimedia Broadcasting (DMB) and so on.

In Korea, mobile phones have been being prevalent in all age groups and genders. According to the report of Korea Communications Commission (2009), subscribers of mobile phone covered about 98 percent of Korean population. This means that many people have more than one phone. Since young users have used the mobile phones from childhood, they are familiar with using mobile phones. However, elderly users have had the difficulty in learning and using the various functions of the mobile phone because they have physical and cognitive limitations.

Many mobile phones for elderly users have been developed by enlarging the letters to compensate the weak eyesight. However, the large font size cannot compensate the decline of cognitive function since elderly users usually have cognitive limitations as well as physical ones. These problems had been discussed by many authors (Fisk et al. 2004, Scholtz et al. 2001, Kroemer & Grandjean 1997, Salthouse et al. 1984).

Therefore, the aim of this study is to evaluate the cognitive ability of various aging groups who use mobile phone. Such quantitative information can be used to design mobile interface for elderly groups and various aging groups.

2 METHODS

2.1 *Definition of age groups*

According to the definition of the United Nations (UN), the age of 65 and above can be considered elderly people. However, age classification was not always so straightforward (Fisk et al. 2004). Therefore, it was needed to clearly classify the age groups for the study. In this study, we had two age groups to be elderly in next two decades. The rest of them were defined as young group. Therefore, age groups were as followed: elderly group (65 and above), high-middle aged group (55–64), middle aged group (45–54) and young group (20–44) (Lee et al. 2009, Kim et al. 2009).

2.2 *Cognitive areas and cognitive design elements*

In order to define cognitive areas for using mobile phone, Mini-Mental State, The 7 Minute Screen and Wechsler Adult Intelligent Scale-Revised were used, which were developed to clinically examine the cognitive ability of people (Koh et al. 2000, Han et al. 1989, Kwon et al. 1989, Wechsler 1981, Folstein et al. 1975, Jeon & Seo 1963, Wechsler 1955). The cognitive areas which were highly related with the use of mobile phone were extracted from them. Simultaneously, the cognitive design elements for mobile phone design were extracted based on previous researches. Then the cognitive areas and the cognitive design elements matched each other. Finally, four cognitive areas such as language, visuospatial ability, perception and memory, and sixteen cognitive design elements were selected (see Table 1).

2.3 *Major cognitive design elements*

It was inefficient to examine all cognitive design elements because the discomfort of mobile phone was not always affected by all cognitive design elements. Therefore, it was needed to find the major cognitive elements affecting the discomfort.

In order to find the major cognitive design elements, a preliminary questionnaire with discomfort questions covering four cognitive areas and sixteen cognitive design elements were made. 240 subjects responded to the survey. Subjects were classified into four age groups consisted of Young Group (YG), Middle aged Group (MG), High-Middle aged Group (HMG) and Elderly Group (EG).

In the results, button comprehension, icon size, icon comprehension, menu breadth, menu depth, menu comprehension, vocabulary and font size were found to be the design elements strongly associated with the discomfort of using mobile phone among aging and aged population in Table 1. Eight major cognitive design elements were finally selected.

Table 1. The discomfort ratio among age groups.

Design elements		YG %	MG %	HM %	EG %
Button	Comprehension	31.7*	30.0*	48.3*	38.3*
	Position	26.7*	21.7	5.0	3.3
Icon	Size	6.7	16.7	20.0	23.3
	Comprehension	36.7*	25.0*	38.3*	26.7*
Menu	Breadth	25.0*	28.3*	33.3*	26.7*
	Depth	25.0*	35.0*	43.3*	31.7*
	Comprehension	26.7*	30.0*	35.0*	35.0*
	Vocabulary	21.7	16.7	11.7	25.0*
Text	Font size	8.3	26.7	48.3*	40.0*
	Color	8.3	3.3	0.0	8.3
	Contrast	6.7	1.7	1.7	3.3
Display	Front display	11.7	30.0*	20.0	13.3
	Main display	10.0	23.3	21.7	16.7
Sound	Volume	41.7*	15.0	20.0	23.3
	Types	11.7	5.0	1.7	1.7
Vibration	Degree	30.0*	21.7	16.7	18.3

*The discomfort ratio is greater than a fourth.

2.4 *Questionnaire development*

Six cognitive design elements except for the elements of menu breadth and menu depth were used to develop the questionnaire for examining the cognitive ability of age groups because the elements of menu breadth and menu depth were dependent on menu structure. Generally, there was no solution for tradeoffs of breadth and depth in the menu structure (Fisk et al. 2004). Questionnaire developed in this study was composed of 40 questions: 9 questions for button comprehension, 8 questions for icon comprehension, 10 questions for vocabulary, 8 questions for menu comprehension, 3 questions for icon size and 2 questions for text size. Table 2 showed the questionnaire with 40 questions.

2.5 *Assessment of cognitive ability*

To assess the cognitive ability among age groups, forty subjects participated in the experiment. Subjects with equal number of male and female were divided into four age groups: young group, middle aged group, high-middle aged group and elderly group in Table 3. The subjects were interviewed with one-to-one due to passive response to written questions, which was very common psycho-social attitude among elderly people in Korea.

3 RESULTS

3.1 *Button comprehension*

The performance time of elderly users was about 3.2 times longer and the incorrect rate was about 7.8 times greater than younger users on average in Tables 4 and 5. Middle aged and high-middle aged users showed similar results each other. The performance time of middle aged and high-middle aged users was about 2.2 and 2.5 times longer and the incorrect rate was about 4.0 and 4.8 times greater than young users.

3.2 *Icon comprehension*

The results of performance time and the incorrect rate were different among age groups in Tables 6 and 7. The performance time and the incorrect rate among age groups were different one after another. The performance time of middle aged, high-middle aged and elderly users was about 1.8, 2.5 and 3.3 times longer than young users. The incorrect

Table 2. Questionnaire with 40 questions.

Design elements	Question lists	
Button comprehension	# 1	Power
	# 2	Send
	# 3	Message
	# 4	Phonebook
	# 5	Menu
	# 6	Navigation
	# 7	Cancel
	# 8	Etiquette mode
	# 9	Lock
Icon comprehension	# 1	Send
	# 2	Service
	# 3	Message
	# 4	Alarms
	# 5	Vibrate
	# 6	Etiquette mode
	# 7	Battery
	# 8	Lock
Vocabulary comprehension	# 1	Send
	# 2	Cancel
	# 3	OK
	# 4	Menu
	# 5	Ring/vibrate
	# 6	Message
	# 7	Speed dial
	# 8	Morning call
	# 9	Display
	# 10	Call list
Menu comprehension	# 1	Add new entry
	# 2	Ringtone Style
	# 3	Alarms
	# 4	Display
	# 5	Lock
	# 6	Search number
	# 7	Message
	# 8	Font Size
Perceived icon size	# 1	Sounds
	# 2	Calendar
	# 3	Settings
Perceived text size	# 1	Text on input button
	# 2	Text on display screen

Table 3. The information of subjects.

	Age	Experience
Age groups	Years	Months
Young Group (YG)	29.7 ± 2.8	116.4
Middle aged Group (MG)	51.1 ± 2.5	115.2
High-Middle aged Group (HMG)	61.6 ± 1.6	116.4
Elderly Group (EG)	70.5 ± 4.3	113.4

rate of young users was almost zero, whereas the incorrect rate of elderly users was about 40% at given questions on average.

Table 4. The performance time in button comprehension test.

	YG	MG	HMG	EG
Questions	s	s	s	s
Power	1.3 ± 1.0	2.9 ± 1.4	2.7 ± 1.5	3.6 ± 2.5
Send	1.2 ± 0.7	2.6 ± 1.2	2.6 ± 1.3	3.8 ± 3.2
Message	1.3 ± 0.6	2.9 ± 1.2	3.0 ± 1.5	5.3 ± 5.6
Phonebook	1.5 ± 0.7	3.2 ± 1.3	4.8 ± 4.9	6.6 ± 6.8
Menu	1.4 ± 0.7	3.4 ± 1.6	3.0 ± 1.4	5.1 ± 2.2
Navigation	1.3 ± 0.7	3.2 ± 1.7	2.3 ± 1.4	3.2 ± 1.8
Cancel	1.9 ± 1.2	3.2 ± 1.5	3.4 ± 2.8	4.7 ± 2.9
Etiquette mode	1.9 ± 0.8	4.4 ± 1.7	4.8 ± 2.6	4.5 ± 2.3
Lock	1.6 ± 0.7	4.0 ± 1.5	7.0 ± 7.6	6.0 ± 2.9

Table 5. The incorrect rate in button comprehension test.

	YG	MG	HMG	EG
Questions	%	%	%	%
Power	0	0	10	0
Send	0	0	0	0
Message	0	0	10	50
Phonebook	0	10	0	20
Menu	0	20	20	30
Navigation	0	10	0	30
Cancel	10	40	20	10
Etiquette mode	10	30	50	70
Lock	20	50	80	100

Table 6. The performance time in icon comprehension test.

	YG	MG	HMG	EG
Questions	s	s	s	s
Send	1.9 ± 0.9	3.6 ± 1.4	4.6 ± 1.8	7.7 ± 5.0
Service	2.3 ± 0.7	3.2 ± 0.7	6.3 ± 3.4	7.2 ± 6.4
Message	2.0 ± 0.6	4.2 ± 1.8	5.8 ± 2.8	8.8 ± 6.4
Alarms	2.4 ± 0.9	4.5 ± 1.2	5.3 ± 2.3	5.3 ± 2.6
Vibrate	2.3 ± 1.1	5.2 ± 1.8	6.3 ± 3.0	10.8 ± 12.2
Etiquette mode	2.3 ± 0.7	3.7 ± 1.1	6.3 ± 2.4	6.3 ± 3.7
Battery	2.2 ± 0.8	3.4 ± 1.8	4.4 ± 1.5	4.5 ± 2.3
Lock	2.0 ± 1.0	3.9 ± 1.4	5.6 ± 2.6	7.8 ± 5.1

3.3 *Menu comprehension*

The performance time of middle aged, high-middle aged and elderly users was about 2.8, 3.1 and 4.6 times longer than young users in Table 8. Therefore, middle aged and high-middle aged users had the similar cognitive ability in performance time. However, in the results of the incorrect rate, they were different each other in Table 9.

Table 7. The incorrect rate in icon comprehension test.

	YG	MG	HMG	EG
Questions	%	%	%	%
Send	0	10	20	20
Service	0	0	10	40
Message	0	10	20	50
Alarms	0	0	0	50
Vibrate	10	40	50	80
Etiquette mode	0	50	60	80
Battery	0	0	10	20
Lock	0	20	30	60

Table 8. The performance time in menu comprehension test.

	YG	MG	HMG	EG
Questions	s	s	s	s
Add new entry	3.0 ± 0.8	7.6 ± 2.6	7.5 ± 4.3	12.5 ± 10.6
Ringtone type	2.9 ± 0.7	6.2 ± 1.9	8.1 ± 4.8	11.0 ± 7.7
Alarms	2.0 ± 0.6	5.8 ± 1.4	6.2 ± 3.0	11.8 ± 8.9
Display	3.0 ± 1.2	8.3 ± 2.6	9.0 ± 4.4	11.0 ± 6.1
Lock	2.2 ± 1.1	6.0 ± 1.9	8.5 ± 6.4	16.1 ± 13.5
Search number	2.3 ± 1.0	6.2 ± 2.6	6.6 ± 4.4	11.9 ± 14.8
Message	2.4 ± 1.0	8.1 ± 4.6	7.9 ± 3.3	9.2 ± 8.7
Font size	2.1 ± 0.9	7.3 ± 3.5	7.9 ± 5.3	7.8 ± 4.7

Table 9. The rate of incorrect answers in menu comprehension test.

	YG	MG	HMG	EG
Questions	%	%	%	%
Add new entry	0	30	40	40
Ringtone type	10	20	20	30
Alarms	0	0	10	10
Display	40	80	40	40
Lock	0	0	10	10
Search number	0	10	10	40
Message	20	0	50	40
Font size	30	10	0	30

3.4 *Vocabulary comprehension*

There was no difference among middle aged, high-middle aged and elderly users in the performance time. The performance time of them was about 1.4 times longer than young users in Table 10.

Table 10. The performance time in vocabulary comprehension test.

	YG	MG	HMG	EG
Questions	s	s	s	s
Send	6.0 ± 1.1	8.0 ± 1.8	8.7 ± 2.0	11.3 ± 8.1
Cancel	6.1 ± 1.9	9.1 ± 2.3	8.2 ± 3.4	10.1 ± 3.9
OK	6.4 ± 1.6	11.0 ± 4.1	10.0 ± 3.0	13.4 ± 7.1
Menu	7.8 ± 1.8	12.2 ± 2.5	12.4 ± 4.6	13.9 ± 7.1
Ring vibrate	9.1 ± 2.8	9.7 ± 1.8	11.7 ± 3.1	10.3 ± 4.4
Message	7.3 ± 1.2	9.6 ± 2.2	8.8 ± 3.8	11.6 ± 5.3
Speed dial	7.2 ± 1.3	9.9 ± 5.4	10.1 ± 4.2	8.8 ± 5.9
Morning call	5.2 ± 0.8	9.2 ± 3.1	8.2 ± 3.6	10.0 ± 6.5
Display	10.2 ± 2.4	13.2 ± 5.0	14.1 ± 4.6	10.2 ± 5.6
Call list	15.4 ± 3.6	19.1 ± 7.2	23.6 ± 6.7	24.5 ± 16.9

Table 11. The rate of incorrect answers in vocabulary comprehension test.

	YG	MG	HMG	EG
Questions	%	%	%	%
Send	0	0	0	20
Cancel	0	10	0	20
OK	0	30	0	30
Menu	0	20	30	30
Ring vibrate	0	0	0	0
Message	0	0	0	0
Speed dial	0	10	10	40
Morning call	0	30	10	30
Display	20	50	70	70
Call list	0	0	0	50

Table 12. The perceived icon size.

	YG	MG	HMG	EG
Questions	mm	mm	mm	mm
Sounds	5.4 ± 0.3	6.4 ± 0.6	8.0 ± 1.6	8.4 ± 1.5
Calendar	5.5 ± 0.4	6.0 ± 0.8	7.1 ± 0.8	7.4 ± 1.3
Settings	5.6 ± 0.5	6.7 ± 0.8	8.1 ± 2.1	8.4 ± 1.6

The elderly users showed high correct rate in some vocabulary, however, they did not in unclear vocabulary in Table 11.

3.5 *Perceived icon size*

The perceived icon size for elderly and high-middle aged users was about 8.1 mm and 7.7 mm on average, whereas about 6.4 mm for middle aged users and 5.5 mm for young users in Table 12. The mean of the perceived icon size of elderly users

Table 13. The perceived text size.

	YG	MG	HMG	EG
Questions	font	font	font	font
Input button	7.4 ± 0.5	9.1 ± 0.9	10.4 ± 0.8	10.9 ± 1.1
Display screen	8.2 ± 0.4	9.1 ± 0.9	10.5 ± 0.8	11.0 ± 1.1

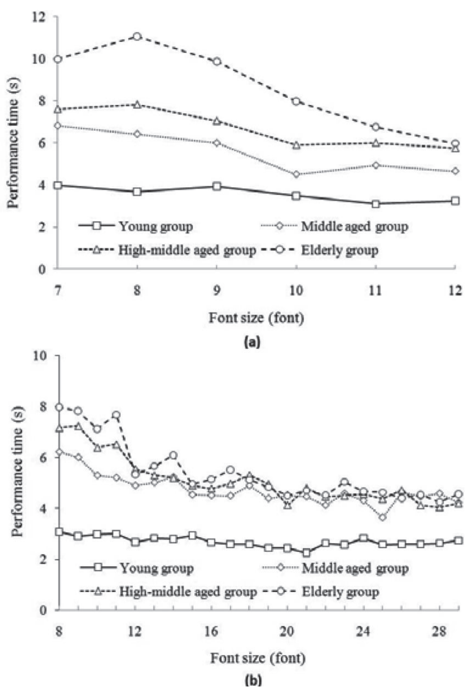


Figure 1. The changes of the performance time as (a) input button text size and (b) display screen text size.

was about 1.5 times and the standard deviation was about 3.8 times greater than young users.

3.6 *Perceived text size*

The perceived text size in input button and display screen was shown in Table 13. The font size for elderly users to well recognize the text on input button and display screen was about 1.5 and 1.3 times greater than young users. The variability of elderly users was about 2.1 and 2.5 times greater than young users.

Figure 1 (a) showed that the performance time decreased as the button text size increased. In particular, the performance time of elderly users varied with big changes. The performance time in display text size was shown in Figure 1 (b). It showed that performance time of middle aged, high-middle aged and elderly users similarly decreased as display text size increased. However, the performance time of young users was not changed.

4 DISCUSSION

In order to evaluate the difference of the cognitive ability among age groups, the contents of questionnaire included the icon, letters and terms extracted from the mobile phone. Since the icon, letters and terms were different among the mobile phone manufacturer, we selected the commonly used items and examined them. The icons used in this study are mostly compatible with mobile phones made by other countries. However, the letters and terms are not compatible each other because of language difference.

The performance time in button, icon, menu and vocabulary test was different among age groups. Middle aged and high-middle aged users showed the similar results in button and menu comprehension test. On the other hand, two groups were different in icon comprehension test. In vocabulary test, they showed the similar results with elderly users. Therefore, middle aged and high-middle aged users had their own characteristics unlike other age groups.

The incorrect rate was also different among age groups. Middle aged and high-middle aged users showed the similar results in button comprehension test but they were different each other in other tests. The incorrect rate increased as age increased on average. In menu and vocabulary comprehension test, the incorrect rate except young users was high when users were asked to answer to unfamiliar or unclear menu and vocabulary. Therefore, the more frequently a person used the functions of mobile phone, the lower the incorrect rate was among groups.

In perceived icon and text size test, the mean size of icon and text increased as age increased. At the same time, the variability of perceived size also increased as age increased. The age group difference between elderly and young users in icon size was shown to be about 1.5 times in terms of mean value and 3.8 times in terms of standard deviation. In text size, the performance time of elderly users was shown to be about 1.4 times greater in terms of mean values and 2.3 times greater in terms of standard deviation value than young users. It indicated that elderly users had a problem due to a personal difference rather than age itself in visual perception. The results showed high-middle aged users were similar with elderly users, whereas they were different from middle aged users.

This study tried to measure the difference of the cognitive ability among age groups in using the mobile phone. Despite of the high penetration rate of mobile phones, the common mobile phones have not been sufficiently considered the difference of the cognitive ability among age groups. Therefore, it was needed to consider the cognitive ability of middle aged and high-middle aged users as well as elderly users.

5 CONCLUSIONS

The performance time and incorrect rate increased as age increased. Middle aged and high-middle aged users have the higher cognitive ability than elderly users in some tests. However, in some tests, they have the similar cognitive ability with the elderly users.

It is expected that the methodology and the result of this study could be used to design future mobile phone for elderly and various aging population.

REFERENCES

Fisk, A.D., Rogers, W.A., Charness, N., Cazja, S.J. & Sharit, J. 2004. *Designing for Older Adults*. New York: CRC Press.

Folstein, M.F., Folstein, S.E. & McHugh, P.R. 1975. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12: 189–198.

Jeon, Y.S. & Seo, B.Y. 1963. KWIS (Korean Wechsler Intelligence Scale). Institute for education: Seoul.

Kim, J.Y., Cho, E.J., Cho, Y.J. & Chung, M.K. 2009. Development of measuring protocol for cognitive ability of elderly people for mobile phone. *Journal of the Ergonomics Society of Korea* 28(3): 115–123.

Koh, H.J., Yoo, T.W., Cho, B.L., Kim, D.K., Kim, K.S., Yang, J.H. & Lee, H.J. 2000. Clinical efficacy of the 7 Minute Screen at primary care setting. *Korean Journal of Family Medicine* 21(9): 1131–1138.

Korea Communications Commission 2009. *Telephone penetration rate and subscribers of mobile phone and high-speed internet*. Seoul.

Kroemer, K. & Grandjean, E. 1997. *Fitting the task to the human: A Textbook of Occupational Ergonomics*. London: Taylor & Francis.

Kwon, Y.C. & Park, J.H. 1989. Korean Version of Mini-Mental State Examination (MMSE-K). *The Korean Journal of Neuropsychiatric Association* 28(1): 125–135.

Lee, D.H., Chung, M.K. & Kim, J.Y. 2009. An investigation of using practices for universal design of information technology products. *Journal of the Ergonomics Society of Korea* 28(3): 103–114.

Park, J.H. & Kwon, Y.C. 1989. Standardization of Korean of the Mini-Mental State Examination (MMSE-K) for Use in the Elderly. Part II. Diagnostic Validity. *The Korean Journal of Neuropsychiatric Association* 28(3): 508–513.

Salthouse, T.A., Rogan, J.D. & Prill, K.A. 1984. Division of attention: Age differences on a visually presented memory task. *Memory & Cognition* 12(6): 613–620.

Scholtz, A.W., Kammen-Jolly, K., Felder, E., Hussl, B., Rask-Andersen, H. & Schrott-Fischer, A. 2001. Selective aspects of human pathology in high-tone hearing loss of the aging inner ear. *Hearing Research* 157: 77–86.

Wechsler, D. 1955. *Manual for the Wechsler Adult Intelligence Scale (WSIA)*. New York: The Psychological Corporation.

Wechsler, D. 1981. *Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. New York: The Psychological Corporation.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A study of different age users' evaluation on museum guide map design

Chun-Ching Chen
National Taipei University of Technology, Taipei, Taiwan

Ying-Ju Chieu
Ming Chuan University, Taoyuan, Taiwan

ABSTRACT: Digital technologies trigger the development of digital guiding systems for the museum, however, printed guide map, providing the visitor convenience and instant information input during the visit, still show its unique in guiding the visitor. This study aims to investigate this issue, through questionnaire surveying museum visitors' experience on using guide maps. The results show that different age groups perform similarly on rating the contents and portability of current guide maps, but have different opinions on the aesthetics and usability, particularly for the aged group. In conclusion, this study demonstrates the effect of age on some attributes of guide map design, and suggests appropriate knowledge levels when designing guide maps for different age visitors.

Keywords: guide map, information design, age

1 INTRODUCTION

1.1 *Background*

Guide map is the first document that the visitor obtains, and allows the user to get start with their move shortly at the museum (Lo, 2004). In other words, museums should provide guide maps with necessary information for visitors during their visits to the place. A user-friendly designed guide map can not only help visitors find the right directions, but also produce satisfactory visiting experience of the place. On the contrary, poor designed guide could make users get lost and spend additional time in finding their way rather than the exhibition (Hsieh, 2000). It is no doubt that guide map design is crucial to visitors and museums.

Guide map can always be found at the entrance of a museum. Normally it comprises of topic and details of the exhibition, floor plan, and other information relevant to the visit. Visitors not only read the instruction at the start, but also refer to the guide throughout their visits. This means the guide map should be compact in size, but rich in content. Additionally, a guide map does not need particular maintenance cost in personnel, software or hardware. This make it being the most common and economical museum guide system, even digital technology has been applied for years.

One of the most important functions for guide map is wayfinding. However, previous studies found it could be confused to understand the floor plan, as different types of space presentation are applied, such as 2D or 3D drawing. Some users are not benefited from the colorful or virtual real looking floor plan instruction, due to different user perception in space transformation, particular for aged people. Designers may try to add "all" information regarding the place to the users, however, this could not be practically useful as users might need only "adequate" information to suit their needs. User experience should be a key input when such guide map is designed. Visitors including children, families, elders or professionals behave variously during their stay in the museum, which indicates that they may have different requirements on the guide map. Therefore, considering visitors' preferences when designing guide maps is necessary.

As mentioned, the development of digital technologies brings various types of guiding systems of museum, such as touch screen, portable voice guide, or other ongoing developing interactive guiding systems. However, the convenience, portability and instant information input of printed guide map, demonstrates its unique function in guiding the visitor, and cannot be replaced. For previous studies concerning museum guiding system design, many focus on web-based digital museum guiding (Lin, 2002, Kao, 2003, Chung, 2004, Huang, 2005), exhibition contents (Fan, 1996, Lin, 2003, Lu, 2004, Hong, 2005, Jesus, Dias, Frias and Correia, 2007), signpost system (Hsieh, 2000, Liu, 2001) or development of digital interactive technologies (Yi, 2003, Mantyjarvi, Paternò, Salvador and Santoro, 2006). Recently some researches are conducted to investigate portable guiding system (Ciavarella and Paternò, 2004, Raptis, Tselios and Avouris, 2005, Herrmann, Kawalek and Stark, 2005, Damala and Kockelkorn, 2006, Kusunoki, Satoh, Mizoguchi and Inagaki, 2007). Although Lo (2004) studied the visual communication of guide map design, he did not address much on different users' perspectives.

1.2 *Aim*

Given the above reasons, it is clear to see that the design of guide map could differ from age in drawing, icons, contents ... etc. Users can only be benefited when considering their needs. However, few works focused on different age users' feedbacks on guide maps. This study aims to investigate this issue, through surveying museum visitors' experience on using guide maps, so as to propose suggestions on design.

2 METHOD

2.1 *Questionnaire*

This research uses questionnaire to survey museum visitors. Visitors from the most popular places in Taiwan, National Palace Museum and Taipei Zoo, were selected to participate the survey. The questionnaire comprises three parts: (1) participant's personal details; (2) evaluation of current guide maps, using 7-point Likert scale to assess design attributes on "contents", "aesthetics", "usability" and "portability" of guide map, (3) preferred detail design information of guide map between different age users, using both close—and open-ended questions.

2.2 *Sampling*

Stratified sampling was adopted to collect participants from 5 groups (including age groups from 10 to 19, 20 to 29, 30 to 39, 40 to 49 and over 50 years old). Ten volunteers were invited for each museum and age group. The results were analyzed using descriptive statistics on users' preference on guide map design, and one-way ANOVA on age group and design attributes and Scheffe Post Hoc to find out the significance. Participants' subjective opinions on design were also analyzed.

3 RESULTS AND DISCUSSION

Questionnaires were carefully examined to pick out those with incomplete feedback. In total 103 effective questionnaires were collected (51 from National Palace Museum and 52 from Taipei Zoo) for analysis.

3.1 *Evaluation of current guide map*

In this part of questionnaire, 7-point Likert scale was used to assess contents, aesthetics, usability and portability of current guide maps from Taipei Zoo and National Palace Museum. The scoring of the questionnaire is "7" for "Agree", "4" for "Undecided" and "1" for "Disagree" (The results are shown in Appendix A.3); scoring between shows the degree.

3.1.1 *Contents*

The questions used to assess "contents" of guide maps include "location" of the visitor, information to prevent from "getting lost" in the museum, information for the visitor to find his/her "destination" and "route", and the marks of "entrance/exit" of the museum. The result shows no significance between different age groups' opinions on guide map contents (see Table 1). Across the age visitors agree with the contents of guide maps despite minor difference.

3.1.2 *Aesthetics*

Evaluation of "aesthetics" on guide maps includes questions of "typography", "graphic design", "color", "layout", and "attractiveness". From Table 2, it is clear to see the significantly different feedbacks in "graphics" and "color".

Further paired analysis on the questions with significance found age group over 50 agree significantly more than age group at 20–29, 30–39 and 40–49 (Table 3), which means elder people in this

Table 1. Significance on "contents" of guide map.

Question	Mean	SD	P value
Location	5.750	1.453	0.065
Getting lost	4.596	1.933	0.827
Destination	5.712	1.486	0.503
Entrance/exit	5.577	1.808	0.142
Route	5.192	1.837	0.318

Table 2. Significance on "aesthetics" of guide map.

Question	Mean	SD	P value
Typography	5.558	1.320	0.720
Graphics	5.558	1.335	0.027*
Color	5.596	1.332	0.038*
Layout	5.288	1.564	0.251
Attractiveness	5.173	1.465	0.259

* shows significance.

survey are more satisfied with the "graphics" and "color" design of the guide maps than the others.

3.1.3 *Usability*

The questions applied to evaluate "usability" attribute of guide maps are "readability", "comprehensibility", "learnability", "error" occurred in using, and the "helpfulness" of the guide map. The result shows significant difference in most questions except for "error" between age groups (see Table 4), although all groups scored positive accordingly.

Similar phenomenon to the "aesthetics", further paired analysis on the questions with significance of "usability" found age group over 50 scored higher than the other age groups, while no significance were found between each other 10–19, 20–29, 30–39 and 40–49, particular on "readability" and "comprehensibility" (Table 5). Age at 10–19 scored close to the elders on "learnability" and "helpfulness", also on the other three questions, which could show similar usability requirement of both groups.

3.1.4 *Portability*

Evaluation of "portability" on guide maps includes questions of "habit" of asking for guide maps, "convenience" of using, "size", "collection" of used guide maps, and "intention" of collecting the current guide maps (Taipei Zoo and National Palace Museum). The result shows no significance between different age groups' opinions on guide map portability (see Table 6). However, lower, but not significantly different scores were given on the "collection" and "intention" questions, which means visitors could get rid of the guide maps after their visits.

Overall, different age groups perform similarly on rating the contents and portability of current guide maps, but have different opinions on the aesthetics and usability, particularly for the aged group. Younger visitors (age 10–19) also score similarly to elderly people in some aspects. This means current guide map design should be focused more on considering needs from the teenagers and elders.

3.2 *Preferable design of different age groups*

To sum up the information collected from the third part of the questionnaire, including subjective suggestions, results of different age groups' preferable guide map design are discussed here.

For contents of guide map, most participants (excluding age 30–39) prefer "all-in-one" design, which means the guide map should provide not only wayfinding, but also detailed information

Table 3. Paired analysis on significance of aesthetics.

Question	Age (I)	Age (J)	Mean (I)–(J)	P value
Graphics	>50	10–19	0.812	0.560
		20–29	1.078*	0.016*
		30–39	1.381*	0.002*
		40–49	1.124*	0.009*
Color	>50	10–19	0.633	0.139
		20–29	1.208*	0.008*
		30–39	1.281*	0.004*
		40–49	0.924*	0.033*

* shows significance.

Table 4. Significance on "usability" of guide map.

Question	Mean	SD	P value
Readability	5.673	1.396	0.005*
Comprehensibility	5.769	1.308	0.002*
Learnability	5.654	1.558	0.028*
Error	4.692	1.832	0.967
Helpfulness	5.981	1.163	0.005*

* shows significance.

Table 5. Paired analysis on significance of usability.

Question	Age (I)	Age (J)	Mean (I)–(J)	P value
Readability	>50	10–19	1.185*	0.011*
		20–29	1.268*	0.009*
		30–39	1.834*	0.000*
		40–49	0.922*	0.048*
Comprehensibility	>50	10–19	1.011*	0.014*
		20–29	1.599*	0.000*
		30–39	1.646*	0.000*
		40–49	0.937*	0.025*
Learnability	>50	10–19	0.748	0.118
		20–29	1.324*	0.009*
		30–39	1.448*	0.004*
		40–49	0.515	0.288
Helpfulness	>50	10–19	0.705	0.087
		20–29	1.170*	0.007*
		30–39	1.633*	0.000*
		40–49	1.034*	0.014*

* shows significance.

Table 6. Significance on "portability" of guide map.

Question	Mean	SD	P value
Habit	5.404	1.871	0.338
Convenience	5.692	1.615	0.499
Size	5.173	1.605	0.267
Collection	4.192	2.327	0.615
Intention	4.442	2.043	0.594

relevant to the introduction of museum exhibition, although wayfinding information is still concluded as the most important. Referring to the result shows in Table 1, in which question "getting lost" is lower scored comparing to the others, it is clear that wayfinding information of current guide maps is not enough and needs to be improved. Additionally, as mentioned, 30–39 year-old users would prefer more educational information, e.g. more knowledge about the exhibition should be presented on the guide.

Preferable design on aesthetics across the age are various. For young visitors (age 10–19 and 20–29), exhibits of digital simulated 3D image are attractive to them, while visitors age over 30 prefer real exhibits photos. More participants like simple and soft toned color rather than rich and colorful style, however, elders hold a contrarily opinion. This result matches the finding in the aesthetic evaluation of current guide, age over 50 rates clearly higher than the other age groups, particular on the "color" and "graphics" questions, as the guide maps of Taipei Zoo and National Palace Museum are both design in rich color.

Participants accordingly indicate that the floor plan should be presented through 3D illustration. This will prevent users from confusion when they are trying to transform their position to a 2D floor plan. Additionally, some participants indicate that they may suffer problems when trying to understand the meanings of the icons. They suggest that more explanation can be added to the icon, accompanying with different colors to distinguish the areas on the floor plan. In such way visitors can learn the map efficiently and find their way out easily.

Regarding the portability of guide map design, visitors have the habit of picking up guide maps at the start and referring to information all the way through. That is to say, the guide map should be portable and it should be folded to a hand carry size—better be smaller than A4. However, the guide map is still a one-off document for the visitor. Concerning sustainability issue, more delicate or special design which can trigger users to keep it after could be considered.

4 CONCLUSION

This study demonstrates the effect of age on some attributes of guide map design. Different age visitors have different requirements on the use of guide map. Clear difference can be particularly found on the usability and aesthetics, for age group over 50. Additionally, the results show similar scores between teenagers and elders. This result agree with previous outcome "the difficulty of guide map information should meet the understanding level of junior high school, if no visiting group is focused for a museum" (Tsau & Chu, 1991). This will meet most visitors' requirements when using guide maps. Future work can be focused on the investigation of the effect of different types of museum on guide map design.

REFERENCES

Chung, K.-W. 2004. The study on visual design in network of museum—In Example of Creative Design of "Betou Hot Spring Museum". Master Thesis of National Taiwan Normal University, Taipei. (in Chinese).

Ciavarella, C. & Paternò, F. 2004. The design of a handheld, location-aware guide for indoor environments. In *Personal and Ubiquitous Computing* 8(2). Springer-Verlag.

Damala, A. & Kockelkorn, H. 2006. A taxonomy for the evaluation of mobile museum guides. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* MobileHCI '06, ACM Press. 273–274.

Fan, C.-W. 1996. *A semantic-cognition oriented computer aided design method for science museum media exhibition*. Master Thesis of National Cheng Kung University, Tainan. (in Chinese).

Herrmann, C., Kawalek, J. & Stark, A. 2005. User-centred system design: addressing user needs with a mobile information system. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* MobileHCI '05, ACM Press. 335–336.

Huang, Y.-J. 2005. *The Application of metaphor and pleasureability in the webpage design of digital design museum*. Master Thesis of National Taiwan University of Science and Technology, Taipei. (in Chinese).

Hung, H.-T. 2005. *Design and application of a game-based mobile guiding system*. Master Thesis of National Taiwan Normal University, Taipei. (in Chinese).

Hsieh, Y.-H. 2000. *A study on the visual information system of the outdoor quasi-museum*. Master Thesis of National Yunlin University of Science and Technology, Yunlin. (in Chinese).

Hsieh, Y.-J. 2000. *A study of demonstrative sign—Travel map*. Master Thesis of National Cheng Kung University, Tainan. (in Chinese).

Jesus, R., Dias, R., Frias, R. & Correia, N. 2007. Identifying locations and GIR applications: Geographic image retrieval in mobile guides. In *Proceedings of the 4th ACM workshop on Geographical information retrieval* GIR '07, ACM Press. 37–38.

Kao, C.-Y. 2003. *A study on the operation of museum website interface for different groups—A case study of the National Museum of Natural Science*. Master Thesis of National Yunlin University of Science and Technology, Yunlin. (in Chinese).

Kusunoki, F., Satoh, I., Mizoguchi, H. & Inagaki, S. 2007. SoundSpot: a next-generation audio-guide system for museums. In *Proceedings of the international conference on Advances in computer entertainment technology* ACE '07, ACM Press. 272–273.

Lin, C.-W. 2002. *A study of icons design for the web browser—Case design of digital museum iconic*. Master Thesis of Dayeh University, Changhua. (in Chinese).

Lin, Y.-M. 2003. *A study of effectiveness of interactive multimedia display at the National Museum of Natural Science*. Master Thesis of National Yunlin University of Science and Technology, Yunlin. (in Chinese).

Liu, C.-J. 2001. *A study of in-house information service systems usability—A case study of the National Museum of Natural Science*. Master Thesis of Tunghai University, Taichung. (in Chinese).

Lo, K.-C. 2004. *The research of map visual communication design and application—Take the China ancient maps and modern amusement park maps as the examples.* Master Thesis of National Yunlin University of Science and Technology, Yunlin. (in Chinese).

Lu, C.-F. 2004. *A study of communication on illustration and text of explanatory labels in "The Physical World" at The National Museum of Natural Science*. Master Thesis of National Yunlin University of Science and Technology, Yunlin. (in Chinese).

Mantyjarvi, J., Paternò, F., Salvador, Z. & Santoro, C. 2006. Scan and tilt: towards natural interaction for mobile museum guides. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* MobileHCI '06, ACM Press. 191–194.

Raptis, D., Tselios, N. & Avouris, N. 2005. Context-based design of mobile applications for museums: a survey of existing practices. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* MobileHCI '05, ACM Press. 153–160.

Tsau, C. & Chu, N.-T. 1991. *Taipei children traffic museum development project.* Taichung: Tunghai University Landscape Architecture Research Center, 14. (in Chinese).

Yi, P. 2003. *The navigation design in virtual design museum*. Master Thesis of National Cheng Kung University, Tainan. (in Chinese).

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Ergonomics challenges in sheltered workshops

Stephen J. Morrissey
Ergonomic-Human Factors Engineering Consultant, Oregon OSHA Consultative Services, Portland, Oregon, USA

ABSTRACT: This paper describes the activities undertaken in developing ergonomics interventions for two different sheltered workshops. Although the initial requests for assistance were to help with ergonomics, the projects eventually involved safety, health, industrial engineering, psychology as well as ergonomics. These projects reflect on the wide range of issues that need to be addressed in many actual interventions.

Keywords: ergonomics, safety, health, sheltered workshop

1 INTRODUCTION

1.1 *Definitions*

A sheltered workshop is an organized work environment where adults with varying degrees of mental handicaps (clients) can perform productive work under the supervision and control of trained persons. This work activity helps in some forms of treatment, provides them with socialization opportunities and lowers the costs associated with caring for the individuals. The work done in these workshops often is very tedious, simple and repetitive. It is also potentially hazardous from safety, health and ergonomic perspectives. A further complication in these environments is that clients may not be able to verbalize or describe pain and discomfort they are experiencing from work activities and may not understand the importance of following specific procedures. In addition, staff or contract agents who get work may not have any understanding of basic safety and health issues in work being done much less of potential issues in ergonomics.

While individual clients have potential for exposures to hazards, for the staff the hazard exposures may be greater as they have to work with the clients, often in very poor postures and often the most rudimentary safety/health protection beyond safety glasses or hearing protection.

Oregon OSHA's Consultation group offers free, non-regulatory and confidential consultation in the areas of safety, health and ergonomics and it was contacted by the two different centers to address "ergonomics" questions or concerns at each of the facilities. The following discussion describes how this was approached using an integrated ergonomics as well as safety, health and management model.

2 CENTER ONE

2.1 *Description*

This center was located in a large warehouse that had been converted to work cells that completed various tasks/orders as local and regional industries sub-contracted out work to the center rather than do it themselves or export it. There could be several different cells and product runs in operation at the same time. The center tries to keep itself open and the clients employee as many days a year as possible. As a result, clients may have to be moved from one project to another, a process that may require re-training. A few of the tasks or contract jobs that were studied over a two year period and associated organization, training, safety and ergonomic risk factors identified included.

2.2 *Making six-packs of small production run or specialty beers*

This had the clients taking one bottle of beer from each of six different cases and then placing each bottle into one package (six-pack). The empty packs were placed at about knee height or on a 1 meter high table and the cases of different style beers were next to the work station at knee height. After a number of six-packs had been created they were then put into a case by the supervisor or worker and moved to another area where they would be palletized, shrink wrapped by hand and then moved to a collection/shipping area.

Risk factors identified:
- Highly repetitive work with hands above mid-chest height and deviated wrists.

- Work with hands below waist level leading to bent postures when packing and when picking up bottles.
- Poor overall workstation design for work being done.
- Significant amounts of work in process transport, hand wrapping by leads or staff.

2.3 *Creating variety cases (24-bottles) of wine*

This task had the clients taking the required number of bottles from different cases usually one at a time and inserting them into the final case to create cases with two, three or four different types of wine. The case was about 33 cm tall and on top of a 1 m high table. The flaps of the box were usually up, creating an addition lift over barrier. Wine cases were on a pallet next to the assembly area. After the case was full, the same client might seal the case and build another or move it to another location where the supervisor would double check and seal the case. A 14 cm high platform had been built in front of the workstation to aid packing.

Risk factors identified:
- Highly repetitive placing of bottle into cases using very poor hand postures.
- Hands at or above shoulder height much of the time.
- Poor overall workstation design.
- Long reaches to cases on ground.

2.4 *Creating conference registration and welcome packages*

This was done by having the client pick up individual sheets of paper or gifts or handouts from piles or boxes of different flyers placed on card tables and placing them into plastic bags for distribution at the meeting or conference.

Risk factors identified:
- Potential for stumbling from everyone walking around the same tables.

2.5 *Labeling wiring harnesses for large trucks and wiring dashboards for trucks*

These tasks were almost exclusively done on tables where clients sat or stood. Some work involved manually following wires and installing labels on the wires or dashboards, other involved using a press to install wire connectors on the ends of wires. Finished work was collected, put in boxes and taken to be palletized and shrink wrapped.

Risk factors identified:
- Several tables were at a poor height for the work done and/or the size of the clients creating poor torso and hand postures.
- Work flows often were congested and created frustration for the supervising staff given their need to supervise/coach individual workers, inspect final work and insure proper transport to the next station.

3 RECOMMENDATIONS AND INTERVENTIONS

3.1 *Discussion*

Over the course of several visits with the center director, the previously noted work activities were observed and discussions were held with floor supervisors and work cell supervisors on what I was observing, my concerns and immediate recommendations to address these issues. In addition to identification of a variety of significant ergonomic issues in work station design, work flow and work postures that were discussed at the time of the visit, these discussions revealed there was a (complete) lack of understanding of ergonomics work principles in the staff and center directors (beyond don't lift too much) and that while very well meaning in their intentions, work had often been set up with what was available (or had been donated), in the most available space and many of the cell supervisors were very overworked.

3.2 *Recommendations*

To address these concerns the following activities/recommendations were made and undertaken.

1. The center director, leads, and supervisory personnel were provided with very basic training on body mechanics, fundamentals of workstation design, work flows and "good working postures".
   a. A suggestion was made to have a workshop on these issues with all staff and with those involved in looking for work.
2. There must be an on-going process to modify workstations based on what is actually being done and observed (on the fly ergonomics).
   a. This was being done after about six-months with varying degrees of success.
3. An integrated materials handling/work flow study needs to be done for any project undertaken.
   a. This has not yet been accomplished due to lack of time, however the center was able to acquire an automated pallet shrink wrap machine and fore and aft conveyors that allowed pallets to be pushed through the system rather than be handled and wrapped by hand.
4. Basic work methods incorporating proper ergonomics methods must be developed and validated before clients are trained.

5. There should be consideration to development of a standard set of work layouts to minimize difficulties in retraining clients.
6. Staff working with the clients must have additional training in recognizing and correcting their own poor work postures and habits as well as how to recognize external signs of musculoskeletal injury that may be expressed by clients.
7. Safety and health awareness must be incorporated into the training of lead and supervisory personnel.
8. The center director must have awareness training and the center sponsors must have greater understanding of both the costs and the benefits of ergonomics and a willingness to both support changes based on ergonomics and in their marketing of the shelter's capacities to potential contract suppliers, be able to ask critical questions on how work can be improved in the shelter with the sponsor's assistance.

4 CENTER TWO

4.1 *Description of facility*

The second workshop to be discussed was a small wood products facility that made a variety of wooden products for industry including pallets, large boxes and some basic table tops. Of these product lines, the wooden pallet fabrication line had the greatest overall production and was the highest concern to the center director and floor manager due to the high volume of work and "issues" with the clients not following directions in work flows, work methods. A parallel process where larger wooden containers were constructed was also of interest for review by the center's staff for "work flow questions".

4.1.1 *Pallet line*

The clients working in this area were adults from 20 to 50 in age, many of whom had worked in this center on the same line for up to 20 years. Equipment used by the clients in the pallet construction line included pneumatic nail guns, hammers, metal strapping tools and pallet jacks. Work began with pre-cut wooden pieces being delivered by pallet to the assembly area where the clients took the individual pieces from the pallets, arranged them on the worksurface then used the pneumatic nail gun and hammers to assemble the pallet. The assembled pallet was then placed on a pallet jack next to the work station and when the pallet was loaded, it was taken to the loading dock where another client secured the pallets with metal strapping tape and the pallet was then pushed into a truck or storage area. The issues of concern expressed by the lead personnel were that workers were no longer following specific instructions to put finished work in specific areas, were adopting poor working postures when doing this and were creating congestion and tripping hazards.

In reviewing the operations, there had been considerable effort made to create a "good" workplace in terms of work surface heights and work flows. The issues identified were that workers were now choosing to place finished pallets in an area used as a passageway. In discussion with the leads, we determined that this behavior began after a bench seat had been moved several feet to allow a faster resupply of materials.

4.1.2 *Solution*

Move the bench back to its original position. While not an ergonomics issue per se, it does reflect that work system design is not a linear or always understood concept.

4.2 *Wooden box construction*

In this area, standard size and custom size wooden shipping containers were made with staff doing cutting of materials, the clients assembling them and moving them to a shipment point. In review of the operations it was observed that work was often done on the floor, necessitating long reaches and much bending; once boxes were finished they were lifted by hand and carried about 3 meters to a staging area by two clients. Some of the boxes were large and quite heavy, about 25 to 35 kg and it was not clear that the workers some smaller, young women working with larger men understood proper lifting or lifting safety. It was also apparent that there was definite competition between the men and women in how hard they could work. The floor and walkway was also congested with air hoses and debris. In discussion with the lead for this area it was determined that they were overworked and were only concerned with power saw safety which was what they assumed was the hazard in the area.

4.3 *Recommendations and interventions*

In discussion with the leads, it was determined that they did not really understand occupational safety issues, that they did not understand the importance of proper materials handling and work station design, and importantly, they did not understand, nor did they recognize, the potentially hazardous outcomes of competition in such a work environment. As such the following recommendations were made:

1. A complete revision of the floor plan needed to be made to insure safe, linear and clean walkways.

2. Work tables needed to be made to put the work at an appropriate height to eliminate bending.
3. There should be no manual handling of large boxes or heavy boxes, these must be moved by carts with loads slid or team lifted onto the cart.
4. There must be some means to try and limit the competition between the workers; this was a more difficult issue as it was not clear the clients even understood what competition was.
5. Leads must have immediate and ongoing training in occupational health, safety and ergonomics training appropriate for their position.
6. Outside safety consultants must be brought in on a regular basis to monitor results. As the State of Oregon provides this as a free service, it needs only to be requested.

5 DISCUSSION AND CONCLUSIONS

Ergonomics interventions can be very straight forward at times, and when these are requested the individual making the request may not understand all the issues that may need to be addressed. In many environments one must be alert to issues in safety, health, industrial engineering and work flows and psychology and group interactions.

In the very challenging world of the sheltered workshop, highly dedicated people were placing themselves and their clients at potential risk because of a lack of understanding of all of these issues. The interventions recommended and the changes made on the floor during the visits served to address these issues.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Discussion of emergency notification device for elderly people

Nanae Shintani
Gifu University of Medical Science, Seki, Japan

Hiroyasu Nagaoka
Graduate School of IPS, Waseda University, Kitakyusyu, Japan

Kazuko Horiguchi
Kobe University School of Medicine Faculty of Health Sciences, Kobe, Japan

Tomoko Morimoto
Gifu University of Medical Science, Seki, Japan

ABSTRACT: As the evolution of an aging society, the number of elderly people who live alone or in a family of two is increasing. Each municipality has a lending program of a home-use emergency notification device (henceforth abbreviated as "device") for elderly people to realize a society of safe living by reducing anxiety in case of emergency. This is a system to inform of any emergency situations by pressing a button on the device to help rapid rescue response. This aims to preserve the security and improvement in QOL by reducing anxiety in elderly people. This device ranks high on elderly people requirements in the area of telecommunication at a survey of the Ministry of Posts and Telecommunications. It is expected that there will be more expansion of demand to the aging society but the diffusion rate is actually low. The purpose of this study is to develop an easy-to-use device for elderly people and to improve operation procedures by clarifying the device usage at an interview in H-city. As a result, it was clarified that there is room for improvement on the operation comprehension and the device shape. In order to utilize the device more usefully, it is clarified that a lecture for understanding device operation should be done together with an acquaintance and continuously, the shape should be considered portability because there were a lot of elderly people with independence in ADL, and the device can be used away from home. In the future, it is required to establish overall services to secure elderly people by introducing ubiquitous tools. It is clarified that automated information tools which can detect and report any emergency situations without the device holders' intention. It is suggested that a guideline to allow monitoring with privacy consideration should be established and receive social consensus, a system to prevent the detected privacy data from being concealed should be developed, and an algorithm to detect emergency situations by checking the attitude and behaviour of elderly people should be established.

Keywords: emergency notification device, elderly, interview

1 INTRODUCTION

The number of people aged 65 and over in our country became 25.67 million people which are more than 20% of the total demographic shifts (127.77 million people) in 2005, so Japan became a country of super-aging society. Subsequently, it is predicted that it will become 35% of the total demographic shifts (89.93 million people) in 2055 and at the national population census in 2005, the ratio of people who live alone or family of two increased consequently due to the rapid lifespan expansion and low birthrate. To help the realization of safe living society by reducing anxiety of increasing elderly people year by year who live alone in case of emergency, each municipality has a lending program of the device for elderly people aged 65 and over who live alone or only elderly people. This is a system to inform a rescue squad of any emergency situations via telephone line to help rapid rescue activities by lending the device to elderly people who live alone or only elderly people. This aims to preserve the security and improvement in QOL by reducing anxiety in elderly people. After launching the care insurance system in 2002, the usage of this kind of system increased because nursing care managers introduced it when they performed a home visiting survey, and aggressively installed mainly at the nationwide municipalities. However, the actual number of people who receive the service throughout Japan is the low diffusion rate of only 6% (approx. 135,000 people) among the high priority elderly people who live alone. On the other hand, this device ranks high on the elderly people requirements in the area of telecommunication at a survey of the Ministry of Posts and Telecommunications. It is expected that there will be further expansion as the evolution of an aging society.

The population of H-city which is the target of this survey was 177,931 people (in September 2008) including 33,080 people (18.6%) of age 65 and over. There were 73,536 households including 6,908 solitary households aged 65 and over, and 5,779 households of only elderly people. There were 515 households (4.1%) who have the device at the end of March 2008.

However, there were a lot of cases that the device could not be used efficiently in emergency situations in spite of lending the device but there are few other studies. Therefore, the purpose of this study is to offer a basic data to improve functions and usability of the device for elderly people by clarifying the actual usage and conditions.

2 METHODS

2.1 *Target and method*

The research was conducted by interviewing 20 elderly people living in H-city who borrow the device. A pre-survey were conducted for the survey contents and items, and then reviewed the survey contents for difficult questions and items. The survey was conducted between March and August 2006.

2.2 *Survey contents*

2.2.1 *Characteristics of target*

There were 15 question items of basic attributes for survey targets which were: age, gender, family cohabitation, current medical treatment condition, current disease, eyesight, hearing ability, degree of dementia, degree of nursing care, possession of physical disability certificate, contents and utilization of in-house care service with/without using care insurance, and social situation.

2.2.2 *Comprehension of device operation*

There were 9 question items which were: installation time, installation circumstance, who lectured, degree of relieving, comprehension degree of device operation, comprehension degree of mobile pendant operation, situations to use the device, cases to think of using the device or not, and number of actual usage.

2.2.3 *Device shape and communication method*

There were 6 question items which were: installation location of the device, reasons for choosing the location, place of situating the mobile pendant, hand-carry status of mobile pendant, easiness of using device, and degree of power to press the button of the device.

2.3 *Ethical considerations*

After explaining purpose of the survey to a district welfare officers in an area, they explained to the target people. Then the survey was conducted by face-to-face interview to people who accept the survey. In addition, the following conditions were explained: answers are anonymous, data will not be used for other purpose than this study, the interviewed people can interrupt anytime when their condition become not good, and do not need to reply any questions if they do not want to reply.

2.4 *Analysis method*

Each item was summarized simply.

3 RESULTS

The results of "Characteristic of targets", "Comprehension condition of the device operation" and "Shape and communication method of the device" are described with variables below.

Total area of the target H-city is 635 sq km. There are three different districts which are: academic district with one university and two colleges, research and agricultural promotion. Since there are a lot of student residents, people who relate to university and production, the population aging rate is 17.7% which is 2.9 points lower than the nationwide average. This study was conducted at the agricultural promotion district where the population aging rate

Table 1. Characteristics of H-city.

Item	Number
Total population (people)	177,245
Population aged 65 and over (people)	31,386
Population aging rate (%)	17,7
National average of population aging rate (%)	20,6
Total number of households (household)	74,928
Family of two aged 65 and over (household)	6,908
People aged 65 and over who live alone (household)	5,779
Number of lending emergency notification devices	573

is high and intercommunion with neighborhood remains.

A majority of the target people was 19 females (95%). Regarding communication amount with family lived separately, nine people (45%) met each other more than once a month and eight people (40%) communicated by phone. Regarding communication with neighborhood, 17 people (85%) kept in touch just when they have troubles and they wanted to chats.

Regarding an explanation at the installation time, 15 people (50%) got explanation from the installation vendor. Two third of the people could understood but one third of the people understood incorrectly or could not understand completely. Regarding the question about cases to think of using the device or not, two people (10%) replied "Yes", 17 people (85%) replied "No", and three people used actually.

Most of the installation place of main device was a living room or a bed room. About carrying the mobile pendant, 16 people (80%) did not carry at all. Regarding ease-of-use, positive answers of "easy to use" and "so-so" were from 14 people (70%), and negative answer of "neither" was from three people (15%). Regarding notification, a half of the people replied "easy to use" and "no problem" because it is just an operation of pressing a button.

4 CONSIDERATIONS

In this section, current status of the device was summarized and described the reason why the device is not fully utilized from the results by considering aspects of "Target characteristics" (Table 2), "Comprehension status of the device" (Table 3), and "Shape and Communication method" (Table 4), then we mention a suggestion at the end.

About current status of the device, there were three examples of actual usage as follows: "when a person broke a bone by falling down", "when a person who has mental problem felt fears" and "False report that a pet pressed a button".

On the other hand, there are 17 people (85%) who did not think of using the device. The reasons are: "There are no experiences of emergency situations", "There was a case but did not use it because of annoyance of neighbors when an ambulance comes", and "Relatives and neighbors are more convenient than the device". It was found that the major reasons why people do not want to use the device are: people are confident of their current health and people have a feeling to reject d an ambulance which will come after using the device. There was an elderly person in this 17 people who fainted on the road in front of the house and found by a neighbor and then sent to a hospital. We asked the elderly person if the person uses the device when the same things happened to the person. The answer is that "the person knows own health condition and found it beforehand but it could not be realized since it became bad suddenly". So there was no aggressive intention to use the device for that person. In this sense, it can be seen that people do not carry a mobile pendant.

Table 2. Characteristics of target.

Item	Quantity (n = 20 people)
Age	81.44 ± SD 8.787 Under 60 (1 person), 60–70 (1 person) 70–80 (4 people), 80–90 (9 people) Over 90 (3 people), Unclear (2 people)
Gender	Female (19 people), Male (1 person)
Cohabitation with families	Alone (17 people), Family of 2 (1 person) Others (2 person)
Dedgree of nursing requirement	Independent (12 people), Need nursing (5 people) Need nursing degree 2 (1 person), Unclear (2 people)
Daily life independence of elderly dementia people	Independent (16 people), I (4 people)
Periodical medical check	Yes (20 people)
Current disease	Affected two diseases in average
Meeting with separated family	A couple of times a month (9 people)
Interaction with neighbors	Can consult when people have problems (8 people) Exchange souvenirs (1 person) Can do small talks (8 people)
Usage of the home servie of nursing insurance	Ambulatory rehabilitation (4 people)
Usage of the home service of other than nursing insurance	Home-helper (3 people)

Table 3. Comprehension status of emergency notification device.

	n = 20 people	
Item	Category	Qty
From whom people get explanation of usage	Municipal employee	1
	Case worker	0
	Care manager	0
	Employee of vendor	15
	Others	3
	Unclear	1
Usage of emergency notification device	Place to push correct	16
	Effect after pressing correct	14
	How to stop announcement after pressing a button correct	10
Usage of mobile pendant	Correct	14
In what situation people use the emergency notification device	Severe pain	7
	No appetite	1
	Feel bad a little	2
	Cannot get up from futon/bed	4
	Too Painful	13
	Run out of kerosene	1
	Feel lonely	1
	At a fire	11
	Cannot move due to falling down	11
	When carring a heavy things	1
Are there any cases to try to use the emergency notification device?	Yes	2
	None	17
	Unclear	1
Number of actually used the emergency notification device	1 time	1
	2 times	1
	More than 3 times	1
	None	17

To summarize above, the case of using device is limited only when a person knows the situation is critical like a threat of life. In other words, people might not use the device if they can not realize the situation is critical even if there is a device nearby.

Next, as you can understand from Table 1, the daily life independence in ADL and dementia among 16 people (80%) were independent. To view the social condition of the target, meetings with family lived separately were more than once a month and phone communication was also more than once a month. In addition, regarding the communication with neighborhood, 17 people (85%) keep in touch just when they have troubles and just for chatting. They have ability to act and communicate because the ADL of the target is almost independent, and there have active relationships with neighbors because the targeted area is agricultural district and continued over many generations. It was considered that these factors improve social conditions.

Regarding the usage status of the nursing care insurance, a half of the certified people of the service used ambulatory rehabilitation. In other housebound services than the nursing care insurance, it was also found that the home helper service is common because elderly people who live alone look forward to waiting for a house helper who does both works to prepare meals and talk to them. According to the "target characteristics" obtained from survey, it is noted that the reason why the survey target did not use the device often is that there are few people who encountered emergency situations because of the high independence degree. The reasons for the low usage that people did not use it when they should use is considered that: communications with neighbors give them a confidence to live alone and the feeling under appropriate control of others by using the nursing care insurance provides a feeling of safety. It is considered that these prevent the mobile pendant from carrying and weaken the existence of the device.

Table 4. Shape and communication method.

	n = 20 people	
Item	Category	Qty
Installation location	Living room	5
	Bed room	3
	Entrance	4
	Others	4
	Unclear	4
Reason of choosing the installation location	Where people stay every time	13
	Where the installation person recommended	3
	Others	4
Usage status of a mobile pendant	Carry every time	0
	Relaxed in a room	0
	During in a toilet	0
	While bathing	0
	While shopping	0
	Go out for a friend's house near by	1
	While sleeping	1
	Never Carried	16
	No answer	2
Easiness to use the emergency notification device	Easy	10
	So-so easy	4
	Nether	3
	Not so easy	0
	Difficult	0
	No answer	3

About device operation comprehension, a half of the people could understand but four people (80%) could not understand entirely where they should press. About the question after pressing a button, 14 people (70%) could understand but six people who understood incorrectly or could not understand entirely are the problem and these six people hardly recognized the installation of device. About the explanation at the installation time, 15 people (75%) got it from an employee of the installation vendor. To consider about a characteristics of elderly people, they have weaken ability to adopt new things. So, it is necessary to give additional explanations and attendance of a person familiar to them like care managers for the explanation of unaccustomed devices, and it is considered that this way will be easier for them to understand and ask questions.

About a question "at what specific situation do you use the device" to the survey target, the answers are: "When it is too hard", "When they cannot stand up after falling down" and "When it is too painful". So, it is found that people only think to use the device when they become severe conditions and put up with to some extent because they do not want to bother others. To this endurable and thoughtful thinking, it is necessary to give explanations to understand: consideration of carrying method which does not bother neighbors as much as possible, the mindset that hesitation to press a button will cause a delay of treatments and bother their family and neighbors, and what kind of people will support after notifying by the device.

Thus, it is a good thing that chances of using device is a little during the daily life of elderly people but it is found that the acknowledgement of device is lowering and fading from consciousness of elderly people due to the low usage. When visiting to the elderly people's house, it was considered that continuation of mindset is necessary that emergency status can be occurred at anytime and efficient use of the device is necessary as well as reconfirmation of the device operation. Furthermore, about the thinking of that "I know my body condition very well" by an elderly person who fainted on a road without using the device and saved life by neighbor's report, it is required to understand that there is a possibility of sudden changes in body condition which differ from that of their young age even if they know their body condition for a long time. It is necessary to think about the characteristic of elderly people from the "problems of education and comprehension", and it was considered to be able to establish conditions to utilize the device securely by adjusting these problems.

Also other than the educational approach, it is suggested that an automated notification device which reports without elderly people's intension is required.

Since there are various installation places of the device, most of the people chose "a place where they always stay". Next, there is an answer that "They were told that it is the best place for wiring by an installation vendor". There will be a case that the installation place is limited for wiring for old houses in rural areas but due to the deterioration in hearing, it was considered that it should be installed to a room where they spend most of the time by improving the device. Regarding the possession status of the pendant, 16 people (80%) replied they do not carry it at anytime and there were some cases that they could not cope with their bad body condition, falling down and broken bones around outside of their houses. The reasons gotten at the interview survey for not carrying the pendant were considered that: "They know their own body condition" as previously mentioned and problems of device itself. There were following device problems: "It is difficult for working", "Hanging from the neck is dangerous", "Anxious if it misreports even when the button is touched softly". There were a lot of targets who were independent in ADL and a lot of elderly people who went out to visit neighbors or engaged in gardening and field works outside. Thus, the pendant type is not suitable for elderly people who are independent in ADL and work especially by using hands very much, so it was considered to incur danger due to blocking and hooking. Therefore, it is found that more than two third of people did not carry it. Recently, there is a watch type notification switch but no one had this in the survey targets. However, the way of notification is just an action of pressing a button, so it is considered easy and there are no problems.

In addition, since the communication method between the pendant and the main part is weak radio wave method, the available working areas are just inside of the house or around 50 m from house without any obstructions. So, it was considered insecure because all areas usually elderly people act are not fully covered. For "the shape and communication method of the device", it was requested to develop a shape to use easily for elderly people who is independent in ADL and a system which can notify anytime anywhere when they encounter bad health conditions.

There are a lot of elderly people who feel anxiety about their health condition as the deterioration of body function according to the aging. Furthermore if family members are not close to them, it is difficult to act rapidly or to confirm the safety at emergency situations like illness or injuries from falling down. Therefore, the need to develop a system to support elderly people at emergency situations is high. The emergency notification device in this study is based on reporting by elderly people by themselves but a system to detect and notify emergency situations without awareness of elderly people is becoming a practical use stage from a researching stage recently. For example, about a method to confirm safety from daily routine works of elderly people, there is the i-pot of Zojirushi Corporation which became popular to monitor the daily life of elderly people. The usage of a pot for boiling water can confirm the safety of elderly people. A communication function built into a pot for boiling water send the information to their family in a distant place. Furthermore, there is a system to sense in-house activities of elderly people who live alone and detect any movement to any rooms. In the past, these systems can just offer activity data to a family to find a chance to notify abnormal conditions. Recently, an automated emergency notification system is studying, which compares non-response times of censors and moving status with normal status data by using the DP matching method or the clustering method and notifies when it judges them different.

Furthermore, there is a system to get health conditions of elderly people directly. It detects health conditions in real time by attaching censors to get vital sign information. However, measuring of data is limited. In case of a device attaching to a body, daily activates are limited by the device. Therefore, it has not become widespread yet. Besides, there are some ingenious attempts of shapes which are a watch type and a roll type on the chest. Furthermore, the movement limitation can be decreased by using a wireless transmission of zigbee or cell phone functions. So there is a possibility that it will be widespread hereafter at once.

The example so far is a limited service basically in the house or areas close to the house. Despite elderly people, a lot of people go out away from their home. Moreover, there are quite a lot of cases that elderly people who tend to hang around or go to places where the family member cannot expect.

A service focused to this, there is also a current location notification service by using a cell phone. Recently, this put into practical use rapidly mainly for the children monitoring system. However other services than using mobile communication terminals or RFID can not clearly identify targets or can detect only a small part of activities. This relates deeply to a background that the problem of privacy violation could not be resolved for a system like monitoring independent person using a camera. This problem is arguing even when setting up a monitoring camera for children. Basically, a monitoring including privacy information should be used by limited individuals and allowed according to a guideline reached consensuses socially, but actually various systems are trying without any guidelines now.

Next, follow-on topics are described to think of these statuses as follows:

5 FOLLOW-ON TOPICS

The following measures are necessary to promote usage of the emergency notification device and make it a more useful system.

5.1 *Education and support by understanding characteristics of elderly people*

Education should be done contiguously with measures like accompanying an acquaintance. So supporting organizations by communities and whole society like neighbors and a residents' association are important. Moreover, it is necessary to establish an atmosphere to press the emergency button freely and do not think about bothering neighbors at emergency situations.

5.2 *Device shape and communication method suitable for elderly people*

Since there are a lot of elderly people who are independent in ADL, it is necessary that the shape should not be a burden to carry and it can communicate away from home.

In the future, it is required that overall services to secure safety for elderly people should be established by combining a ubiquitous system which is developed rapidly. An automated system to judge and inform of emergency statuses without person's intension anytime, anywhere is mandatory. A function to judge "emergency status" objectively is also needed and it is required to get vital sign data for elderly people themselves in addition to a function to catch surrounded circumstances of elderly people. However since the collected data contains sensitive personal information, a social consensus is needed for the data collection and data processing. Therefore, the following three points are proposed:

1. Establish a guideline to allow monitoring including privacy data and get social consensus.
2. Develop a system to conceal detected data related to privacy.
3. Establish an algorism to judge emergency status from poses and moving patterns.

REFERENCES

ITPT 1999. Utilization trend for disabled people and elderly people, *Post laboratory monthly report* (January), Institute for Posts and Telecommunications Policy (ITPT). Retrieve from *http://www.umin.ac.jp/choju/abstract.htm*

MHLW 2006. National population census in 2006: The first basic summary results. Ministry of Health, Labor and Welfare (MHLW), Japan. Retrieve from *http://www.mhlw.go.jp/shingi/2006/11/dl/s1114–5b.pdf*

Miyazaki, S. 2002. Development of emergency notification system for elderly people who live alone by using the pulse wave—basic consideration. *Biomechanics* 6(1): 32–39.

MPTCPB 1995. *Aging society and multimedia to realize information longevity society*. Ministry of Posts and Telecommunications, Communications Policy Bureau (MPTCPB), Information Management Division, New Japan law publication.

Shinagawa, Y., Kishimoto, T. et al. 2006. Development of automated notification algorism by using non-response time of infrared sensors. *Kawasaki medical care welfare journal* 15(2): 553–563.

Zojirushi Corp. 2001, Mimamori hotline. Zojirushi Corporation. Retrieve from *http://www.mimamori.net/*

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Cognitive aging and determine of an instructional media type on procedural task learning for elderly people

Ding-Hau Huang & Wen-Ko Chiou
Chang Gung University, Tao Yuan, Taiwan

ABSTRACT: It is well-known that cognitive factors decline with age which due to older people tends to have more difficulty learning compliable. Although some researchers indicated that Cognitive Load Theory (CLT)-based training formats meet the cognitive abilities of elderly learners particularly well, the mapping of instructional design principle and their combination of procedural instruction for elder learner is still not clear. Therefore, the aim of this article is to review the literature on how properly design principle and what type of instruction can be better support elderly people in completing procedural task. To conclude, our literature review shows that the key to optimal procedural learning in the elderly might be determined as combination of dynamic and static visualizations by using the cognitive related theory to supporting skill acquisition.

Keywords: cognitive aging, cognitive load, procedural learning, instructional design, multimedia

1 INTRODUCTION

Different modalities of information such as pictures, drawings or dynamic videos are often designed to facilitate learning and instruction and have been found to have a positive effect. It's because these instructions were designed by cognitive based theory such as Cognitive Load Theory (CLT) (Sweller, 1988) and cognitive multimedia theory (CTML) (Mayer, et al., 2001), CLT and CTML based on a cognitive architecture that consists of a limited working memory, can provide guidelines to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance. (Kirschner, 2002; Paas, et al., 2004; van Gerven, et al., 2006; Wilson & Wolf, 2009; Wong, et al., 2009).

Many studies also have explored the effectiveness of presenting procedural learning information through comparing different media and their combinations (Arguel & Jamet, 2009; Brunye, et al., 2006; Höffler & Leutner, 2007; Wong, et al., 2009). The topic which comparing the effectiveness of static versus dynamic visualizations has been more concerned (Ayres & van Gog, 2009), but continues to result in seemingly contradictory findings unless under certain conditions. In particular, the effect was greater when the animation was representational, when it was highly realistic, and when procedural-motor knowledge was to be acquired (Höffler & Leutner, 2007).

However, from the age point of view have less attention in this context. It is well-known that physical, sensory, and cognitive factors decline with age (Craik & Salthouse, 2000; Scialfa, et al., 2004). These declines precipitate after individuals reach their mid-forties which due to older people tends to have more difficulty learning compliable (Kang & Yoon, 2008) and procedural tasks (Vakil & Agmon-Ashkenazi, 1997). It is because these tasks are concerned with cognitive learning and retention and be affected by cognitive aging. Thereby, some researchers started to connect cognitive aging and CLT. For instance, van Gerven, et al. (2002) claimed that CLT-based training formats meet the cognitive abilities of elderly learners particularly well. That is, cognitive aging brings about several declines of working memory, which impede the acquisition of complex cognitive skills. By making an optimal use of the 'remaining' cognitive resources, learning can be enhanced. They also indicated that existing principles of general instructional and multimedia design can be evaluated and used to accommodate the needs of elderly learners (van Gerven, et al., 2006).

Nevertheless, the mapping of specific media type and their combination of procedural instruction for elder learner is still not clear. Therefore, the aim of this article is to review the literature on how properly design principle and what type of instruction can support elderly people in completing procedural task.

2 COGNITIVE PERSPECTIVE

2.1 *Cognitive aging*

Cognitive aging usually refer to a decline of processes that contribute to the efficiency of information processing. van Gerven, et al. (2006) mentioned that there are two main phenomena of cognitive aging included reduction of general cognitive speed and cognitive control. The cognitive speed affects every cognitive function at every level of processing. In the most speed cognitive tasks, the reaction time of the elderly is dramatically slow down than young. It's the most central mechanism of cognitive aging (Salthouse, 1996).

A second general decline is cognitive control that entails a whole group of mechanisms aimed at manipulating information in working memory and behavior. One of the most prevalent views of cognitive aging was that it primarily involves a reduction of the storage capacity of working memory (Hasher & Zacks, 1988). Mechanisms such as task switching, updating, searching, integration, coordination, and selection have been identified as separate cognitive control processes, most of which are affected by cognitive aging (Fisk & Sharp, 2004).

However, Design changes that provide environmental support for declining cognitive, perceptual, and psychomotor abilities can serve as a powerful intervention for maintaining and improving older adult performance. Training is also can improve performance at both the basic ability level and the level of task performance (Charness, 2008). By some cognitive theory can improve quality of life for an aging population.

2.2 *Cognitive Load Theory*

During the past two decades, CLT has become an influential theory in the fields of educational psychology and instructional design (Paas, et al., 2004; Sweller, 1988; Sweller, et al., 1998; van Merriënboer & Sweller, 2005).

CLT is aimed at developing training material that efficiently makes use of the available cognitive processing capacity and stimulates the learner's ability to use acquired knowledge and skills in new situations.

CLT provides a theoretical foundation for designing instructional materials to best enhance learning. The basic premise of this theory is that learning will be hindered if the instructional materials overwhelm a learner's cognitive resources. Research has indicated that several features of human cognitive architecture are especially important in instructional design. Specifically, cognitive load theory is based on a cognitive architecture consisting of a limited working memory that interacts with an unlimited long-term memory (Chandler & Sweller, 1992; Sweller, 1988). The constrains of working memory highline the importance of reducing unnecessary cognitive demands on individuals learning.

The central idea of CLT is that working memory plays a highly significant role in learning. However, because WM is very limited in both capacity and duration, learning can be seriously inhibited if instructional designers fail to take account of these limitations.

According to Sweller's CLT identifies three categories of cognitive load: extraneous, intrinsic and germane. Intrinsic cognitive load is considered as determined largely by element interactivity. On the other hand, extraneous cognitive load is determined by how the information is presented when intrinsic load is high, a high level of extraneous cognitive load can be a critical factor for successful learning (Carlson, et al., 2003). Thus, the instructional format (e.g., animations or static pictures) might influence the learning efficacy of a learning environment. By reducing extraneous cognitive load and increasing germane cognitive load the third type of cognitive load, referring to the effort involved in the processing, construction and automation of schemas more efficient learning may be possible.

2.3 *Working memory and procedural task*

Working memory is the resource through which people manipulate and actively keep information available for on-line processing. Working memory is activated in nearly all complex cognitive tasks, particularly those that require holding information in the face of distraction, or conducting simultaneous activities. However, procedural tasks are one of typically activity which need working memory. Procedural learning involves learning to carry out a series of acts or operations in the proper order. Whether rebuilding a car engine, folding a paper airplane, assembling a stereo surround sound system, or following a recipe, the overall goal is similar: to follow a series of steps through to a final product. Thus, an important issue in the study of procedural learning involves determining effective methods for helping individuals build strong, manipulate representations of task steps that will lead to reproducible, reliable outcomes especially in the future aging life.

3 PROCEDURAL INSTRUCTION DESIGN

3.1 *The general instructional design theory*

There are two general instruction design theory. CLT developed by John Sweller and colleagues

(Sweller, et al., 1998; van Merriënboer & Sweller, 2005). The second is Richard Mayer's (2001) CTML. The combination of the two theories yields a powerful tool for supporting skill acquisition in the elderly. Both CLT and CTML are based on a cognitive architecture in which a capacity-limited working memory is connected to an unlimited long-term store. Whereas CLT emphasizes the use of multimedia as a means to control cognitive load, CTML is more concerned with attention aspects of learning.

In 1998, CLT had been used almost exclusively to study instruction intended to decrease extraneous cognitive load. Some of the major effects that yield better schema construction and higher transfer performance and that may be attributed to a decrease in extraneous cognitive load are briefly summarized in Table 1 (Sweller, et al., 1998).

van Merrienboer & Sweller (2005) outlined the possible implications of CLT and CTML for elderly learners in multimedia-based learning environments. According to these authors, it is important to map systematically age-related cognitive declines on the potentially compensatory strategies offered by existing instructional theories. For that purpose, they formulated four groups of age-related cognitive declines: (a) Reduced processing capacity of working memory, (b) reduced cognitive speed, (c) reduced inhibition of irrelevant information, and (d) reduced coordination and integration of different information sources. Subsequently, they surveyed those CLT- or CTML-based instructional design principles that might compensate for each of these declines. An extended overview of this "mapping" approach is organized in Table 2 (van Merriënboer & Sweller, 2005).

3.2 *The property of different kinds of instructions*

Wileman (1993) categorized static visual information into three major ways to present objects, progressing from concrete to abstract, as pictorial symbols (photograph, illustrations or drawings), graphic symbols (pictogram, signs and icons) and verbal symbol (text). And the choice of which

Table 1. Some effects studied by Cognitive Load Theory.

Effect	Description
Goal-free effect	Replace conventional problem with goal-free problems that provide learners with an a-specific goal
Worked example effect	Replace conventional problems with worked examples that must be carefully studied
Completion problem effect	Replace conventional problems with completion problems, providing a partial solution that must be completed by the learners
Split attention effect	Replace multiple sources of information (frequently pictures and accompanying text) with a single, integrated source of information
Modality effect	Replace a written explanatory text and another source of visual information such as a diagram (unimodal) with a spoken explanatory text and a visual source of information (multimodal)
Redundancy effect	Replace multiple sources of information that are self-contained (i.e., they can be understood on their own) with one source of information

Table 2. Age-related cognitive declines on the potentially design principle.

Age-related cognitive decline	Compensatory multimedia strategy
Reduced processing capacity	Bimodal (audiovisual) presentation Worked examples instead of conventional practice problems Goal-free instead of goal-specific practice problems Presenting instruction in a parts-whole sequence Omitting redundant information
Reduced cognitive speed	Bimodal (audiovisual) presentation Enhanced timing Omitting redundant information Presenting instruction in learner-controlled segments
Reduced inhibition	Omitting redundant information Attention support
Reduced coordination and integration	Bimodal (audiovisual) presentation Enhanced timing temporal Enhanced layout Omitting redundant information Presenting instruction in a parts-whole sequence

of these symbols to use is directly related to the major objective of communication and the specific information that is suitable to the specific situation. For instance, the more abstract symbols, pictograms, are often used in situations where the meaning of a message needs to be comprehended quickly, since well-designed pictograms are intended to convey meanings perceptively (Yamazaki, et al., 2008). Yamazaki et al. (2008) examined the effectiveness of pictograms representing actions in lathe procedures and for steps in manufacturing procedures. The results showed that the pictograms developed for lathe instructions conveyed intended meanings as effectively as common public signs.

Furthermore, being well planned and designed for the visual content and detail is very important. Well-displayed designs are very important for multimedia learning. Michas & Berry (2000) demonstrated that enhanced line-drawings which use symbols, such as arrows and highlights, can help to make temporal and spatial relationships within graphical representations clearer and to better convey information about how to get from one step of the procedure to the next.

In the learning and instruction domain, the visual instructions often are distinguished into dynamic and static. Several studies compared the learning effectiveness between dynamic and static instruction. Static information has benefit for the learning efficacy, and its effectiveness can possibly be increased by using certain key pictures that illustrate very specific moments of the process or the procedure to be learned. In the study by Boucheix & Schneider (2009), one of their experiments indicated that animated as well as integrated sequential static frames enhanced comprehension in learning dynamic mechanical systems.

Most studies noted that dynamic visual information was superior to the static visual information, especially in specific areas, under specific circumstances. For instance, the results of the meta-analysis conducted by Höffler & Leutner (2007) reveal the greater benefits of animations when procedural-motor knowledge rather than problem-solving knowledge or declarative knowledge is being taught. Wong et al. (2009) also mentioned that animated instructions are superior to static graphics for cognitively based tasks that involve human movement. On the other hand, a recent study indicated that a combination of instructional animation with static pictures was more efficient in promoting learning than visualizations composed only of videos (Ayres & van Gog, 2009).

From the above literature the property of different kinds of visual information can be discussed from three points of view in the future. Firstly, the most discussed issue is dynamic and static presentation, the second is planning and reorganizing the visual information presentation style, and the third is comparison of the concrete and abstract content.

3.3 *The mapping effective procedural instructional design principle and media types on the cognitive constraints of older learners*

The procedures are characterized by the existence of a beginning and an end, and between these two extreme points, there are a succession of steps describing each action or some steps of the procedure. The order of these steps is very important. Therefore, using multimedia presentations instead of paper-based documents could potentially provide an alternative way for helping the learning of procedural contents. Arguel & Jamet (2009) combination of dynamic and static visualizations, was more effective than either format alone. The finding was that the best strategy to present the statics one at-a-time (dynamically) in synchronization with the animation content. This kind of format (See Fig. 1) is especially fit for procedural content learning.

In Table 2, there are eight compensatory multimedia strategies for the four age-related cognitive declines which mean that the elderly took more advantage of the instructional strategy than their younger counterparts. In the following, the implications of the instructional design principles given in Table 1 would be detail discussed on the determining procedural learning instruction by (Arguel & Jamet, 2009).

First, bimodal presentation of instructional material stimulates a more efficient use of working memory capacity by using two instead of one modality-specific slave systems (Mayer, et al., 2001). In the case of the instruction, animations are accompanied by both spoken text and static pictures. In this way, the depiction of micro-steps from the procedure and its natural development are maintained. Moreover, the static pictures are visible throughout the animation and act to limit the transience of the animation. Relieving the visual system is especially beneficial for the elderly because of age-related capacity limits.

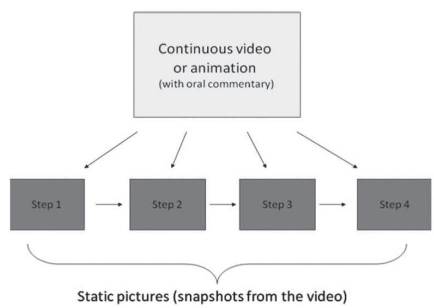


Figure 1. Schematic representation of the learning material (Arguel & Jamet, 2009).

A second strategy to compensate for reduced processing capacity is the use of worked examples instead of conventional practice problems. Worked examples reduce extraneous activity in working memory which again is especially beneficial for older learners (van Gerven, et al., 2002; van Gog, et al., 2004). The video plus static picture format in this case which induces a capacity demanding means-ends strategy, whereas the key steps pictures draw attention to problem states and operations to progress toward a solution.

Third, goal-free instead of goal-specific practice problems doesn't be applied. The fourth is presenting instruction in a parts-whole sequence. In the case of video plus static picture instruction, the static pictures are the simple part minimizes working memory load in the earlier stages of learning and then see the whole process of video can maximizes the chance of correctly combining these parts at a later stage. And if it is also organizing the learning material into a parts-whole sequence helps elderly learners integrate the different parts of the learning material in an effective step-by-step way.

Fifth is the omitting redundant information or so called split attention effect. The negative impact of the split attention effect (Ayres & Sweller, 2005; Sweller & Chandler, 1994) between static pictures and animation can be avoided by using pictures that are exact copies of frames from the video, so there is no need to integrate this identical information. Sixth is the presenting instruction in learner-controlled segments which also not emphasis on this examples. Seventh is the enhanced timing, Integration of these information sources can be optimized by enhanced timing, the case of video plus static picture format taken more learning time by watch the video and understand the static pictures.

The last is enhanced layout, the spatial layout of the material might support information integration by grouping related elements and separating unrelated elements. The case of video plus static picture format fit in with this principle. As Arguel & Jamet (2009) the act of presenting static pictures near videos could be beneficial because the pictures highlight to learners the crucial steps of an unknown procedure, rather than maintaining transient information.

4 DISCUSSION AND FURTHER RESEARCH

Procedural learning involves learning to carry out a series of acts or operations in the proper order and to follow a series of steps through to a final result. It's closely related to the capacity of working memory which people manipulate and actively keep information available for on-line processing. Furthermore, learning procedural task is very important for the elderly in daily life which would like have an independent aging life. Therefore, using the cognitive related theory to supporting skill acquisition and learning complex procedural task in the elderly should be more concerned. According to the theory of CLT and CTML, the instruction media types are complementary for the elderly people learning with cognitive declining. By the classical view of the CLT, learning materials should keep the learners' extraneous cognitive load at a minimum and germane load at a maximum during the learning process. The form of instruction material play very important role.

The limited capacity of working memory highlights the need for effective design that minimizes the extraneous cognitive demands placed on patients and frees their mental resources to better process important information. When applying a cognitive factors approach to the development and refinement of procedural instruction, it is also important to note the circumstance in which the material will likely be presented. Even when designers use optimal strategies to promote effective learning, a message's modality can affect cognitive load and subsequent elderly people understanding and recall of material.

Table 3. The key to optimal procedural learning in the elderly by mapping effective procedural instructional design principle and media types.

Potentially design principle	The strategy of the determined instruction media type
Bimodal (audiovisual) presentation	Animations are accompanied by both spoken text and static pictures
Worked examples instead of conventional practice problems	Key steps pictures draw attention to problem states and operations to progress toward a solution
Presenting instruction in a parts-whole sequence	Integrate the different parts of the learning material in an effective step-by-step way
Omitting redundant information	Animation can be avoided by using pictures that are exact copies of frames from the video
Enhanced timing	More learning time by watch the video and understand the static pictures
Enhanced layout	Presenting static pictures near videos

To conclude, our literature review shows that the key to optimal procedural learning for the elderly might be determined the instruction media type as combination of dynamic and static visualizations which based the age-related cognitive declines on the potentially design principle, the mapping as the Table 3.

Much of the research in the field have compared static pictures with animations only, and have not combined the two formats. However, combining them together under certain conditions such as procedural learning and specific population such as elderly people seems to have produced the best learning outcomes than either format presented individually. Consequently future research should be conducted to test the effectiveness by the truly empirically studies.

REFERENCES

Arguel, A. & Jamet, E. 2009. Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior* 25(2): 354–359.

Ayres, P. & Sweller, J. 2005. The split-attention principle in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*: 135–146. New York: Cambridge University Press.

Ayres, P. & van Gog, T. 2009. State of the art research into cognitive load theory. *Computers in Human Behavior*, 25(2): 253–257.

Boucheix, J.-M. & Schneider, E. 2009. Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction* 19(2): 112–127.

Brunye, T., Taylor, H., Rapp, D. & Spiro, A. 2006. Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology* 20(7): 917–940.

Carlson, R., Chandler, P. & Sweller, J. 2003. Learning and understanding science instructional material. *Journal of Educational Psychology* 95(3): 629–640.

Chandler, P. & Sweller, J. 1992. The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62(2): 233–246.

Charness, N. 2008. Aging and human performance. *Human Factors* 50(3): 548.

Craik, F.I.M. & Salthouse, T.A. 2000. *The Handbook of Aging and Cognition* (Second ed.). New Jersey: Lawrence Erlbaum associates.

Fisk, J.E. & Sharp, C.A. 2004. Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology* 26(7): 874–890.

Höffler, T.N. & Leutner, D. 2007. Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction* 17(6): 722–738.

Hasher, L. & Zacks, R.T. 1988. Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* 22: 193–226. New York: Academic.

Kang, N.E. & Yoon, W.C. 2008. Age-and experience-related user behavior differences in the use of complicated electronic devices. *International Journal of Human-Computer Studies* 66(6): 425–437.

Kirschner, P.A. 2002. Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction* 12(1): 1–10.

Mayer, R.E. 2001. *Multimedia learning*. Cambridge: Cambridge University Press.

Mayer, R.E., Heiser, J. & Lonn, S. 2001. Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology* 93(1): 187–198.

Michas, I.C. & Berry, D.C. 2000. Learning a procedural task: effectiveness of multimedia presentations. *Applied Cognitive Psychology* 14(6): 555–575.

Paas, F., Renkl, A. & Sweller, J. 2004. Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science* 32(1): 1–8.

Salthouse, T. 1996. The processing-speed theory of adult age differences in cognition. *Psychological review* 103(3): 403–428.

Scialfa, C.T., Ho, G. & Laberge, J. 2004. Perceptual aspects of gerotechnology. In D. Burdic & S. Kwon (Eds.), *Gerotechnology*: 18–41. New York: Springer Publishing Company.

Sweller, J. 1988. Cognitive load during problem solving: Effects on learning. Cognitive Science: *A multidisciplinary journal*, 12(2): 257–285.

Sweller, J. & Chandler, P. 1994. Why some material is difficult to learn. *Ethics & Behavior* 12(3): 185–233.

Sweller, J., van Merriënboer, J.J.G. & Paas, F. 1998. Cognitive architecture and instructional design. *Educational Psychology Review* 10(3): 251–296.

Vakil, E. & Agmon-Ashkenazi, D. 1997. Baseline performance and learning rate of procedural and declarative memory tasks: Younger versus older adults. *The Journals of Gerontology: Series B*, 52(5): P229.

van Gerven, P.W.M., Paas, F. & Tabbers, H.K. 2006. Cognitive aging and computer-based instructional design: Where do we go from here? *Educational Psychology Review* 18(2): 141–157.

van Gerven, P.W.M., Paas, F., van Merriënboer, J.J.G. & Schmidt, H.G. 2002. Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learning and Instruction* 12(1): 87–105.

van Gog, T., Paas, F. & Van Merriënboer, J.J.G. 2004. Process-oriented worked examples: Improving transfer performance through enhanced understanding. *Instructional Science 32*(1): 83–98.

van Merriënboer, J.J.G. & Sweller, J. 2005. Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review* 17(2): 147–177.

Wileman, R.E. 1993. *Visual Communicating*. New Jersey: Educational Technology.

Wilson, E.A.H. & Wolf, M.S. 2009. Working memory and the design of health materials: A cognitive factors perspective. *Patient Education and Counseling* 74(3): 318–322.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G.A. Paas, F., et al. 2009. Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior* 25(2): 339–347.

Yamazaki, A.K., Goto, K., Taki, H. & Hori, S. 2008. An effectiveness study of pictogram elements for steps in manufacturing procedures. *Lecture Notes in Computer Science* 5179: 680–686.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Physical load reduction at the waist of a care worker with wearable robot

Takashi Inoue, Masato Sakurai & Sakae Yamamoto
Tokyo University of Science, Tokyo, Japan

Miwa Nakanishi
Keio University, Kanagawa, Japan

ABSTRACT: To examine the physical load at the waist of a care worker when shifting the position of a patient in with and without a wearable robot (Muscle Suit), we measured observer's caring posture during the task. Using Muscle Suit may reduce risk of backache. In the result, it is suggested that there is a possibility to reduce the risk of low back pain for care worker in daily task.

Keywords: care worker, backache, wearable robot, body mechanics, motion analysis

1 INTRODUCTION

1.1 *Aged societies and care work*

It is well-known that elderly people are increasing with year and year. According to world population prospects from the United Nations Population Division database revised in 2008, the world population over the age of 65 will reach about 523 million, 7.6% of the total world population in 2010. In Japan, over the age of 65 people reach 22.6% of the total already in 2010 [United Nation Population Division (2008)]. As the functions of human body decline with aging and elderly people need nursing care, the role of care workers becomes more and more important.

However, recently, there is a problem that the work of nursing care imposes a care worker on physical damage [National Institute for Occupational Safety and Health (1997)]. Especially, when the care worker supports patient's body, the damage to waist is bigger. Actually in Japan, 47.4% of the total in care workers is backache [Ryo Minematsu (2005)]. One of the reasons is a posture of highly bending in care worker when nursing care for the patient. The highly bending posture during the work nursing care work has a high risk to cause backache.

1.2 *Applying a wearable robot to nursing care scene*

To solve the above problem, this paper proposes to introduce a wearable robot to care workers when nursing care. The wearable robot supports muscles to the wearer. The wearable robot; Muscle Suit is developed at the Kobayashi lab, Tokyo University of Science [Aida et al. (2009)]. By shrinking the McKibben tube of the Muscle Suit, it can support muscle of wearer. In the work of nursing care, wearing Muscle Suit has possibilities to reduce the risk of backache because of supporting the muscle of the care worker. Figure 1 shows the Muscle Suit. A previous study examined the physical load at the neck and waist with or without Muscle Suit when making a bed and it is suggested that Muscle Suit is effective in reducing the physical load at the waist of wearer [Nakanishi and Yamamoto (2009)].

1.3 *Evaluation of Muscle Suit and body mechanics*

In the evaluation of Muscle Suit, it is discussed how much reduce the muscular load with Muscle Suit compared without it mostly. In general a care worker often postures a highly bending at the waist during the task of nursing care. As previously mentioned, the highly bending posture may cause the care worker backache. In this paper, it takes into account to reform the posture in the work of nursing care with Muscle Suit.

Also, there is body mechanics in the skill of nursing care. Body mechanics is the skill to reduce an imposition of a care worker [Gassett et al. (1996), Karahan et al. (2009)]. Using body mechanics in personal care task leads to precaution of backache. For example body mechanics, approach patient's body, use muscle in every part of the body and lowered center of gravity. If this paper think

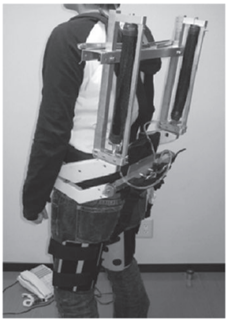


Figure 1. Muscle Suit.

introduction of Muscle Suit to caring scene, this paper need to examine an effect to body mechanics in wearing the Muscle Suit.

2 PURPOSE

The purpose of this paper is to reveal the effect of the wearable robot when carrying out the work of nursing care. In addition, when a care worker carries out the work of nursing care with the wearable robot, this paper earns the wearable robot (This paper uses Muscle Suit.) rating from care workers and then suggests an important to development team.

3 METHODOLOGY

3.1 *Task*

The task is to turn a patient on their back and to the right on a bed. This task called "Changing body position". Changing body position does 4 to 7 times toward one patient in a day. Figure 2 shows each scene of task in this experiment. By Figure 2, it is obvious that a care worker postures that damages his waist. One of the load of waist is long hours of deep bending posture. The observers were asked to hold a left arm of a patient with their left hand and a left side of back or left leg of a patient with their right hand in this task. It was used a model of human body (*Sakura*, KYOTOKAGAKU) as a patient.

3.2 *Conditions*

The condition was with and without Muscle Suit. Each observer performed this task five times for both conditions of with and without a Muscle Suit, respectively. The interval task to task was 15 sec. Thinking order effect in two conditions, observers were divided into two groups, one doing the task with Muscle Suit at first and the other doing the task without it at first. The inter-condition interval had enough time to refresh from the physical load

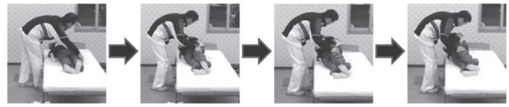


Figure 2. Task of this paper.

in each condition. To learn how to use Muscle Suit, each observer had an adequate practice with it.

3.3 *Apparatus*

Measurement records a posture and movement of observers in task. For record, this paper uses three camcorders. Figure 3 shows an arrangement of apparatus.

Muscle Suit (3.5 kg), Compressor "JUN-AIR" (Kuroda International): Compressor send air to the McKibben tube, and support muscle to wearer, Camcorder (HDR-S12, SONY), Dummy "*Sakura M75*" (21 kg, KYOTOKAGAKU), Bed (Height 48 cm), Tripod (SILK).

3.4 *Measurement*

In a task such as changing body position in this experiment, one of the body mechanics is a low center of gravity for care worker. A low center of gravity recommended with bending at the knees and ankles when performing the task. According to this, the waist of a care worker goes down, and it is not necessary to maintain the highly bending posture that compares not to use body mechanics. As we evaluation a posture of patient in task, we use an image in camcorder 1. To evaluate the relationship between the uses of Muscle Suit and body mechanics, the knee and ankle angle of each observer was measured with and without Muscle Suit in the task. Figure 4 shows a measurement part. Knee angle and ankle angle defined Figure 4. From image of camcorder 1, the knee angle is the angle between the right thigh and leg with the right knee in the vertex, the ankle angle is the angle between the right leg and vertical axis with the right ankle in the vertex.

3.5 *Observer*

Observers were a care worker in nursing facilities (Ginnofune Yokohama, Yokohama) to should be able to get care worker's view and do according to the procedure to which the work of the experiment task is done on the scene of nursing care. Ten people took part in this experiment. Table 1 is observer's profile. It experimented after it explained the content of the experiment enough to each observers and the observer's agreement was obtained.

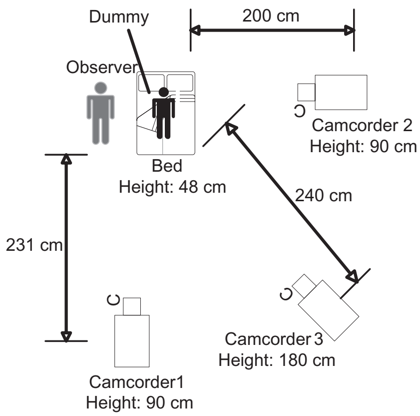


Figure 3. Experiment surrounding.

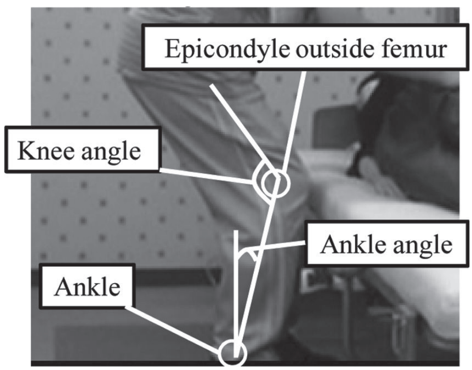


Figure 4. Knee angle and ankle angle.

Table 1. Observer's profile.

People	Three males, seven females	
	Average	S.D.
Age	39.1 age	11.4 age
Care term	3.5 years	0.78 years
Height	161 cm	6.37 cm
Backache	Yes 7 people, No 3 people	

3.6 *Subjective evaluation in care worker*

When people evaluate Muscle Suit, people often concentrate in muscular power assistance. But the evaluation of Muscle Suit seen from wearer's position is not done. We think that it is necessary to take wearer's opinion into consideration. In this paper, as think introduction of Muscle Suit to caring scene, we decided a care worker's opinion to be obtained. As a method, after the experiment, each observer was asked to fill out the questionnaire regarding to the movement when wearing Muscle Suit, the capability in nursing facilities and more. This is an important evaluation from the point of view of a care worker.

4 RESULTS AND DISCUSSIONS

4.1 *Knee and ankle angles*

Figure 5 and Figure 6 are the results of the knee and ankle angles for each observer based on the average of five repetitions. The horizontal axis of these figs shows angle without Muscle Suit, respectively. The vertical axis of these figs shows angle with Muscle Suit, respectively.

4.1.1 *Knee ankle*

In Figure 5, the right-bottom region refers to the use of body mechanics when wearing the Muscle

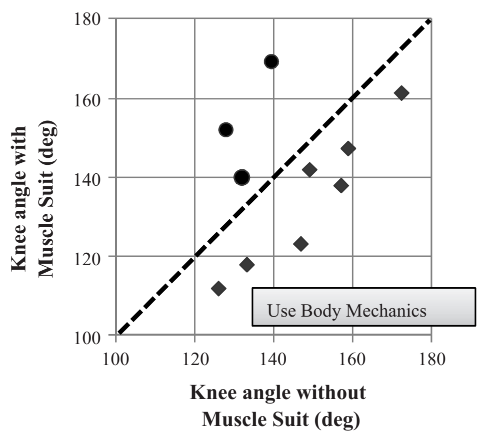


Figure 5. Muscle Suit compared.

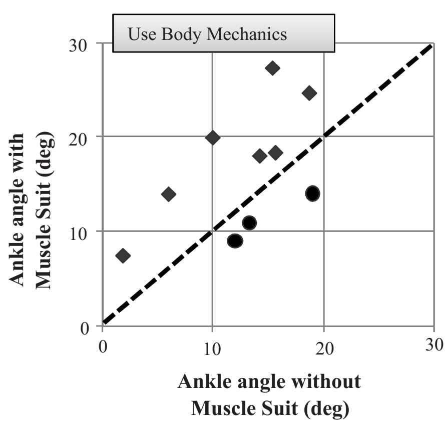


Figure 6. Compare ankle angle.

Suit compared of the results without it since decreasing the knee angle of the observer lowers their center of gravity. We think that seven observers of ten use body mechanics to do changing body position in wearing Muscle Suit.

4.1.2 *Ankle angle*

In Figure 6, the left-top region refers to the use of body mechanics when wearing the Muscle Suit compared to the results without it since increasing the ankle angle of the observer lowers their center of gravity as well. As knee angle, we think that seven observers of ten use body mechanics to do changing body position in wearing Muscle Suit. These seven observers are same person to use body mechanics in knee angle. We said that seven observers work to bend their knees and work to bend the ankle at the same time.

4.1.3 *Discussions to use of body mechanics*

This is thought that the influence by the mechanism of the Muscle Suit is large. Figure 7 shows the one that the mechanism and the effect of the muscle suit were described. When making power, the wearer deeply bends its knee to use the power of the Muscle Suit to its maximum. That is to say, wearing the Muscle Suit improves posture, and it leads to the prevention of backache. It can be said that wearing the Muscle Suit will make the wearer use body mechanics above. But there are a lot of kinds of body mechanics, and it is understood that it is not enough only by this experiment [National Institute for Occupational Safety and Health (1994)].

4.2 *Subjective evaluation in care worker*

Figure 8 shows the results of the questionnaire regarding the effect at the waist and the ease of movement with the Muscle Suit based on ten observers responses. In Figure 8, 70% of all the observers indicated an effect at the waist with the Muscle Suit. On the other hand, the 30% of the females (three) did not feel it. The mean height of these females is 152 cm, much lower than that of the average of 10 observers. Therefore, it is considered that the Muscle Suit used in this experiment does not work well for lower height observers because of the size. The size adjustment of the Muscle Suit used this time was two stages, and it was not possible to match it to the physique of a lower woman. Therefore, it is necessary to make the Muscle Suit with abundant size kinds.

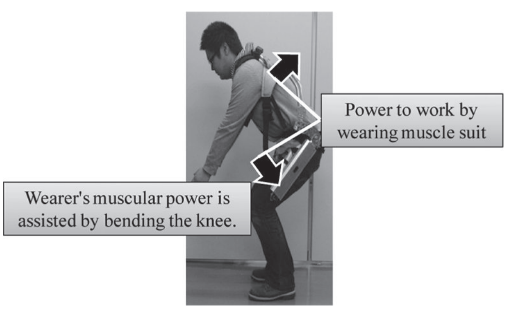


Figure 7. Effect of Muscle Suit.

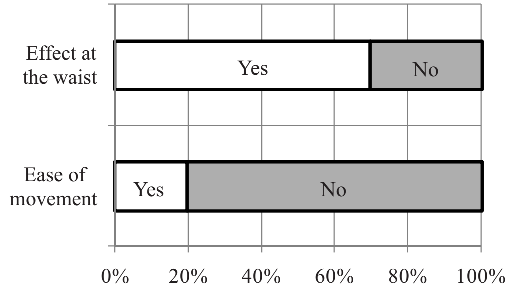


Figure 8. Observer's subjective evaluation.

As shown in Figure 8, most of observers feel a limitation in the movement of the upper body with the Muscle Suit. The Muscle Suit used in this experiment is strapped to the shoulder by the observer in Figure 1; consequently it is not easy to move their upper body. From the capability point of view in nursing facilities, the variable size of the Muscle Suit and the free movement of the upper body when wearing the Muscle Suit will be future work.

5 CONCLUSION

The Muscle Suit makes the wearer use the body mechanics when working about the work of nursing care changing body position. The Muscle Suit has been understood also though is effective of the posture improvement in addition to muscle support. But, because the body mechanics has a lot of things, it is necessary to examine other one. From the subjective evaluation of a care worker, most of observers feel the effect at the waist of observer with the Muscle Suit indicating the possibility for use of the Muscle Suit in nursing facilities.

ACKNOWLEDGEMENT

Muscle Suit is receiving of Mr. Hiroshi Kobayashi‘s (Mechanical Engineering Section, Department of Engineering, and Tokyo University of Science) cooperation. It tenders an apology here.

REFERENCES

Aida, T., Nozaki, H. & Kobayashi, H. 2009. Development of muscle suit and application to factory laborers. In *Proceedings of the 2009 IEEE International Conference on Mechatronics and Automation* (Changchun, China, 2009) 1027–1032.

Gassett, R.S., Hearne, B. & Keelan, B. 1996. Ergonomics and body mechanics in the work place. *Orthop Clin North Am* 27(4): 861–879.

Karahan, A., Kav, S., Abbasoglu, A. & Dogan, N. 2009. Low back pain: prevalence and associated risk factors among hospital. *J of Advanced Nursing* 65(3): 516–524.

Nakanishi, M. & Yamamoto, S. 2009. Effectiveness of using muscle suit in daily nursing care scenes. *J of Ergonomics in Occupational Safety and Health* 11 (Supplement): 90–93 (in Japanese).

NIOSH 1997. *Musculoskeletal disorders and workplace factors*. National Institute for Occupational Safety & Health (NIOSH) Publication No. 97–141.

NIOSH 1994. Work place of back belts. Department of Health and Human Service, National Institute for Occupational Safety & Health (NIOSH) No. 94–122.

Minematsu, R. 2005. Current state of backache measures in care worker. *Daily Employment disease medical journal* 53: 35–38.

UNPD 2008. *World population prospects: The 2008 revision population database*. United Nations Population Division (UNPD), retrieved from http://esa.un.org/unpp/

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Using a muscle suit in nursing care to prevent caregivers' backache

Miwa Nakanishi
Keio University, Yokohama, Japan

Masato Sakurai & Sakae Yamamoto
Tokyo University of Science, Tokyo, Japan

ABSTRACT: In recent years, patients and caregivers at nursing homes have expected the development of wearable robots that assist the muscles of their wearer hoping that such devices would address the physical difficulties experienced by caregivers and the psychological problems experienced by patients as a result of them. In this study, we use the muscle suit that is under development to experimentally examine its application not only to work that requires temporary use of large muscles (i.e., transfer assistance), but also to work that requires repetitive use of comparatively small muscles. Specifically, we focus on changing sheets, which is one of the tasks that often gives rise to complaints from workers of heavy physical loads, and we evaluate the capacity of the muscle suit to mitigate the difficulty of the task by ElectroMyoGraphy (EMG).

Keywords: muscle suit, wearable robots, nursing care, backache, EMG

1 INTRODUCTION

The percentage of senior population is rapidly increasing in many countries, including Japan, and the nursing-care load has become a pressing health issue that needs to be addressed immediately. The current problems in nursing care can be divided into two categories: those that affect the caregivers and those that affect the patients. A significant problem for caregivers is the physical damage that they suffer as a result of their nursing-care activities (M. Billinghurst et al., 1997; G. Bjorn and D. House, 2003; R. Kishi et al., 2002; B. Cameron, 1967; J.T. Best, 2002; Z. Roupa, 2008). In particular, they cannot avoid suffering large muscle loads in the waist area when they hold or carry a patient. and these loads cause backaches which may leave them in chronic pain. As a result, recent years have seen the development of welfare machinery such as lifts or automatic beds. A significant patient problem is mental discomfort (R. Campbell, 1984; P.M. Peter et al., 1997; M. Guha, 2007). Many elderly patients experience fear, confusion, and loneliness when handled by mechanical devices instead of by human caregivers and, with the spread of welfare machinery, the problem could grow.

Originally, it was assumed that wearable robots would be applied to the fields of construction or civil engineering (H. Kobayashi and H. Nozaki, 2007). However, the application of wearable robots that assist the user's muscles are now focused on the nursing-home issues (T. Tanaka et al., 2008; K. Yamamoto et al., 2002, Y. Sankai, 2006; H. Kobayashi and T. Tsuji, 2007), because it is expected that as new welfare machinery, they can improve the working conditions and health of caregivers and improve the living conditions of patients by enabling care gives to deliver humane, hands-on care.

In previous studies (M. Nakanishi and S. Yamamoto, 2009), we applied one of the most promising wearable robots, the "muscle suit," to the task of assisting the transfer of a patient from a bed to a wheelchair and to get in and out of a bathtub, and we confirmed that the muscle load experienced by caregivers wearing the muscle suit was reduced.

Because nursing care covers all aspects of the life of elderly people, there are many tasks other than those just mentioned that subject the caregiver to significant physical loads, but the most serious occupational hazard that results is backache. Because the muscle suit can be easily attached and removed, is sufficiently safe, and is waterproof, it is said to be an appropriate wearable robot for nursing-home care. Thus, in the present study, we expand the field of application of the muscle suit from work requiring temporary use of large muscles (i.e., patient-transfer assistance) to work involving repetitive use of comparatively small muscles, and we examine the possibility of making nursing care more comfortable and more satisfying. In particular, we focus on changing sheets, which is one of the tasks that results in the most complaints from caregivers for heavy physical loads, and we experimentally investigate the question of whether the muscle suit can mitigate the painfulness of this task.

2 MUSCLE SUIT

Several types of wearable robots that assist the users' muscles are currently under development. In the present study, we examine the applicability of the muscle suit (H. Suzuki and H. Kobayashi, 2004; H. Kobayash et al., 2006; H. Kobayashi et al., 2004), which was designed and implemented by Kobayashi Laboratory of the Tokyo University of Science, from the viewpoint of ergonomics. The main part of the muscle suit is the tube-type actuator, which is called a McKibben tube (B. Tondu and P. Lopez, 1997; S. Eskiizmirliler, 2002; C.P. Chou and B. Hannaford, 1996; F. Daerden and D. Lefeber, 2001). The McKibben tube, which is contracted and relaxed by compressed air, functions as an artificial muscle to increase the user's physical strength. Fig. 1 shows the overall Muscle-Suit system. The muscle suit is composed of McKibben-type artificial muscles, an air compressor, a pressure regulator, and a commercial PC. The McKibben-type artificial muscles can be installed into the parts of the muscle suit that correspond to the human muscles that to be assisted. For the present study, which aims to reduce the load on the waist, two McKibben-type artificial muscles were installed to assist the user's lower back muscles. In this configuration, the muscle suit has a total mass of approximatley 3.5 kg, and can deliver a maximum force of about 90 N at the waist.

3 METHOD

3.1 *Subjects*

The subjects were 17 female nursing students who were skilled in nursing care. Their average age was 24.5 years old (SD 7.12), their average height was 160.0 cm (SD 6.7), and their average weight was 54.9 kg (SD 10.74).

3.2 *Experimental task*

The subjects' task in this experiment was to change sheets. They removed a sheet from a bed and replaced it with a different sheet, tucking in each corner of the new sheet under the mattress in the order shown in Fig. 2. They were required to complete the task as illustrated in Fig. 3, which shows the procedure they learned for tucking in a sheet at a mattress corner that is illustrated in a textbook (M. Okazaki, 1998). The bed frame occupied a space of $190 \times 91.5 \times 45$ cm^3, and the mattress was 6 cm thick. The sheet dimensions were 294×84 cm^2.

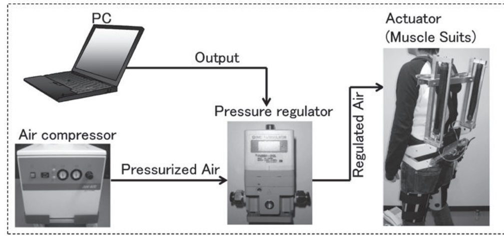


Figure 1. Overall system of muscle suit.

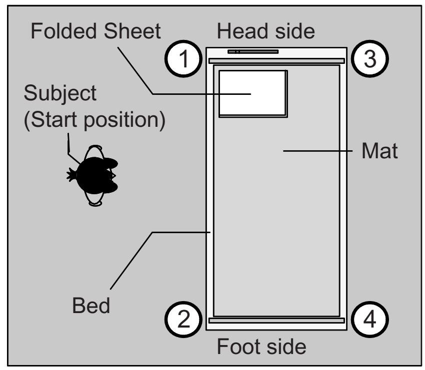


Figure 2. Experimental environment.

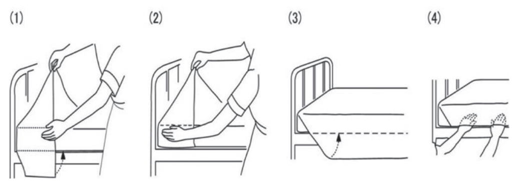


Figure 3. Process of setting a sheet at a corner (quoted from the reference [23]).

3.3 *Control of muscle suit*

To ensure that the experimental conditions were the same for all the subjects, the muscle suit was controlled by software on a commercial PC. High-pressure air (2 atm = 101,235 PA) was pumped into the McKibben artificial muscles so that the subject's bent-forward posture was maintained as shown in Fig. 4. The subjects commanded a control pro-

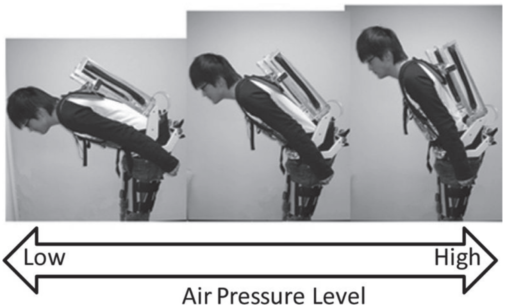


Figure 4. Subjects' posture when high-pressured air was infused to McKibben artificial muscles.

gram to contract and relax the McKibben artificial muscles, and the delay between their command and the completion of the actual contraction or relaxation was approximately 1 s.

In this experiment, a JUN-AIR (KURODA International Inc.) and an electro-pneumatic regulator (Hitachi Medical Corporation) were used as air compressor and pressure regulator, and ThinkPad T41 (made by IBM) was used as a PC.

3.4 *Experimental conditions*

The experiment was explained to the subjects, who had agreed to participate in advance. Each subject performed the given task ten times with and ten times without the muscle suit, which was done before any data were recorded. They were divided into two groups to eliminate the order effect. Thus, while one group of 9 subjects performed the task ten times with the muscle suit, the other group of 8 subjects performed it without the muscle group, and then the group of 9 subjects performed the task ten times without the muscle suit, while the other group performed it ten times with it. Every subject rested 60 minutes between the sets of 10 tasks to recover from muscle fatigue.

3.5 *Measurements*

To evaluate their muscle load in the waist and in the shoulders during the task, ElectroMyoGraphies (EMGs) were recorded specifically on the lumbar erector spinae muscle, which was expected to be supported by the muscle suit, and the upper trapezius muscle, which was expected to suffer from the weight of the muscle suit (C.W. Thompson and R.T. Floyd, 1998). Surface electrodes were placed on the points shown in Figs. 5a and 5b. To record the EMG signals, a WEB5500 (Nihon Kohden) multi-telemeter system was used. The sampling rate was set at 2 kHz, the time constant at 0.01 s, and the high-cutoff filter at 500 Hz.

3.6 *EMG reference data set*

The EMG reference data on the lumbar erector spinae and on the upper trapezius muscles were recorded before the task, both with and without the muscle suit. Data with the muscle suit were recorded for 30 s, while the subjects remained standing and holding a 5-kg weight, as shown in Fig. 6a. Data without the muscle suit were recorded for 30 s, while the subjects held their arms up at 90 degrees, as shown in Fig. 6b.

3.7 *Analysis*

Because the subjects tended to bend forward more to tuck in the sheet at the third and forth corners, we analyzed the EMG data recorded during these sections. For the analysis, we used the Integrated EMG (IEMG), which is defined as the area under the curve of the rectified EMG signal or, in other words, the integral of the absolute value of the raw EMG signal. We calculated the IEMG per

Figure 5a. Points of surface electrodes to record EMG of lumbar erector spinae (quoted from the reference by Thompson, C.W. and Floyd, R.T., 1998).

Figure 5b. Points of surface electrodes to record EMG of upper trapezius muscle (quoted from the reference by Thompson, C.W. and Floyd, R.T., 1998).

Figure 6a. Subjects' posture when reference EMG of lumber erector spinae was recorded.

Figure 6b. Subjects' posture when reference EMG of upper trapezius muscle was recorded.

second from data recorded for each subject from the waist and the shoulder during one ten-task session and from reference data recorded before the task and calculated the ratio of the former to the latter. We defined the average ratio defined over 10 trials by each subject in each condition (with and without the muscle suit) as the muscle-load index on the waist and the shoulder. Using the t-test, we tested for statistically significant differences between the tests with and without the muscle suit.

4 RESULTS

Figure 7 compares the muscle load on the right shoulder for tests with and without the muscle suit. In this figure, the x axis of each plot corresponds to the average ratio of the IEMG during the task to the reference data for each subject's right shoulder when they did not use the muscle suit, and the y axis shows the same ratio for test in which the subjects did use the muscle suit. Fig. 8 also compares the muscle load on the left shoulder between the same two conditions. These data show that the muscle load in the shoulders was reduced or did not significantly increase for 16 out of 17 subjects. From this result, we see that muscle suit adds little load to the shoulders in this task. Furthermore, to examine why the muscle load on the shoulders was reduced for some of the subjects with the muscle suit (even though it assists the waist and not the shoulder), we recorded their task on video and analyzed their motions during the task. We find that the subjects whose muscle load on the shoulders was reduced by the muscle suit tended not to move their arms more than necessary when they wore it. This strategy apparently leads to the mitigation of the load on the shoulders.

Fig. 9 compares the muscle load on the right side of the waist for tests with and without muscle suit. In this figure, the x axis of each plot corresponds to the average ratio of the IEMG during the task to the reference data for each subject's right-side waist area when they did not use the muscle suit, and the y axis shows the same ratio for test in which the subjects did use the muscle suit. Furthermore, Fig. 10 compares the muscle load on the left-side waist area between those two conditions. The data indicate that the muscle load on one or both sides of the waist was significantly reduced for 14 out of 17 subjects. When the subjects tuck in a sheet at a corner of a bed without the muscle suit, they are obliged to bend forward and

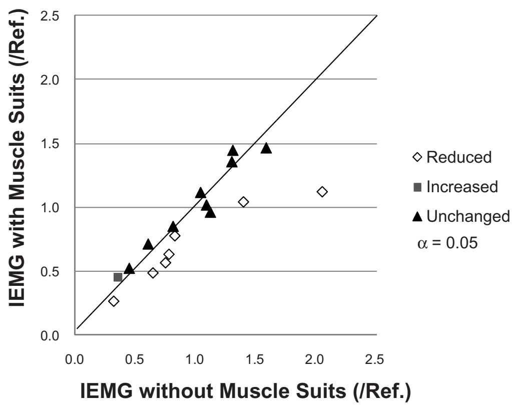


Figure 8. Comparison of muscle load on the left shoulder between the conditions with and without muscle suit.

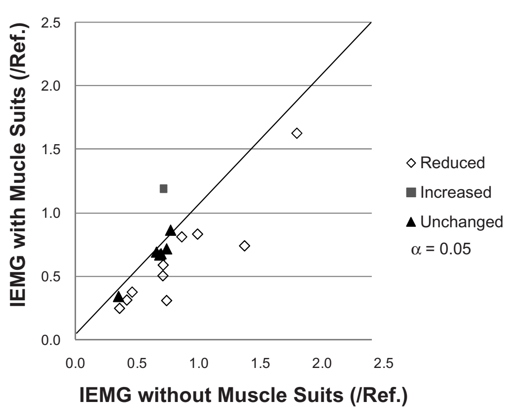


Figure 7. Comparison of muscle load on the right shoulder between the conditions with and without muscle suit.

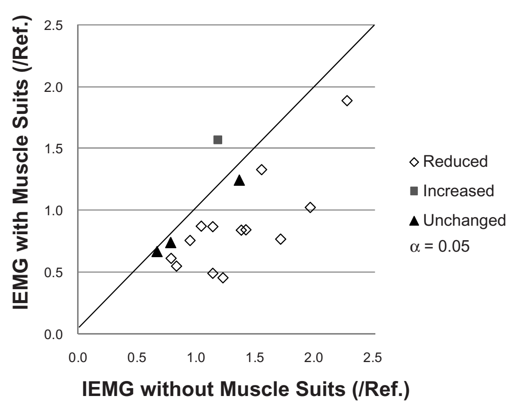


Figure 9. Comparison of muscle load on the right side of the waist between the conditions with and without muscle suit.

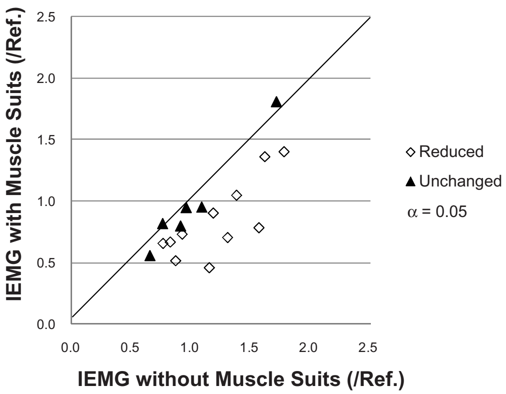


Figure 10. Comparison of muscle load on the left side of the waist between the conditions with and without muscle suit.

hold their own upper body weight with the waist, so that excessive load is placed on the waist muscles. However, when using the muscle suit, part of their upper body weight is held by the muscle suit, and the muscle load on the waist is mitigated.

5 CONCLUSION

In this study, we explored the possibility of expanding the field of application for wearable robots that assist their users' muscles. One such device is the muscle suit, which assists in tasks that requires a fair amount of muscle load repeatedly as well as tasks where significant muscle load is temporally required. In particular, we apply the muscle suit to the task of changing sheets, which is a necessary and frequent task that occurs in daily nursing-care situations, and we evaluate the muscle load on the shoulders and waist through an experiment in which we use EMG to record the muscle load. The results demonstrate that wearing the muscle suit can reduce caregivers' muscle load on the waist without increasing the load on the shoulders. Furthermore, we find that it helps caregivers naturally stabilize their position so that the weight is evenly distributed on each side of the waist. Thus, we believe that the muscle suit could effectively prevent backaches, which are caused by repeatedly assuming an unnatural body position such as bending forward or twisting.

ACKNOWLEDGEMENT

We thank Dr. Hiroshi Kobayashi, who designed and developed muscle suit, and gratefully acknowledge the help received from Ms. Yoko Seki and Mr. Yoshiaki Numa.

REFERENCES

Best, J.T. 2002, Understanding spinal stenosis. *Orthopaedic nursing* 21(3): 48–56.

Billinghurst, M., Weghorst, S. & Furness III, T. 1997. Wearable computers for three-dimensional CSCW. *First International Symposium on Wearable Computers* 39.

Bjorn, G. & House, D. 2003. Multimodality and speech technology: verbal and non-verbal communication in talking agents. *Eurospeech*: 2901–2904.

Cameron, B. 1967. The OR nurse and backache. *AORN* 6(5): 40–52.

Campbell, R., 1984. Nursing homes and long-term care in Japan. *Pacific Affairs* 57(1): 78–89.

Chou, C.P. & Hannaford, B. 1996. Measurement and modeling of McKibben pneumatic artificial muscles. *IEEE transactions on robotics and automation* 12(1): 90–102.

Daerden, F. & Lefeber, D. 2001. The concept and design of pleated pneumatic artificial muscles. *International Journal of Fluid Power* 2(3): 41–50.

Eskiizmirliler, S. 2002. A model of the cerebellar pathways applied to the control of a single-joint robot arm actuated by McKibben artificial muscles. *Biological Cybernetics* 86(5): 379–394.

Grove, A.T. 1980. Geomorphic evolution of the Sahara and the Nile. In M.A.J. Williams & H. Faure (eds), *The Sahara and the Nile*: 21–35. Rotterdam: Balkema.

Granata, K.P. 2004. Influence of fatigue in neuromuscular control of spinal stability, *The Journal of the Human Factors and Ergonomics Society* 46(1): 81–91.

Guha, M. 2007. The psychology of aging. *Journal of Mental Health*, 16(5): 691–694.

Jappelli, R. & Marconi, N. 1997. Recommendations and prejudices in the realm of foundation engineering in Italy: A historical review. In Carlo Viggiani (ed.), *Geotechnical engineering for the preservation of monuments and historical sites*; *Proc. intern. symp., Napoli, 3–4 October 1996*. Rotterdam: Balkema.

Johnson, H.L. 1965. Artistic development in autistic children. *Child Development* 65(1): 13–16.

Kishi, R., Kitahara, T., Masuchi, A. & Kasai, S. 2002. Work-related reproductive, musculoskeletal and mental disorders among working women. *Industrial Health* 40(2): 101–112.

Kobayashi, H. & Nozaki, H. 2007. Development of muscle suit for supporting manual worker. *Proceedings of the 2007 IEEE/RSJ international conference on intelligent robots and systems* (CD-ROM).

Kobayashi, H., Suzuki, H. & Nozaki, H. 2006. Development of muscle suit. *Japan society of mechanical engineers symposium on welfare engineering* 1: 244–247.

Kobayashi, H. & Tsuji, T. 2007. Development of power assist system for caregiver by muscle suit. *Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation*: 180–185.

Kobayashi, H., Uchimura, H.A., Isihida, Y., Shiiba, T., Hiramatsu, K., Konami, M., Matsushita, T. & Sato, Y. 2004. Development of a muscle suit for the upper body—realization of abduction motion, *Advanced Robotics* 18(5): 497–513.

Nakanishi, M. & Yamamoto, S. 2009. Evaluation of muscle suit in designing processes from ergonomic viewpoint—Applicability to nursing care. *Journal of Ergonomics in Occupational Safety and Health* 10(1): 13–23.

Okazaki, M. 1998. Basic Nursing Technique (1). *Medical Friend*: 38–44.
Owen, B.D. 1980. How to avoid that aching back. *The American Journal of Nursing* 80(5): 894–897.
Peter, P.M., Edward, J.D. & Martin, S. (eds) 1997. *Quality care for elderly people.* Chapman and Hall, London.
Polhill, R.M. 1982. *Crotalaria in Africa and Madagascar*. Rotterdam: Balkema.
Pope, M.H., Goh, K.L. & Magnusson, M.L. 2002. Spine ergonomics. *Annual Review of Biomedical Engineering* 4: 49–68.
Roupa, Z., Vassilopoulos, A., Sotiropoulou, P., Makrinika, E., Noula, M., Faros, E. & Marvaki, C. 2008. The problem of lowe back pain in nursing stuff and its effect on human activity. *Health Science Journal* 2(4): 219–225.
Sankai, Y. 2006. Leading edge of cybernics: Robot suit HAL. *SICE-ICASE international joint conference*: 1–2.
Suzuki, H., Kobayashi, H. 2004. Development of a muscle suit. *Proceedings of the human interface symposium* (CD-ROM).
Tanaka, T., Satoh, Y., Kaneko, S., Suzuki, Y., Sakamoto, N. & Seki, S. 2008. Smart suit: Soft power suit with semi-active assist mechanism—Prototype for supporting waist and knee joint. *International conference on control, automation and systems 2008* (CD-ROM).
Thompson, C.W. & Floyd, R.T. 1998. *Manual of structural kinesiology*. McGraw-Hill Education (ISE Editions); International 13 Revised.
Tondu, B., Lopez, P. 1997. The McKibben muscle and its use in actuating robot arms showing similarities with human arm behavior. *Industrial Robot* 24(6): 432–439.
Yamamoto, K., Hyodo, K., Ishii K. & Matuo, T. 2002. Development of power assisting suit for assisting nurse labor. *JSME International Journal Series* 45(3): 703–711.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Thermal stimulation presenting method to develop a deaf-blind communication device

Yasunori Iwasaki, Hiroshi Horio & Chikamune Wada
Kyushu Institute of Technology, Wakamatsu-ku, Kitakyushu, Japan

ABSTRACT: In order to develop a communication device which the deaf-blind person can use to handwrite characters on their palm, we have been studying apparent movement induced by cold stimulus on the palm. The results of our previous study showed that apparent movement was not induced on some areas of the palm. We considered that the reason for this was that the characteristics of cold tactile sensation differed over the palm area. In this paper, we therefore proposed and investigated a new cold stimulation presenting method based on the cold tactile characteristics of the palm. In the first experiment, we measured response time as a cold characteristic; that is, the time taken by subjects to feel coldness after a cold sensation was presented. Next, we compared our two (former and new) methods. From the results, it was found that the frequency of apparent movement increased when using the new method.

Keywords: deaf-blind, apparent movement, cold tactile characteristic, stimulus duration, palm

1 INTRODUCTION

The deaf-blind suffer from both visual and auditory impairment. The communication methods they use vary according to the nature of each person's disability and there are several methods available, such as handwriting or finger braille, sign language, and so on. In order to further facilitate effective communication between the deaf-blind and the non-disabled, we would like to develop a communication support device which can handwrite kana (Japanese syllabic characters) on a human palm.

When we write a character on the palm, the finger moves continuously to write the character on the palm. If the tactile stimulus were able to move continuously on the palm, handwriting on the palm would be realized. Incidentally, if a mechanism such as an X-Y plotter were used, then handwriting would be realized easily. However, we thought it would be difficult to make a device which was small and lightweight. To solve this drawback, we proposed and investigated the use of apparent movement of cold stimulus induced by Peltier element because 1) the Peltier element is both small and light, 2) apparent movement can make continuous movement virtually, and 3) the density of the cold spot is higher than that of the warm spot. If virtual continuous movement for cold sensation could be realized, it could cause tactile sensation in the areas where Peltier elements do not exist. Therefore, the number of elements required would decrease and the device would be lighter.

According to our previous study, apparent movement by cold stimulus was not induced in some areas of the palm. We considered that the reason for this was that the characteristics of cold tactile sensation differed over the palm area (Horio et al. 2007). Because our previous tactile stimulation presenting method (defined as the "former method") did not consider the characteristics of the cold threshold which varied over the area of the palm, the frequency of induced apparent movement varied over the area of the palm.

In this paper, we propose a new tactile stimulation presenting method (defined as the "new method"), which calibrates the stimulus condition with the area of the palm.

First, we measured the characteristics of cold tactile sensation on the palm. Next, we compared the frequency of induced apparent movement of the former method with that of the new method. From these results, we were able to ascertain the efficacy of our new method.

2 MEASUREMENT OF CHARACTERISTICS OF COLD SENSATION

We measured the characteristics of cold sensation on the palm. In this study, we measured the time duration (defined as "response time") before subjects started to feel coldness after a cold stimulus was presented to the palm surface. We considered this response time as the cold threshold.

2.1 *Experimental set-up*

The measurement system is shown in Figure 1 and the cold sensation stimulator is shown in Figure 2. The Peltier element was controlled by a signal from the computer through a D/A converter and amplifier. One thermocouple was attached to the Peltier element (8 mm × 8 mm) to measure temperature of the Peltier element's surface (left in Figure 2). A second thermocouple was placed on the palm surface to measure the temperature of the palm surface (right in Figure 2). This second thermocouple was set 5 mm apart from the Peltier element, in order to prevent the Peltier element from influencing the results. The voltage of two thermocouples was converted to temperature text data by the I-7018 (ICP DAS Co.). The sampling rate was 10 Hz. When the Peltier element did not present a cold stimulus, the temperature of the Peltier element was controlled to make it equal to the temperature of the palm surface.

While measuring response time, the cold sensation stimulator was pressed upon the palm moderately without pain.

2.2 *Experimental method*

Five subjects (all males, average age 23.6 years) took part in this experiment. All subjects gave their informed consent to participate in this experiment.

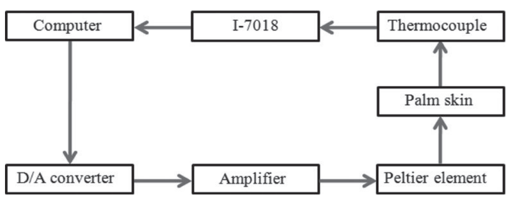


Figure 1. Measurement system used to calculate response time to cold sensation.

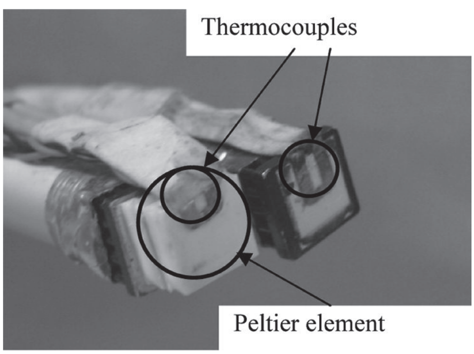


Figure 2. Cold sensation stimulator.

The measurement area used to calculate response time is shown in Figure 3. The cold stimulus was presented to the right palm of all subjects. First, a line was drawn between the metacarpophalangeal joints of the second metacarpal bone and the fifth metacarpal bone (handbreadth line). Next, a line was drawn to pass vertically through the midpoint of the handbreadth line. These two lines were used as basis lines. Then, a grid (8 mm × 8 mm) was made based on the basis lines. This grid was used as the measurement area in our calculations of response time. The number of measurement areas was 72 (9 rows and 8 columns, for four subjects) or 63 (9 rows and 7 columns, for one subject).

The applied voltage used to control the temperature of the Peltier element was constant so the temperature decreased at a constant rate. In this experiment, the voltage was −3V and the temperature decreased at a rate of −5 degrees Celsius per second when the Peltier element was not attached to the palm. This condition was the same as that in our previous study.

During this experiment, each subject was asked to turn their palm upwards and the cold sensation stimulator was placed on one area of the grid. Before the experiment, each subject was asked to wait until the temperature of the Peltier element became equal to that of the palm surface. After the temperature was equal to that of the palm temperature, the experiment started with a start signal being given by the subject When the subject felt coldness, the subject was asked to press a key on the computer keyboard as soon as possible. The time duration (response time) was measured when subjects felt coldness after the cold stimulus was presented to the palm surface. This procedure was executed on all areas of the palm and 10 trials were done for each area.

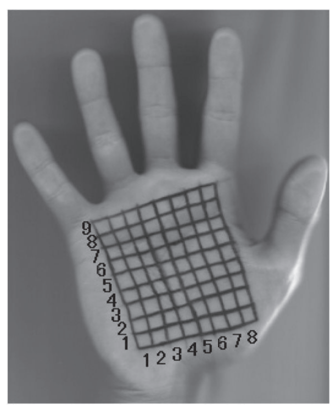


Figure 3. Measurement area.

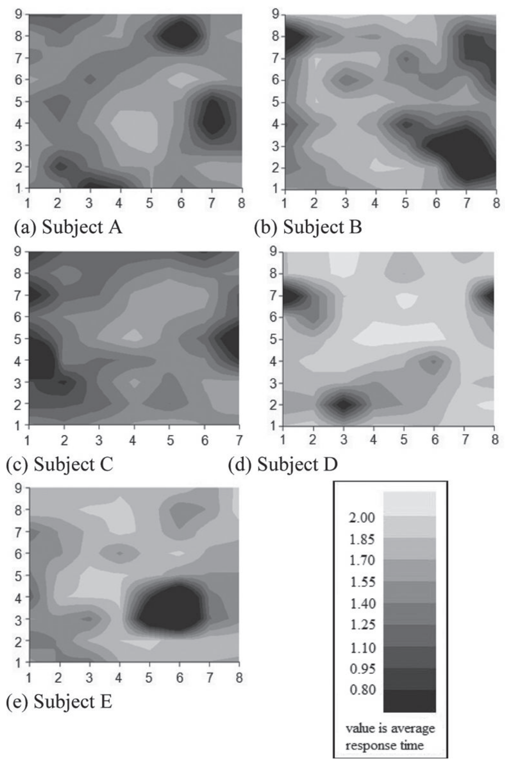


Figure 4. Results of response times to cold sensation.

2.3 *Results*

The response times for each subject are shown in Figure 4. The vertical axis shows the lengthwise direction (1st–9th row) of the palm; the 1st row is on the wrist side. The horizontal axis shows the crosswise direction (1st–8th column or 1st–7th column) of the palm; the 1st column is on the fifth finger side. The legend shows a shading scale which corresponds to the average response time in the second of 10 trials on each area of the palm. When the mean time became shorter, the standard deviation of average response time became shorter. Whereas, when the mean time became longer, the standard deviation became longer. From this figure it can be seen that the response time was different among areas and among subjects. For each subject, the shortest response time was under 1 second and longest response time was over 2 seconds.

3 COMPARISON BETWEEN FORMER METHOD AND NEW METHOD

We proposed a new stimulation presenting method based on the results of this experiment. We investigated the efficacy of the new method by comparing it with our former method.

3.1 *Tactile stimulation presenting method for apparent movement*

In our former stimulation presenting method which was used our previous study, two Peltier elements were placed on the palm with a certain distance between them (see Figure 7). The details of this former presenting method are shown in Figure 5, in which "1" and the solid line denotes the temperature change of the first Peltier element and "2" and the dotted line denotes the temperature change of the second Peltier element. In addition, "a" denotes Stimulus Onset Asynchrony (SOA), which is the time interval between the time when the first Peltier element started presenting the cold stimulus and the time when the second Peltier element started presenting the cold stimulus; "b" denotes stimulus duration, which is time duration used to present the cold stimulus. Thus, in this experiment, the first Peltier element presented the cold stimulus for "b" seconds and the second Peltier element began to present the cold stimulus for "b" seconds after "a" seconds had passed from the first Peltier element presenting the cold stimulus, then, apparent movement was induced. In this experiment, "b" was set to 3 seconds.

From the results of chapter 2, we found that when the same time duration was presented to all areas, the magnitude of perceived stimulus differed with each area. In former method of our previous study, the time durations for all areas were the same. Therefore, because the magnitude of perceived stimulus differed, we thought apparent movement was not induced in some areas in our previous study using former method. We therefore proposed a new method based on response time (shown in Figure 6).

As stated above, in the former method, "b" was fixed to 3 seconds for all areas. However, in the new method, the response time for each area (c1, c2) was added to "b" and "c1" and "c2" which were recorded (see Figure 4). Figure 6 shows that the

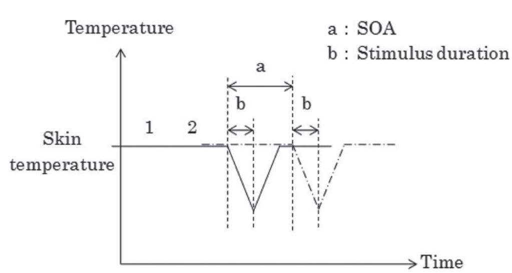


Figure 5. Former stimulus presenting method.

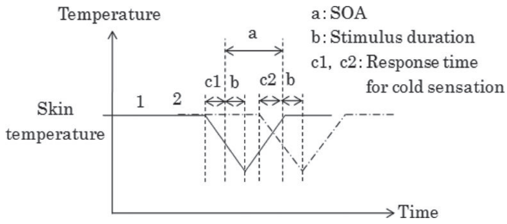


Figure 6. New stimulus presenting method.

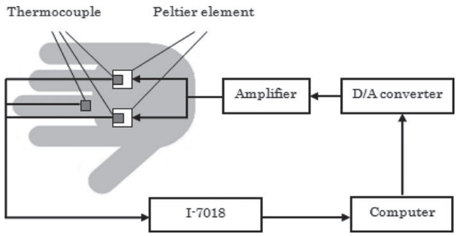


Figure 7. Experimental set-up for apparent movement.

first Peltier element's time duration was "b + c1" and that the second Peltier element's time duration was "b + c2". Therefore, we hypothesized that the subject could feel the cold stimulus for 3 seconds in all areas.

3.2 *Experimental set-up*

The experimental set-up which was used in former method and new method is shown in Figure 7. The method used to control the temperature of the Peltier element was essentially the same as that used for the experiment described in chapter 2. However, a thermocouple was set on the base of the third finger to prevent the Peltier elements from influencing the results. The other temperature control method was the same as that described in chapter 2. In this experiment the SOA was varied from 0.4 second to 6.0 seconds every 0.4 second.

3.3 *Subjects*

The five subjects were the same as those described in chapter 2.

3.4 *Stimulation patterns of cold stimulus*

To investigate the difference between the former method and the new method, we prepared four stimulation patterns of cold stimulus based on the results of chapter 2. One of these results is shown in Figure 8. This is the result for subject A. The vertical axis and horizontal axis are the same as in Figures 3 and 4. The average response time was calculated from all data shown in Figure 8. In the case of subject A, the average response time was 1.43 seconds. Data shorter than 1.43 seconds is hatched, whereas data longer than 1.43 seconds is non-hatched. Four stimulation patterns were prepared as follows:

1. Two response times ("c1" and "c2" in Figure 6) for two Peltier elements were selected from hatched data;
2. Two response times ("c1" and "c2" in Figure 6) for two Peltier elements were selected from non-hatched data;
3. First response time "c1" was selected from hatched data and second response time "c2" was selected from non-hatched data; and
4. Response time "c1" was selected from non-hatched data and "c2" was selected from hatched data.

These four stimulation patterns are summarized in Table 1.

9	1.8	1.7	1.6	1.2	1.3	1.7	1.4	1.4
8	1.3	1.5	1.3	1.2	1.7	2.8	1.3	1.2
7	1.5	1.3	1.4	1.5	1.4	1.3	1.3	1.4
6	1.2	1.4	1.6	1.3	1.2	1	1.2	1.3
5	1.5	1.7	1.4	1.2	1.2	1.3	2.1	1.6
4	1.4	1.5	1.3	1	1	1.3	2.3	1.6
3	1.3	1.4	1.5	1.1	1	1.4	1.7	1.7
2	1.2	2	1.5	1.3	1.2	1.3	1.1	1.2
1	1.4	1.4	2.2	1.9	1.2	1.5	1.3	1.2
	1	2	3	4	5	6	7	8

Figure 8. Response time for cold sensation of subject A.

Table 1. Stimulation patterns and examples.

	First stimulus	Second stimulus	An example
①	Below average	Below average	9th row (7th column and 4th column)
②	Above average	Above average	8th row (5th column and 2th column)
③	Below average	Above average	6th row (6th column and 3th column)
④	Above average	Below average	3th row (7th column and 4th column)

3.5 *Estimation of induced apparent movement*

Each subject was asked to estimate induced apparent movement according to the following procedure: when the subject felt strong/weak apparent movement, he was asked to answer “2 points”/“1 point”, respectively. If he did not feel apparent movement, he was asked to answer “0 points”.

3.6 *Experimental results in the case of the former method*

Subjects were split into two groups (A and B) according to the results. In the case of group A, apparent movement was not induced when SOA became 0 seconds, but apparent movement was induced strongly when SOA ranged from 1.6 to 2.4 seconds. Four subjects were in group A. In the case of group B, apparent movement was induced strongly when SOA ranged from 0 to 1.6 seconds. One subject was in group B.

The average score, i.e., frequency of induced apparent movement was calculated. The average score is shown in Figures 9 and 10 for groups A and B, respectively. In both figures, the vertical axis shows the average score and the horizontal axis shows the SOA. From these graphs, it can be seen that the percentile frequency rate when the subject’s answer was “2 points” was obtained when the graph had its peak. The percentile frequency rate was defined as “induced rate” and the SOA when the graph had its peak was defined as “Scorepeak”. In the case of stimulation pattern (1) for group A, the Scorepeak was 1.7 points when the SOA was 1.6 seconds (see Figure 9). The induced rate was 71%, whereas the frequency rate for both “1 point” and “0 points” was 29%. When the induced frequency of apparent movement was over 60%, apparent movement was thought to be induced completely, because the induced frequency of apparent movement for tactile stimulus with optimum conditions is 60% (Hulin 1927).

In the case of group A, for pattern (1), the Scorepeak was 1.7 points and the induced rate was 71%. It was therefore assumed that apparent movement was induced completely. On the other hand, the Scorepeak was 1.3 points and rate was 23% for pattern (2), the Scorepeak was 1.3 points and the rate was 28% for pattern (3), and the Scorepeak was 1.4 points and the rate was 43% for pattern (4). From these results, we concluded that apparent movement was hardly induced for patterns (2), (3) and (4).

In the case of group B, for pattern (1), the Scorepeak was 1.6 points and the induced rate was 60%, and apparent movement was therefore assumed to have been induced completely. On the other hand,

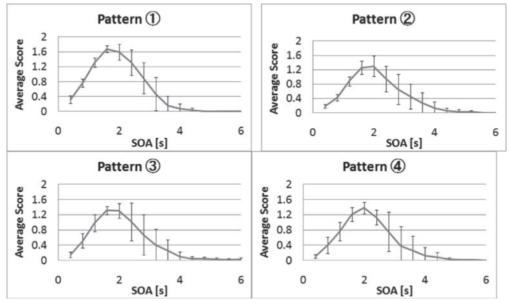


Figure 9. Experimental results of induced apparent movement with former method for Group A.

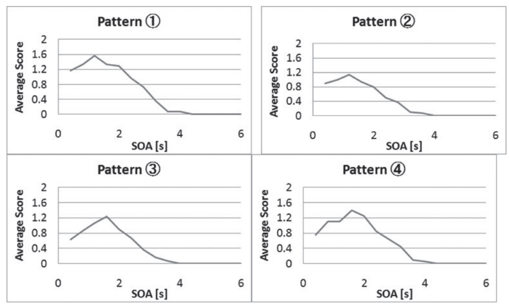


Figure 10. Experimental results of induced apparent movement with former method for Group B.

the Scorepeak was 1.1 points and the rate was 25% for pattern (2), the Scorepeak was 1.2 points and the rate was 30% for pattern (3), and the Scorepeak was 1.4 points and rate was 40% for pattern (4). We concluded that apparent movement was hardly induced for these patterns.

From these results it can be seen that, when using the former method, apparent movement was induced completely for pattern (1). However, apparent movement was hardly induced for the other patterns. We therefore investigated whether the induced rate increased or not when the new method was used for stimulation patterns (2), (3) and (4). Our findings are presented in the next section.

3.7 *Experimental results in the case of new method*

We adopted the new method for stimulation patterns (2), (3) and (4) and investigated the induced rates. The results for groups A and B are shown in Figures 11 and 12, respectively. In both figures, the vertical axis shows the average score and the horizontal axis shows the SOA. The solid line shows the results obtained by using the former method and the dotted line shows the results obtained by

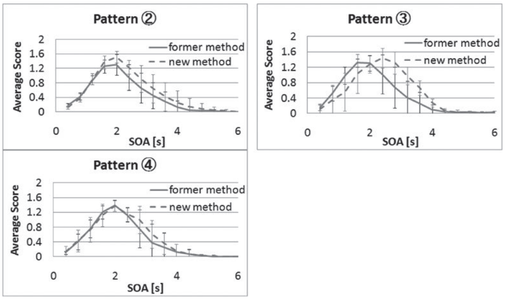


Figure 11. Experimental results of induced apparent movement with new method for Group A.

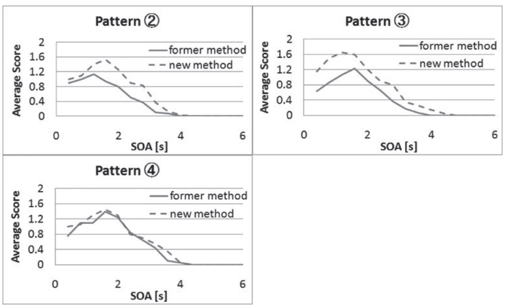


Figure 12. Experimental results of induced apparent movement with new method for Group A.

using the new method. These graphs show the induced frequency rate calculated for each stimulation pattern and each group.

In the case of group A, for pattern (2), the Score peak increased to 1.5 points and the induced rate increased to 36%, but apparent movement was not thought to be induced completely. Moreover, the Scorepeak was 1.3 points and the rate was 28% for pattern (3), and the Scorepeak was 1.4 points and the rate was 43% for pattern (4). These induced rates were almost the same as those achieved when using the former method.

In the case of group B, the Scorepeak increased to 1.5 points and the induced frequency rate increased to 57% for pattern (2), the Scorepeak increased to 1.6 points and the rate increased to 47% for pattern (3), and the Scorepeak increased to 1.5 points and the rate increased to 55% for pattern (4). Thus, the induced rate became larger than when the former method was used. However, apparent movement was not considered to be induced completely.

From these results, we concluded that apparent movement was induced effectively by using the new method only when stimulation pattern (2) was used in the case of group A. However, apparent movement was not induced completely because the induced frequency rate was under 60%. As for the other patterns in group A, the induced rate did not vary according to method used. In the case of group B, the induced frequency rate increased when using the new method for all patterns.

3.8 *Considerations*

3.8.1 *The influence of masking*

a. *Temporal masking*

We hypothesized that the difference in graphs between groups A and B was due to the influence of masking. The time interval was short between two stimulations when SOA was short. Therefore, we concluded that group A did not exhibit induced apparent movement at SOA ranged from 0 to 1.6 seconds because of the effect of masking. On the other hand, the induced apparent movement for group B at SOA ranged from 0 to 1.6 seconds may not have been affected by masking.

b. *Masking due to differences of sensation intensity*

For patterns (3) and (4), the induced frequency rate did not vary between the new method and the former method. A possible explanation may be that the results were affected by masking due to differences of sensation intensity. In the experiment, subjects reported that "stimulus is strong in the area where response time is short, and that stimulus is weak in the area where response time is long". So, we considered that sensation intensity differed according to palm area despite the thermal stimulus being the same intensity. Thus, sensation intensity differed between two cold stimulations for patterns (3) and (4), and the stimulus which was presented to areas where response time was short may have become masked. As a result, for patterns (3) and (4), the new method was minor benefit.

3.8.2 *Possibility of adapting experimental results to elderly people*

In general, the elderly have a high thermal threshold. The cold thresholds on the palms of elderly people have been measured (Uchida et al. 2007). They measured the thermal threshold of body parts of subjects aged in their 20 s, 60 s, 70 s and 80 s. The results show that subjects in their 20 s perceived a −1 degree Celsius change in skin temperature, whereas subjects in their 60 s, 70 s and 80 s perceive a −1.5 degrees Celsius change in skin temperature. We consider that our new method could be used for the elderly people, because our new method presented cold stimulus over 3 seconds with a cooling rate of −4 degrees Celsius per second, and provided a cold stimulus which changed skin temperature by over −10 degrees Celsius.

4 CONCLUSION

In this study, we investigated a stimulation presenting method to induce apparent movement by cold stimulus in order to develop a communication support device which could enable the deaf-blind to handwrite kana on their palm.

In the first experiment, we measured a cold tactile characteristic, namely the time taken by subjects to feel coldness after a cold sensation was presented to about 70 areas on the palm. From the results, it was found that response time varied with palm areas and subjects.

In the second experiment, we proposed a new stimulation presenting method that considered response time, and compared the induced rate of apparent movement of the former method with that of the new method. From the results, it was found that the frequency of apparent movement increased under certain experimental conditions when using the new method.

In order to realize the development of a new communication device for the deaf-blind, we will study the influence of sensation intensity and masking on cold stimulus, and propose a more effective stimulation presenting method in the next experiment in order to increase the induced rate of apparent movement.

REFERENCES

Horio, H. & Wada, C. 2007. Relationship between the palmar skin's thermal perception and its mechanical characteristics. *International Conference on Control, Automation and Systems*: 1893–1896.

Hulin, W.S. 1927. An experimental study of apparent tactual movement. *Journal of Experimental Psychology* 10(4): 293–320.

Uchida, Y. & Tamura, T. 2007. Regional difference of thermal sensitivity on the skin surface of elderly. *Journal of the Japan Society of Home Economics* 58(9): 579–587.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

A questionnaire to determine nursing and child care students' understanding of welfare equipment

Tomoko Morimoto
Gifu University of Medical Science, Seki, Gifu, Japan

Shoko Kamimotono
Nara Medical University, Kashihara, Nara, Japan

Nanae Shintani
Gifu University of Medical Science, Seki, Gifu, Japan

ABSTRACT: "Welfare Equipment" is generally recognised as being tools used for disabled persons and elderly persons. However, if we were to widely view the concept to see welfare equipment as tools to support daily life, this concept could be adapted to include tools used for self-sustaining infants as well as disabled persons and elderly persons. The purpose of this paper is to understand the awareness amongst students studying nursing and childcare regarding welfare equipment and utilise this research for education in their study area.

The survey was conducted with 300 students who study at nursing and childcare universities and vocational schools in the Kinki area. The collection rate was 70%. Oral explanation of the purpose of this survey and methodology was implemented. Answer sheets were collected only from students who agreed with this purpose. 13 welfare equipment (diaper, etc.) used for both elderly persons and infants as, well as 7 (activities singing etc.) used for rehabilitation were selected for this research. Questions addressing these activities could be seen as welfare and rehabilitation factors, were asked in this survey with 5 scales. This research was conducted in September and October 2008. Mann-Whitney-U test was used to analyse each piece of equipment and activity. Significance level was set under 5%.

A trend indicating that these items are for elderly persons amongst students studying nursing in universities and studying childcare in vocational schools is recognised. As for students studying at childcare college, they recognise some of these and items to be used not only for elderly persons also for infants.

It is impossible to widely disseminate welfare equipment without recognition and knowledge of these equipments, even though we have the political policy to develop and use this equipment. It is necessary to consider creating a curriculum to make nursing and childcare students understand clearly the purpose of welfare equipment in their education.

Keywords: technical aids, nursing, childcare, education

1 INTRODUCTION

The law on development and spread for the promotion of technical aids in Japan is set as "Technical aids shall be aids to accommodate elderly persons whose physical and mental functions have decreased and who have difficulty in their daily lives ("elderly person") or people with disabilities or aids or adaptive equipment for their functional trainings". We recognize that technical aids are used for persons with disabilities or the elderly. When we recognize that technical aids are a helpful thing to support daily life, we could understand that this would apply even to children who are on the way to being self-reliant, as well as persons with disabilities and the elderly. We surveyed items on technical aids and rehabilitation in order to understand what recognition students who are studying nursing and childcare have on technical aids and to use the surveyed results for education in nursing and childcare.

2 SUBJECTS AND METHODS

The subjects consist of 80 third-grade students at a nursing university in Kinki area, 110 students who study childcare at a vocational school ("childcare vocational school students") and

110 students who study childcare at a junior college ("childcare junior college students").

The purpose and methods of the survey were explained orally and the survey form was collected from only students who gave consent on the same day.

We selected common items in the elderly and children, from the ones which are well known as technical aids, and for rehabilitation ("rehabilitation", and gain-of-function for children).

The items are 13: disposable diapers, portable toilets ("portable toilets", a "potty" for children), auxiliary toilet seats, feeding cups with a holder ("feeding cups with a holder"), chopsticks designed to be easily (correctly) held ("designed chopsticks"), chairs with a table ("chairs with a table"), dishes designed so that food can be easily scooped ("designed dishes"), spoons designed so that food can be easily scooped ("designed spoons"), chairs for bath ("bath chairs"), aprons during meals ("aprons"), thickened meals or meals that foods are cut into small pieces meals, and baby food for children ("designed meals"), clothing that is not separated into tops and bottoms ("boiler suits"), wheelchairs, and strollers for children ("wheelchair"). The following 7 items are for rehabilitation: articles made out of paper by hand, mainly Origami, and Origami for children (article made out of paper by hand), playing mainly writing and drawing, and drawing for children ("writing and drawing"), playing with hands (fingers), playing with songs ("playing with songs"), playing with play equipment ("playing with play equipment"), walk and physical contact. Abbreviations used for Tables are presented in the brackets. We asked students whether they think these 13 items and 7 items can be used as technical aids or be elements of rehabilitation when the items are used for the elderly and children. The reply to the question was set to 5-stage scale: 1: Disagree, 2: Slightly disagree, 3: I neither agree nor disagree, 3: Strongly agree, and 5: Agree. The survey period was September and October in 2008.

The collection ratio was 95% of nursing university students, 48% of childcare vocational school students and 69% of childcare junior college students.

We used (1) SPSS as a discussion method and employed the Mann-Whitney-U test as to replies to the elderly and children from each subject group. Items where differences were observed were extracted. The significance level was set to less than 5%.

A factor analysis was conducted to observe structures of recognition of the elderly and children to (2) technical aids. After an analysis was conducted using a principle factor method, varimax rotation was applied.

3 RESULTS AND DISCUSSION

1. Items in each group where significant differences were found are exhibited. Items where significant differences were observed in nursing university students are: disposable diapers, potable toilets, auxiliary toilet seats, feeding cups with a holder, designed chopsticks, designed dishes, designed spoons, bath chairs, designed meals and wheelchairs. The students recognized many surveyed items as elements of technical aids and for rehabilitation to the elderly. Items where significant differences were observed in childcare vocational college students are: disposable diapers, auxiliary toilet seats, holder feeding cups, designed dishes, designed spoons, bath chairs, designed meals and wheelchairs. As in the results of the nursing university students, these students also recognized the surveyed items as elements of technical aids or rehabilitation for the elderly. Items where significant differences were found in childcare junior college students are: auxiliary toilet seats, chairs with a table, bath chairs, boiler suits, wheelchairs, playing with fingers, playing with songs and playing with play equipment. Compared to the replies from nursing university students and childcare vocational school students indicating that more surveyed items are elements of technical aids or rehabilitation for the elderly, replies from childcare junior college students divided items into the ones with elements of technical aids or rehabilitation for the elderly and the ones with elements of technical aids or rehabilitation for children. An auxiliary toilet seats, bath chairs and wheelchairs can be elements of technical aids or rehabilitation for the elderly and chairs with a table, aprons, boiler suits, playing with fingers, playing with songs and playing with play equipment would be elements of technical aids or rehabilitation for children (Table 1).
2. Results of the factor analysis extracted 2 factors with the eigenvalue of 1 or higher in the recognition to technical aids for children. Table 2 shows items with factor loading of 0.5 or more to each factor. On the other hand, the recognition to technical aids for adults extracted 4 factors with the eigenvalue of 1 or higher. Table 3 shows items with factor loading of 0.5 or more to each factor.

"Interpretation of factors" (1) The first factor of structures of recognition to technical aids for

Table 1. Technical aids and rehabilitation where significant differences were observed by subject groups.

Subject group / Items	Nursing university students		Childcare students Vocational school		Childcare students Junior college	
	level of significance	Group where sum total of orders of replies is large	level of significance	Group where sum total of orders of replies is large	level of significance	Group where sum total of orders of replies is large
Disposable diapers	**	Elderly	**	Elderly		
Portable toilets	*	Elderly				
Potty seat	**	Elderly	**	Elderly	**	Elderly
Feeding cups with a holder	*	Elderly	**	Elderly		
Designed chopsticks	**	Elderly				
Chairs with a table					*	Children
Designed dishes	**	Elderly	**	Elderly		
Designed spoons	**	Elderly	*	Elderly		
Bath chairs	**	Elderly	**	Elderly	**	Elderly
Aprons						
Designed food	**	Elderly	*	Elderly		
Overalls					**	Children
Wheelchair	**	Elderly	**	Elderly	**	Elderly
Article made out of paper by hand						
Writing and drawing						
Playing with fingers					**	Children
Playing with songs					*	Children
Playing with play equipment					**	Children
Walk						
Physical contact						

** $p < 0.01$, *$p < 0.05$

Table 2. Factor matrix after application of varimax rotation (children).

	Extracted factors		n = 300
Items	1	2	Communality
Playing with songs	.906	.330	.931
Playing with play equipment	.900	.283	.891
Playing with fingers	.881	.315	.876
Physical contact	.879	.287	.855
Drawing	.863	.321	.848
Walk	.857	.372	.873
Designed spoons	.264	.773	.666
Bath chairs	.118	.627	.408
Aprons	.445	.598	.555
Baby food	.546	.553	.604
Strollers	.417	.529	.453
Contribution ratio (%)	57.713	10.013	
Cumulative contribution ratio (%)	57.713	67.726	

Table 3. Factor matrix after application of varimax rotation (adults).

	Extracted factors				n = 300
Items	1	2	3	4	Communality
Playing with songs	.818	.140	.119	.266	.774
Playing with fingers	.789	.162	.169	.146	.699
Writing and drawing	.770	.171	.109	.224	.684
Article made out of paper by hand	.728	.229	.007	.234	.637
Playing with play equipment	.696	.189	.144	.014	.541
Designed spoons	.083	.895	.164	.194	.873
Designed dishes	.141	.881	.155	.189	.855
Chairs with a table	.244	.632	.128	–.159	.501
Designed chopsticks	.173	.584	.308	–.158	.491
Feeding cups with a holder	.166	.533	.369	–.033	.449
Wheelchair	.121	.110	.668	.153	.496
Disposable diapers	–.045	.079	.579	–.047	.346
Portable toilets	.134	.224	.555	–.017	.377
Designed food	.254	.334	.525	.101	.461
Walk	.506	.049	.104	.781	.880
Physical contact	.505	.046	.096	.657	.699
Contribution ratio (%)	35.576	14.088	7.529	5.741	
Cumulative contribution ratio (%)	35.576	49.664	57.193	62.934	

children makes up 58%. Items with factor loading of 0.5 or more are “playing with songs”, “playing with play equipment”, “playing with fingers”, “physical contact”, “drawing” and “walk” and we named the items together as “playing and physical contact”. The second factor makes up 10%. Items with factor loading of 0.5 or more are “designed spoons”, “bath chairs”, “aprons during meals”, “baby food”, and “strollers”. We named the items together as “childcare aids”.

(2) The first factor of structures of recognition to technical aids for adults makes up 36%. The items with factor loading of 0.5 or more are “playing with songs”, “playing with fingers”, “writing and drawing”, “article made out of paper by hand” and “playing with play equipment” and we named the items together as “playing”. The second factor makes up 14% and the items with factor loading of 0.5 or more are “designed spoons”, “designed dishes”, “chairs with a table”, “designed chopsticks”, and “feeding cups with a holder”. We named the items together as “meal supportive aids” The contribution ratio of each third factor and fourth factor was 7% and 6% respectively and we determined to omit their results.

The results of this survey present similar tendency in nursing university students and childcare vocational school students on recognizing items as technical aids or rehabilitation for the elderly. Although significant differences are observed in replies from childcare junior college students, they recognized some items as elements of technical aids or rehabilitation for children as well as for the elderly. We did not obtain structures or details of the curriculum provided to the subject students. As recognition by childcare junior college students are found in items for the elderly and children, these differences in recognition by subjects seemed to be affected by the education, including the curriculum.

According to the results of the factor analysis, "playing" appears to be an important factor for recognition to technical aids for children and adults.

4 CONCLUSION

Although development or use of technical aids is specified in a policy measure, it is impossible to recommend the aids without recognition or knowledge of the aids. The results of this paper are obtained from the survey targeting students who study nursing or childcare, and indicate that items of technical aids and rehabilitation vary depending on education styles.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Relationship between depersonalization and experience of filling out incident report

Miyuki Sugiura & Masataka Hirosawa
Juntendo University of Health and Sports Science, Chiba, Japan

Hideko Aida
Juntendo University Hospital, Tokyo, Japan

Aya Okada
Juntendo University Nerima Hospital, Tokyo, Japan

Motoki Mizuno1, Yasuyuki Yamada & Kawata Yujiro
Juntendo University of Health and Sports Science, Chiba, Japan

Hiroshi Ihara
Dokkyo Medical University Koshigaya Hospital, Saitama, Japan

Yukihiro Kobayashi
Juntendo University Hospital, Tokyo, Japan

ABSTRACT: The purpose of this study is to investigate the possible relationship between depersonalization and medical errors. Participants were 721 female nurses (mean age: 28.65 ± 6.69) from two general hospitals. Participants completed questionnaires which included the Cambridge Depersonalization Scale (CDS), questions regarding experience of filling out incident/accident report and socio-demographic data. The investigation was performed in 2009. In this study, the prevalence of possible depersonalization disorder, estimated by the CDS, was 4.6%. Scores of global CDS, "anomalous body experience (CDS subscale)", "emotional numbing (CDS subscale)" and "alienation from surroundings (CDS subscale)" in the nurses who have experience of filling out incident/accident report were higher than those of the nurses who have no experience filling out the report ($p < 0.01$) but there was no significant difference among two groups regarding "anomalous subjective recall (CDS subscale)". These results suggested that depersonalization is possibly one of the factors related to medical error.

Keywords: depersonalization, medical error, incident report

1 INSTRUCTIONS

Medical errors made by medical staff members not only have a negative influence on their professional and personal status, confidence, and function, but also bring to patients serious risks and increase healthcare costs (Kohn et al., 2000, West et al., 2006).

Depersonalization is characterised by persistent or recurrent episodes of 'detachment or estrangement from one's self.' Although the individual may feel like an automaton or there may be sensation of being an outside observer of one's own mental process, patients have enough ability to recognize the real world during the depersonalization experience (American Psychiatric Association, 2000). Common symptoms of depersonalization can be categorized into aspects of visual depersonalization, altered body experience, emotional numbing, loss of agency feeling, and changes in subjective experiencing of time (Sierra & Berrios, 2001).

It is well known that depersonalization accompanies most mental diseases, emotional stress, somatic diseases, and exhaustion (Nuller, 1982). Considering the fact that the nursing profession is both cognitively taxing and stressful, it may be an occupation which contains the risk of experiencing depersonalization. Some of them may develop symptoms of clinical depersonalization disorder (Hirosawa et al., 2009).

In a previous study (Hirosawa et al., 2009), we had opportunities to treat 10 nurses who were advised to consult a psychiatrist by their supervisor due to frequent incidents and/or accidents. In the clinical interview, they mentioned that their errors occurred under depersonalization experience. In those cases, depersonalization was considered one of the critical factors which had an influence on their errors. Moreover, if their experience of medical accidents resulted in severe conditions for patients, it may have been so stressful for them that they developed a higher likelihood of experiencing more severe depersonalization. However, until now, there has been no empirical study which has verified the relationship between depersonalization and medical errors. The purpose of this study was to investigate the relationship between depersonalization and experience of filling out incident/accident report among nurses working in general hospitals.

Regarding the term "depersonalization," it is also used as one of the domains of burnout syndrome. According to the Maslach Burnout Inventory (MBI) (Maslach et al., 1996). which is widely used to estimated burnout syndrome, depersonalization is characterized by negative, cynical attitudes and feelings about one's recipients and it is composed of items such as; "I feel I treat some recipients as if they were impersonal object". The term depersonalization, which is used in identifying burnout syndrome, seems to represent misanthropic attitudes (Ebihara et al., 1979) and so it is different from the clinical term of depersonalization which relates to depersonalization disorder (American Psychiatric Association, 2000, World health Organization, 1992) In the present study, we use the term depersonalization, only as it is described in depersonalization disorder in DSM-IV-TR.

2 METHODS

2.1 *Participants*

Participants were 1379 female nurses from two university hospitals; located in the Tokyo metropolitan area and in the suburbs of Tokyo. The study was approved by the ethics committee of Juntendo University Hospital. Nurses participated in this investigation after informed consent was obtained. The investigation was performed from 2008 to 2009. 721 nurses completed in this study. The mean age was 28.65 years (SD = 6.69).

2.2 *Measures*

2.2.1 *The Cambridge Depersonalization Scale (CDS) (Sierra & Berrios, 2000)*

The CDS comprehensive instrument contains 29 items addressing the complaints associated with depersonalization disorder based on a comprehensive study of the phenomenology of this condition. Each item was rated on two Likert scales for frequency and duration of experience, and the sum of these two scores generated an index of item intensity (range, 0–10). The global score of the scale was the arithmetical sum of all items (range, 0–290). It had a high internal consistency (Cronbach α and split half reliability of 0.89 and 0.92) (Sierra & Berrios, 2000). In addition, the CDS has four subscales; anomalous body experience, emotional numbing, anomalous subjective recall, and alienation from surroundings (Sierra et al., 2005). The original study aimed to distinguish inpatients with depersonalization from inpatients with panic disorder, generalized anxiety disorder and temporal lobe epilepsy and recommended a cut-off point of 70. In this study, we used a cut-off point of 60, which is more effective to distinguish depersonalization disorder from the general population under non-clinical conditions (Sugiura et al., 2009). The cut-off point of 60 was shown to yield a sensitivity of 100% and a specificity of 96% by distinguishing depersonalization disorder from healthy controls.

2.2.2 *Socio-demographic variables*

We asked participants about socio-demographic variables; age, years of nursing experience, work place (outpatient clinic or word) and job position (staff member or administrator). Administrators included chief nurses and head nurses.

2.2.3 *Experience of filling out incident/accident report*

In order to assess the experience of medical errors, we asked participants about whether or not they had experience of filling out an incident/accident report about error in the past 6 months. Here we adopted the length of 6 months in consideration of the length of depersonalization required in the CDS.

2.3 *Statistical analysis*

Statistical analysis in this study was carried out with SPSS, version 18.0. Firstly, in order to obtain fundamental information about depersonalization among nurses, we examined the prevalence of possible depersonalization disorder among nurses by using the recommended cut-off point of 60. Secondly, in order to reveal associated factors with experience of filling out incident/accident report, we carried out the t-test, Mann-Whitney U test, Chi-square test and logistic regression analysis for score of global CDS, its subscale, socio-demographic variables. All tests were two-tailed. The significance level was considered to be $p < 0.05$.

3 RESULTS

3.1 *Possible depersonalization and socio-demographic factors*

Among all participants, the prevalence of those who got more than 60 points in the CDS was 4.58%. A number of 402 (55.76%) had no experience and 319 (44.24%) had experience of filling out an incident/accident report about error in the past 6 months.

3.2 *Association of experience of filling out incident/accident report with depersonalization and socio-demographic factors*

Regarding depersonalization, Chi-square test showed that 1.99% of nurses who had no experience of filling out the incident/accident report (non-report group) and 7.84% of nurses who had experience of filling out the report (report group) had possible depersonalization disorder and there was significant association between experience of filling out the incident/accident report and possible depersonalization disorder ($\chi^2 = 13.92$, $P<0.001$). Additionally, Mann-Whitney U test showed that the report group scored significantly higher than the non-report group in the score of total CDS and "anomalous body experience," "emotional numbing" and "alienation from surroundings" ($P<0.01$), but there was no significant difference between these two groups in the score of "anomalous subjective recall."

Regarding socio-demographic variables, there were significant differences between these two groups in regard to work place and job position ($P<0.001$), but no significant difference between these two groups in regard to age and years of nursing experience.

As shown in Table1, logistic regression analysis showed that experience of filling the report was associated with possible depersonalization disorder and working in ward after adjusting for other factors (OR = 4.00, 95% CI = 1.69–8.99, $P<0.001$; OR = 2.15, 95% CI = 1.37–3.38, $P<0.001$, respectively).

Table 1. Factors related experience of filling out incident/accident report.

	B	Odds rate	(95% CI)	
age	–0.03	0.97	(0.93–1.01)	*P* = n.s.
years of nursing experience	–0.01	0.99	(0.94-1.04)	*P* = n.s.
work place				
outpatient clinic				
ward	0.77	2.15	(1.37–3.38)	*P*<.001
Job position				
staff member				
administrator	–0.50	0.61	(0.31–3.38)	*P* = n.s
CDS Score				
<60				
≥60	1.36	3.90	(1.69–9.00)	*P<.01*

Hoser and Lemeshow test, χ 2 = 9.74, df = 8, *P* = 0.28

4 DISCUSSION

4.1 *Prevalence of depersonalization disorder and its recognition*

In this study, the prevalence of possible depersonalization disorder was 4.58%, according to the recommended CDS cut-off point. The precise prevalence of depersonalization disorder among the general population is unknown, but it has been estimated at 1.7–2.4% by using diagnostic interviews (bebbington et al., 1997), and at 2.7% among Japanese university students (Sugiura et al., 2008) by using the CDS. Possible prevalence in this study is somewhat higher than among the general population and Japanese university studies.

The CDS was established by using diagnostic interview and the reliability and validity of its cut-off point is sufficient to distinguish depersonalization disorder from that among healthy people (Sierra & Berrios, 2000, Sugiura et al., 2009). Of course, there is a possibility that symptoms of some nurses who were evaluated as having depersonalization disorder according the CDS do not meet complete diagnostic criteria of depersonalization disorder but this implies that there is a considerable percent of nurses who have depersonalization disorders which need to be treated psychiatrically. Therefore, we have to consider depersonalization disorder as one of the notable mental problems which appear among nurses.

4.2 *Relationship between socio-demographic factors and medical experience of filling out an incident/accident report*

Working in a ward related to experience of filling out the report after adjusting for influence from other factors (Table 1).

Some studies have demonstrated that extended-duration work shifts among health care providers significantly increase fatigue and impair performance and safety (Lockley et al., 2007). In this study, the ward had a night duty and rotating shift, but the outpatient department, except for the emergency section, did not have a night duty and its office hours were approximately constant. Also, the condition of patients in wards is more severe than in the outpatient clinic. Therefore, these working conditions in wards may lead to medical errors more frequently than in the outpatient clinic.

4.3 *Relationship between depersonalization and experience of filling out an incident/accident report*

Possible depersonalization disorder was related to experience of filling out an incident/accident report after adjusting influence from other factors (Table 1).

It has been reported that complaints of occupational impairment are very common during depersonalization experiences (Simon, 2004). Some neuropsychological studies (Guralnik et al., 2000; Guralnik et al., 2007) have indicated that in depersonalization disorder, slower processing speed, poorer perceptual organization, vulnerability to distraction, and problems with immediate recall can be found. These subjective distresses and deficits in cognitive functioning seem to prevent nurses from accomplishing their day-to-day business and sometimes may lead to medical errors.

Additionally, it has been pointed out that, psychologically, depersonalization disorder should be exclusively experienced as a "subjective experience", and that we cannot identify it without individuals' own expressions (Yasunaga, 1987). Considering the fact that these "subjective experiences" are very difficult to express in plain language (Hirosawa, 2005), depersonalization disorder is also very difficult to recognize and identify, not only by suffering nurses but also by their colleagues, if they are not well informed about this condition. Moreover, it has been pointed out that daily behaviors of individuals with depersonalization may not necessarily appear strange (Yasunaga, 1987). Some of those individuals showed a high capacity at work (Matumoto, 1998) and their overall intelligence remained normal (Guralnik et al., 2007). Accordingly, some nurses with depersonalization disorder were considered by others to be performing their daily nursing work adequately. These findings indicate that colleague nurses will often regard those nurses who have depersonalization disorder as being healthy enough to achieve their daily work (Hirosawa et al., 2009). For this reason, support systems from co-workers and supervisors regarding medical error may not function adequately.

On the other hand, experience of filling out an incident/accident report gives a great shock to the nurses concerned and it is considered to be one of the major causes of stress for nurses. Regarding the repeatedly described characteristics of depersonalization that are often accompanied by psychological stress (Nuller, 1982) some depersonalization disorder may result from stress after an error has occurred. Therefore, when an error occurs, we need to pay particular attention to the appearance of depersonalization disorder and depersonalization-like experience among nurses and to help prevent them from making further errors.

Regarding the symptoms of depersonalization disorder in the CDS, scores of "anomalous body experience," "emotional numbing" and "alienation from surroundings" were significantly higher in the report group than in the non- report group.

The "anomalous body experience" assessment was composed of items regarding the feeling of out of body experiences, loss of agency feeling, and automatic and mechanical body feeling (Sierra et al., 2005). Persons experiencing these symptoms may feel as if their own body is moving automatically and they are not able to control their own behaviour. This may lead error associated with experience of automaton and having uncontrollable movement.

The "emotional numbing" assessment was composed of items regarding numbing of affection accompanied with own behaviour, affection for close person and sensation of taste and smell (Sierra et al., 2005). Loss of familiar affection toward patients may lead to the loss of prudent attitude toward action. Numbing of affection may make it difficult for them to notice emergency signs and to react immediately and adequately to these signs.

The "alienation from surroundings" experience assessment was composed of items regarding feelings of not being real, being cut off from the world, and having a veil between oneself and the outside world (Sierra et al., 2005). In cases of "alienation from surrounding" experience, it is also difficult for them to notice emergency signs and react to them. Additionally, even if colleagues were able to notice some troubles in the nurse's mind and gave her some useful advice, their words of advice did not reach her mind.

It is not clear why "anomalous subjective recall" is not related to medical error. However, many nurses are careful not to depend on memory only but also take written notes. These coping behaviours for anomalous subjective recall may help to prevent making errors.

In conclusion, the factors closely related to experience of filling out an incident/accident report were possible depersonalization disorder, working in a ward. Under depersonalized conditions, preventing errors may be very difficult not only for the nurses themselves but also when they receive the help of colleagues. Therefore, we need to detect depersonalization in nurses as early as possible, and to help prevent nurses from suffering from depersonalization disorder.

4.4 *Limitations*

A limitation of this study is that we used a self-rating scale in order to estimate depersonalization. Although we consider the CDS to be the most useful measure to investigate depersonalization at present and that using a self-rating scale is the best way to estimate mental problems of working nurses, the CDS cannot detect depersonalization disorder completely. Additionally, in the present study, the number of nurses whose CDS scores were higher than the cut-off points was too small to permit us to discuss their characteristics. In future studies it will be necessary to increase samples or to use diagnostic interviews. Moreover, this study was a cross-sectional study, so we could not reveal causality. In future studies we need to carry out longitudinal study.

ACKNOWLEDGEMENT

This work was supported by Grant-in-Aid for Scientific Research [KAKENHI] from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (20592514).

REFERENCES

American Psychiatric Association. 2000. *Diagnostic and statistical manual of mental disorders.* Washington, DC: APA.

Bebbington, P.E., Marsden, L. & Brewin, C.R. 1997. The need for psychiatric treatment in the general population: the Camberwell Needs for Care Survey. *Psychol Med* 27: 821–834.

Ebihara, H. 1979. Misanthropie und endogene depression. *Seishin Shinkeigaku Zasshi* 81: 275–287 (in Japanese).

Guralnik, O., Giesbrecht, T., Knutelska, M., Sirroff, B. & Simeon, D. 2007. Cognitive functioning in depersonalization disorder. *J Nerv Ment Dis* 195: 983–988.

Guralnik, O., Schmeidler, J. & Simeon, D. 2000. Feeling unreal: Cognitive processes in depersonalization. *Am J Psychiatry* 157: 103–109.

Hirosawa, M. 2005. Depersonalization. In Uezima, K. (ed.), *New approaches for psychiatric treatment Vol.3: Neurotic disorders and stress related disorders.* Tokyo: MEDICAL VIEW 137–42 (in Japanese).

Hirosawa, M., Sugiura, M. & Okada, A. 2009. Depersonalization disorder; a hidden psychopathology related to medical errors–retrospective investigation of 10 nurses' case reports. *Jpn J Gen Hosp Psychiatry* 21: 32–42.

Kohn, L.T., Corrigan, J.M. & Donaldson, M.S. 2000. *To err is human: building a safer health system.* Washington, DC: National Academy Press.

Lockley, S.W., Barger, L.K., Ayas, N.T., Rothschild, J.M., Czeisler, C.A. & Landrigan, C.P. 2007. Effects of health care provider work hours and sleep deprivation on safety and performance. *Jt Comm J Qual Patient Saf* 33: 7–18.

Maslach, C., Jackson, S.E. & Leiter, M.P. 1996. *MBI manual 3rd edition.* Palo Alto. CA: Consulting Psychologists Press.

Matumoto, M. 1998. Psychopathological characteristics of "double consciousness" appeared in a patient with depersonalization- from a 37 years clinical course a serious depersonalization. In Matumoto, M. (ed.) *Clinical Psychopathology of Schizophrenia* 1: 263–291. Kyoto: Jimbunshoin (in Japanese).

Nuller, Y.L. 1982. Depersonalization -symptoms, meaning, therapy. *Acta Psychiatr Scand* 66: 451–458.

Sierra, M. & Berrios, G.E. 2001. The phenomenological stability of depersonalization: comparing the old with the new. *J Nerv Ment Dis* 189: 629–636.

Sierra, M., Baker, D., Medford, N. & David, A.S. 2005. Unpacking the depersonalization syndrome: an exploratory factor analysis on the Cambridge depersonalization scale. *Psychol Med 35: 1523–1532.*

Sierra, M. & Berrios, G.E. 2000. The Cambridge depersonalization scale: A new instrument for the measurement of depersonalization. *Psychiatry Res* 93: 153–64.

Simeon, D. 2004. Depersonalisation disorder: A contemporary overview. *CNS drugs* 18: 343–354.

Sugiura, M., Hirosawa, M., Tanaka, S., Nishi, Y., Yamada, Y. & Mizuno, M. 2009. Reliability and validity of a Japanese version of the Cambridge depersonalization scale as a screening instrument for depersonalization disorder. *Psychiatry Clin Neurosci* 63: 314–321.

Sugiura, M., Hirowawa, M., Nishi, Y., Yamada, Y., Mizuno, M. & Tanaka, S. 2008. Development of a Japanese version of the cambridge depersonalization scale and application to Japanese university students. *Juntendo Institute of Mental Health Annual Reports* 93–101.

West C.P., Huschka, M.M. & Novotny, P.J. et al. 2006. Association of perceived medical errors with resident distress and empathy: A prospective longitudinal study. *JAMA* 296: 1071–1078.

World health Organization. 1992. *The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines.* Geneva: World Health Organization.

Yasunaga, H. 1987. Depersonalization. In Doi. T., Kasahara, M., Miyamoto, T. & Kimura, B. (eds.) *Abnormal Psychology Vol. 4 Neurosis and psychosis.* Tokyo: Misuzu Shobo 212–253 (in Japanese).

Zhan. C. & Miller M.R. 2003. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA* 290: 1868–1874.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Acceptance of tattoo culture from different lifestyle groups

Jui-Che Tu
Department of Creative Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

Shu-Ping Chiu
Graduate School of Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

ABSTRACT: According to historical records, tattoo, which originates from punishment, and leaves a negative impression on common people. However, as time passes and social formation changes, tattoo has gradually combined with popular culture and become the token of adolescents' expressing individuality, acting cool and controlling over their bodies. This kind of alternative sub-culture-and-art makes the body become a canvas for imprinting ego.

This study aimed to find out that whether different lifestyle groups' acceptance of tattoo culture varies. It used questionnaire survey as the research tool and took questionnaire statistics by SPSS 17.0. To achieve the study purposes, this paper adopted research methods including independent samples test, One-way ANOVA and frequency distribution. A total of 389 usable responses were received, representing a 96.5% response rate. With a confidence level of 95%, the sampling error was ±4.97%. Overall reliability α value was 0.724 and validity KMO value was 0.86. The main study results were summarized as follows:

1. The highest proportion of tattoo acceptance was 39.2% and mainly located on tip-end sites such as ankle.
2. 25.5% of testees liked abstract totem and this percentage was the highest. Then comes totem or written word with special meanings, which accounted for 20.3%.
3. 50.9% of testees believed that tattoo was a kind of artistic expression and 40.3% though that celebrity or stars could push forward tattoo prevalence.
4. Testees having tattoos on their bodies felt that tattoo was positive (with a mean value of 3.69) while those with no tattoos didn't think so (with a mean value of 2.91). There was significant difference between the two.
5. Testees with tattoos approved of their families or boyfriends/girlfriends' tattoo (with a mean value of 3.69) and those with no tattoos (with a mean value of 2.61) had contrary opinion. The two sides' opinions were significantly different.

Study suggestions: The results show that people with tattoos and without tattoos have different acceptance of tattoo culture to some extent. Tattoo advocates and enthusiasts should emphasize on how to convert people's negative impression of tattoo into positive attitude.

Keywords: tattoo culture, lifestyle, alternative sub-culture

1 INSTRUCTIONS

1.1 *Research motivation and purpose*

Tattoo culture has a long history. Tattoo endows people with social marginalizing imagery, is an indelible punitive mark given to criminals. Tattoo develops from the status symbol of tribal society into the symbolic sign of middle classes, lower classes, underworld gangs and specific industries in Taiwan, representing not only a token of self-physical control but also a symbol of love and popularity. Iconolatry and the media exert deep effect on tattooing. In a study on tattoo as subculture in teenagers (Tsai, 2006) conducts a qualitative research and focuses on three sections: the construction of self-identity; interpersonal relationship and groups identities; and the representation of tattoos and bodies based on the filed of tattoos. Her research results indicate tattoo has the following important meanings: the individualized style is represented; the emotion leads off; increases the friend; and the health acts on one's own. This study aims to analyze and generalize the historical development of tattoo through literature collection, and use quantitative research to analyze differences in acceptance of tattoo culture with different population variables and different life style groups. Then, it further discusses tattooed and untattooed people's views of tattoo culture, so as to know the evolving condition and trend analysis of tattoo culture.

2 LITERATURE DISCUSSION

2.1 *What is tattoo?*

Tattoo derives from the word of "Tatau" in Polynesian. It is written as tattooing in English, Tatowierung in German and tatouage in French. In Japan and Korea, tattoo is also called "emblazonry", "needle", "piercing", etc.

Webster's Dictionary (1993) defines "tattoo" as: a character, symbol, pattern or sign that is permanently marked on people's skin or body using a needle and pigment (Huang, 1987) classifies tattooing into four kinds: primitive cutting and piercing, branding, needle sewing, and needle piercing. Modern tattooing is developed from the fourth method.

Piercing has many variant forms, e.g. teeth destroying, lip punching, teeth engraving, scar branding, etc., all for making scars on the body just as tattooing.

2.2 *Cultural history of tattoo*

Getti Tattoo has an origin lost in the mists of antiquity. Among the mummies 3000 to 4000 years old, patterns of horses appeared on the textiles or skins as if they had been staying by their masters during lifetime (Schildkrout, 2004) Some data discover that, in Southeast Asia, doctors and Buddhist monks make tattoo on believers, so that super-natural power can protect them and others who approach them can keep kind-hearted (Tannenbaum, 1987). Tattoo is also used as a religious mark by early Christendom. Moreover, some medial astrologers tattoo cosmology symbols on their bodies, hoping that astrology can change their fates. Greeks, Roman and Celts use tattoo as a sign of criminals or properties (Schildkrout, 2004). In China, pre-Zhou-dynasty historical data recorded that primitive Wu nationality, living south of the Yangtze River, tattooed colorful emblazonry to scare piranhas. However, this kind of art was lost gradually. Until the thirteenth century, tattoo was just used as a mark of criminals.

Tattoo is common in ancient civilization and aborigines' culture, regardless of the West or East and Brazil, Africa or Japan. Tattoo is usually a decoration, ritual or a kind of identity. People's tattoo, a mark just like their name, represents people themselves and often adopts geometrical shapes, dots and lines (Bell, 1999).

Before 1880, tattoos were mostly made on criminals, sailors and the laboring class. In the late 1880s, tattoo suddenly became popular with the upper class of Britain and America, and this trend lasted one or two decades since then. The development of electronic tattooing machine was one of the reasons. The invention of this machine accelerated the speed of tattooing, alleviated the pains, made the colors and shadows more refined, and brought about better-quality paintings and coloring. Besides, O'Reilly and his partners began to design and sell tattoo design and stenciling, which made tattooing more and more standardized without a demand on masterful tattooing skills. At the same time, he introduced Japanese design from America. These three kinds of pioneering work made pains of tattooing reduced, and design diversified and more delicate, and thus, tattoo became welcome among the upper class in late 19th century (Fisher, 2002).

Tattoo is called "character piercing" in Taiwan and has a long history there. The earliest character tattooing records of Taiwan aborigines suggest that they have a long-history custom of making tattoos on skin. At the beginning, aborigines tattooed patterns representing "brave man and virtuous woman". In early times, affected by Japanese bushido culture, some people imitated samurais' tattoos so as to show their own "bravery". This group was mainly constituted of "underworld brothers" who formed subculture in early prisons and made the society think that tattooed people are not good (Du, 2003). In 1990s, because of open social mood and frequent contact with the West, Taiwan people lapped up European and American fashion. With propagation of the media, tattoo was combined with new ideas of "popularity", "vogue", etc, and tattoo shops sprang up like mushrooms. Nowadays, many underworld guys or local ruffian still tattoo traditional figures like dragon and phoenix as the group symbol. Tattoos are separated or excluded by different groups (Chuang, 2006).

2.3 *Definition of life style*

The basic concept and theory of "life style" was firstly proposed by psychologists Alder in 1927, and he believed that life style was constructed in terms of people's central goal (Hall, 1957). Afterwards, many scholars discussed this topic from different views and established abundant theoretical basis. The concept of "life style" originates from psychology and sociology. After 1960, (Lazer, 1963) introduced an idea of "life style" into marketing field. He followed above theoretical basis and integrated this with consumers' behavior at the same time. Scholars of marketing agree that lifestyle variables can better predict and understand customers' behavior and preference than demographic variables, and life style can reflect consumers' judgment and acceptance of everything. (Wind, 1974) classify the descriptions and measuring methods of individual life style into five measuring standards:

1. Measure the products consumed by individual and after-sale service.
2. Measure individual activity, interest and opinion: aiming at the initiative or passivity of activity, generation process and purposes of interest, as well as emotional cognition and opinion of attitude, i.e. A.I.O variables.
3. Measure consumers' values system.
4. Measure consumers' personality characteristics and self conceptions.
5. Measure individual attitudes towards different products and brands as well as the benefits they pursue.

2.3.1 *A.I.O framework*

(Plummer, 1974) proposed AIO (Activities, Interest, Opinion) variables scale and he thought that demographic variables should be added to the three items of AIO, constituting four perspectives.

Among the four perspectives, AIO is the most common measuring variable. (Engel, 1978) defined AIO as:

1. Activities: a concrete action, such as media watching, shopping, etc.
2. Interest: the degree of being excited at some thing, object or topic, and being continuous to pay special attention to it.
3. Opinion: individual oral or written response to an exciting situation, i.e. people's explanation, expectation and evaluation of things.

2.4 *Formation of life style groups*

Different people live under different environment. Individuals with similar living environments always have close personality and generate similar behavior response, and this is called formation of "group" (Tu, 2004). Under complicated environmental factors, the conditions used for separating groups include (Mahon, 2003):

1. Age: People growing up contemporarily can naturally form particular, similar personality tendency.
2. Gender: usually causes differences in physiology, psychology and social role. Though men and women have equal rights, inborn differences indeed exist.
3. Geography: Difference also results from different-geographical conditions of living environment, such as different nationalities, towns as well as eating and consuming habits of tropical and glacial zones.
4. Economic capability: The quality and quantity of consuming is affected by differences in economic situation, and thus generates different life quality.
5. Race: Racial cultural concepts and traditions result in differences, such as differences between
6. Educational status: Different educational backgrounds cause different cognitive ideas.
7. Occupational difference: Different occupations result in different habits and customs. For example, teachers and merchants have differences in working and resting schedule as well as consuming habits.
8. Interests and hobbies: People with different interests and hobbies purchase different products, and difference is formed.
9. Personal character: Every one has his own character, which causes differences.
10. Other universal factors: People can be classified into various sub groups according to diversified factors such as their different backgrounds chances, growing backgrounds, etc.

3 RESEARCH METHODS

This study carries out research through questionnaire survey. The questionnaire is designed in accordance with Likert five-point scale, and each question has five degrees including strong disagreement, disagreement, no opinion, agree and strong agreement. Answerers check the answers most appropriate for them. Questionnaires are sent to answerers in their face and are chosen through snowball sampling. Questionnaire statistics is analyzed through SPSS 17.0.

4 RESULTS AND DISCUSSIONS

This study makes generalization according to research purposes and literature discussion, investigates the quantitative parts through questionnaires, integrates the results for achieving research goals, and finally proposes original ideas as well as suggestions for successive researches in terms of the conclusions.

4.1 *Framework of questionnaire research*

Questionnaire of this study is divided into three parts:

1. Demographic variables: demographic variables are set in accordance with research purposes, so as to know the differences in different people's acceptance of tattoo culture. The variables include: gender, marriage, occupation, educational qualification, residence, professional background, age, income and bedtime.

Table 1. "Validity and reliability" analysis scale.

Reliability		Validity	
Overall perspective reliability (α)	Aspect reliability (α)	KMO value	0.865
0.724	Tattoo culture identity 0.866	Bartlett Sphericity test	4720.09
	Social image of tattoo 0.69	Significance	0.000

2. Life style: contains various life styles such as individual special interest and taste, habits and customs, as well as preferences for tattoo culture.
3. Acceptance of tattoo culture: includes questions like identifying with tattoo, cultural meaning of tattoo or social cognition.

4.2 *Questionnaire callback and statistics*

This study conducts questionnaire investigation through snowball sampling. Among the 403 questionnaires send out, 389 ones are usable with an effective response rate of 96.5%. Under the confidence level of 95%, confidence interval is ±4.97%.

4.3 *Validity and reliability test*

Reliability: Devellis (1991) believes that α value between 0.6 and 0.7 is the acceptable minimum value, and α value between 0.7 and 0.8 is fairly good. The Cronbach's Alpha value of questionnaire items is 0.724 in this study, representing a good consistency between the question items.

Validity: This study carries out factor analysis through Reduction Factor. Two factors are extracted, each corresponding to more than three questions. KMO value is 0.865. Total Variance Explained of the two factors is 56. 87%, which is greater than 50% and indicates the two factors are representative (Table 1).

5 RESULTS AND SUGGESTIONS

This study investigates different life style groups' different acceptances of tattoo culture by means of questionnaire survey. Some main conclusions are summarized as follows:

1. The highest proportion of tattoo acceptance was 39.2% and mainly located on tip-end sites such as ankle.
2. 25.5% of testees liked abstract totem and this percentage was the highest. Then comes totem or written word with special meanings, which accounted for 20.3%.
3. 50.9% of testees believed that tattoo was a kind of artistic expression and 40.3% though that celebrity or stars could push forward tattoo prevalence.
4. People have no tattoos consider that tattoo has underworld image (with a mean value of 2.61), and tattooed people think the underworld image of tattoo is low (with mean of 1.77). The mean of untattooed people is significantly higher than that of the tattooed.
5. Tattooed and untattooed people both agree that tattoo is an expression of art. Those with tattoos are highly in favor of a viewpoint that tattoo is a kind of art (with a mean of 4.31), and their identification degree is remarkably different from people without tattoos.
6. Testees having tattoos on their bodies felt that tattoo was positive (with a mean value of 3.69) while those with no tattoos didn't think so (with a mean value of 2.91). There was significant difference between the two.
7. About the prospect and market of tattoo industry, tattooed people think highly of it (with a mean value of 3.85) and people without tattoo remain hesitant (with a mean of 2.95).

REFERENCES

Bell, S. 1999. Tattooed: A Participant Observer' exploration of meaning. *Journal of Amernal of American Culture* 22: 53–58.

Chuang, N.-N. 2006. *Body, gender and tattooing: The gender construction and practices of female tattooing*. Shih Hsin University: Taipei (Master thesis).

Du, R.D. 2003. *Substitute for military service tattoos behavior and related factors*. National Chung Cheng University: Tainan (Master thesis).

Engel, J.F., Blackwell, R.D. & Kollat, D.T. 1978. *Consumer behavior*. Illinois: Hort, Rinehartand winston.

Fisher, J.A. 2002. Tattooing the body. *Marking Culture Body & Society* 8(4): 91–107.

Hall, C.S.L. 1957. *Theories of personality*. New York: Wiley Press.

Huang, H.-N. 1987. *Lost Tattoo Culture* Taipei City: Long-Ho.

Lazer, W. 1963. *Life style concept and market*. Chicage: Stephen Greyser, Taward Scentific Marketing.

Mahon, N.E., Yarcheski, T.J. & Yarcheski, A. 2003. The revised personal lifestyle questionnaire for early adolescent. *Western J Nurs Res* 25: 533–547.

Plummer, J.T. 1974. The concept and application of lifestyling segmentation. *Journal of Marketing* 38: 33–37.

Schildkrout, E. 2004. Inscribing the body. *Annual Review of Anthropology* 33: 319–344.

Tsai, H.-H. 2006. *Tattoo as subculture in Taiwanese teenagers.* Yuan-Ze University (Master thesis).

Tu, J.-C. 2004. *Style life design: Analysis of culture, life, consumption and products*. Taipei: Asia Pacific Press.

Wind, Y.G. & Green, P.E. 1974. Some ceonceptual measurement and analytical problem in life style researc. In Willian D. Wells (ed.,). *Life style and Psychographics*. Chicago: AMA.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

An investigation on design education in elementary school in Taiwan under the trend of postmodern art: A case study of the design activity for the students in the formal operational stage

Li-Chun Huang
Graduate School of Design, National Yunlin University of Science & Technology, Yunlin, Taiwan

Chao Hsi Ho
Department of Creative Life Design, National Yunlin University of Science & Technology, Yunlin, Taiwan

ABSTRACT: The 21st century is a commercialization era of art and design, in which artists and designers compete in the same area, and the educational theory is also influenced by postmodernism. However, in Taiwan, the current instruction of elementary school in visual art still stays in the traditional teaching and learning of art and crafts, and the instruction activities relevant to design are seldom. The writer is an instructor as well as a researcher of art education in an elementary school. Therefore, action research was employed in this investigation. Except carrying out in person an instruction activity of designing T-shirt pattern, she also studied the senior students of their representation of design and aesthetic capability in the elementary school, understood deeply the design concepts of the little designers and their gains from and feelings of the learning activity by means of case interview.

Keywords: postmodern art, design education in elementary school, activity of design, visual art

1 INTRODUCTION

21st century is a commercial era for art and design full influenced by the post modern art trend. Designers and artists are competition on the same stage. The trend is also taking effect in school education. Multiple values take place of single value in the past. Teaching is focusing on the students must understand environment context, cultural communication and interpretation, deconstruction and construction of meaning. Art education is more concerned about social effect and combination of art and daily life. New Grade 1–9 Curriculum was introduced to elementary and middle schools in Taiwan in 2001. Traditional visual art education is put into Arts and Humanities Area. The curriculum of design education of elementary and middle schools in Taiwan is not very often match activities of school routine or traditional festival and not fits in post modern art education which requires adaptation to local conditions, individual differences and compromises. The writer is a researcher for design education of elementary school as well as a teacher. She employed an action research to carry out one of teaching activities is called "T-shirt design" personally. After reading the analysis of the aesthetic feelings and expression abilities of students' pattern design and taking individual interviews, the writer was able to understand students' concepts, experiences and feelings gained from the learning activity. She wishes this research could enrich contents of design education in Taiwan and enhance the students adapt to the capacity of post-modern art trend (Figs. 1–2).

Figure 1. The pattern was designed by the student whose number is 2.

Figure 2. The photo of T-shirt.

2 RESEARCH METHOD, STAGE AND GLOSSARY

2.1 *Research method and stages*

This research uses literature research method to analyze trends of post modern art, documents and related information of Taiwan art education. It states a brief history of Taiwan art education and stages of globe trend of post modern art. A school-based teaching activity of design education is introduced to students with action research. Individual interviews are also taken for understanding design concepts of young designers. The aim of this research is to make flexible use of knowledge for design education and intend to build a design education theory.

2.2 *Glossary*

2.2.1 *Postmodern art*

Michael Woods has defined the post modern art that postmodernism rejects modernism's grand narratives of artistic direction, eradicating the boundaries between high and low forms of art, and disrupting genre's conventions with collision, collage, and fragmentation (Woods, 1989). Postmodern art holds that all stances are unstable and insincere, and therefore irony, parody, and humor are the only positions that cannot be overturned by critique or revision. "Pluralism and diversity" are other defining features (Sandler, 1996).

2.2.2 *Formal operational stage*

The formal operational period is the fourth and final of the periods of cognitive development in Piaget's theory with long term observation. This stage commences at around 11 years old (puberty) and continues into adulthood. In this stage, youngsters begin to think abstractly, solve problem in accordance with the hypothetical experiments, as well as apply formal logic rules to problem thinking (Borich & Tombari 1997).

2.2.3 *Activity of design*

The word 'design' means different things to different people. "Design" consists of "de" and "sign"."De" stands for destroy and "sign" stands for marking. Therefore the English word "design" is to destroy existing objecting and to create, to invent or to reconstruct new objecting (Esslinger, 1990). Activity of design is defined of taking practical action of using design concept and making positive or right direction to human life.

3 DOCUMENTS DISCUSSION AND RESEARCH

3.1 *Brief of post modern art trend*

"Modern art" in early 20th century, "art" was creators' the expressions of self centered and aestheticism and analytical "form structure" was part of modern art structures. People get the ability of aesthetic judgment from creation and enjoyment the form of art works. Post modern art trend has been developing since 1970s. The art works have been going to the trends of anti-form, anti-aesthetic, deconstruction and more frequently using multimedia, diverse and inclusive forms of artistic expressions when post modern art trend introduced. The performances of post modern art shows more social activities and arts are required integration in the people life. Sometimes viewers are able to take part in the art shows. The multimedia, forms and contents showed in post modern art are not only broken the rigid form and boundary of art but also totally denied the specific existence of art. It makes art turn into a kind of concept (Chang, 1996). The trend of post modern art is remaining in 21st century. It makes changes and complexities of art and design even more different than the past. Hence, the author states the consequences and influences of the 20th century art trend and wishes people have basic knowledge or understanding of the rising of the trend of post modern art after reading this article.

3.1.2 *Continuation of post modern art trend*

In 1919 Dadaism artist Duchamp (1887–1968) added beard on the duplicate of Mona Lisa in his works "*L.H.O.O.Q.*" (1919). When the work showed in an art exhibition the world was embracing a new meaning of art: destroy the delicate and original art' principals people pursued, change the conception of art and subvert the modernist to analyze art works in the way of theoretical system. At that time art could be an incident or experience and creation art had become a game. Behaviors of daily life were also an art and artists were no longer masters but creators embraced by ordinary people. Viewers could explain the meaning of works from their own views. Post modernism art was founded by the rising of Dadaism and Surrealism in the early 20th century. Post modernism art allows people to find beauty, enjoy beauty and concern the environment through art when Pop, Graffiti, Conceptual and Installation art appeared and multimedia and interactive art moved into art gallery and museum after 1960s. Post modern art liberates the imaginations and creativities and amazingly becomes popular in the world not only now but also in the future (Li, 2007).

3.2 *Development of modern art education in Taiwan*

In the early 20th century, globe major trends of art education are in following stages: children-centered art theory in 1940s, discipline-based art education theory in 1960s, community-based art education in 1980s and post modern art education theory in 1990s (Chen & Huang, 1995). The development of world art education theory has enriched the content of Taiwan art education. The trend of post modern art education influences the art and design education in Taiwan in 21st century too. This essay divides four stages of art education of elementary and briefly describes art education theories and courses in Taiwan. At the end of description is the existing circumstance of art education of Taiwan elementary under the trend of post modern art.

3.2.1 *Development of modern art education of elementary in Taiwan*

In 1895 Taiwan was a Japanese colony and its art education theory or formal art education course were all copied from Japan. Taiwan started establishing own art education curriculum after 1945. Therefore this essay is using successively revised art education curriculum standards and its reforms in Taiwan after 1945 as main references besides art education theories in the world. Stages are divided into the following: the implementation of nine-year national education (1968), the lifting of martial law (1987), establishment of education reform committee (1994) and the implementation of Grade 1–9 Curriculum (2001). These stages are Taiwan brief history for elementary art education.

(1) 1945–1968

Lowenfeld (1947) from US and Read (1958) from UK proposed child-centered art education theory to the world. They believed that people had creative powers and learning abilities by nature. Children contact outside environment via sensory organs and learning. They can get more learning opportunities if teachers notice individual differences and give proper guides by way of art. Children's paintings can reflect the stage of cognitive development of children and also be easy to distinguish the various stages works of children (Lowenfe & Brittain, 1982). This theory was introduced to Taiwan in 1960s and influenced Taiwanese visual art education for thirty years (Chen & Huang, 1999). The theory emphasizes the teaching role in art creativity too much and ignores the children aesthetic abilities, attitudes, developing art interests and acknowledge of traditional culture. Art education becomes a tool just for children natural growth of creativities and loses its unique functions (Kuo & Liu, 1991). In 1948 the education department of Taiwan enacted curriculum same as the one used in mainland China, transferred Drawing course into Art which included appreciation lessons, publish and researches, combined Art and Labor courses as Work course for young children. In 1962 the targets and hours of Art, Labor and Work curriculum remained unchanging but other curriculum were amended because of the time of domestic and international situation and education thoughts (Wang, 2008).

(2) 1968–1987

In 1968 the national compulsory education was extended from six years to nine years. The first six-year was elementary school education. The next three-year was middle school education. In 1975 Art and Labor courses merged into one course called Art and Labor, and reduced course hours from 120 to 80 minutes for low grade students and 150 to 120 minutes for middle and high grade students. After merger the courses, Art and Labor course added more hours for design and consisted of six parts of drawing, sculpture, design, craft, gardening and housework (Wang, 2008). In 1960s Eisner et al. stated that children- centered was the cause of teacher becoming useless, criticized the low teaching qualities and wanted return of Discipline-Based Art Education (DBAE) (Clark et al. 1987). The educational content of DBAE included four sections of aesthetics, arts criticism, art history and artistic creation. The learning materials are chosen from ancient to modern, west to east and practical art to pure art. Adult art works are the key part of teaching materials management and united four parts of the educational content. DBAE was rigorous curriculum center oriented, believed that teachers should plan the teaching program and stay in the program, and encouraged teachers to connect or merge art from other subjects (Chen & Huang, 1999). Taiwan art education was still based on children-centered although new theory of art education of world was established at that time.

(3) 1987–1994

In the late of 1980s DBAE was full of opinions of feminism, multi-culture, social Deconstruction etc., and fit into courses of folk custom and public art besides fine art (Delacruz et al. 1996). Through continuing development of DBAE Hamblen called the post modern trend of DBAE as Neo-DBAE (Hamblen, 1993). In 1993 and 1994 elementary and middle school curriculum of art education emphasized the combination of teaching and real life practices in appreciation, performance, and aesthetic judgment and so on. The whole course plan broke away the line drawing, color painting and material performance planed for the core of whole curriculum. These changes were influenced by Neo-DBAE (Wang, 2008).

(4)1994-now

McFee (1970) recommended community-based art education because of the DBAE disconnection concept of society and community. She considered art education was the process of discovery, cognition and appreciation for community value. Art education was able to push students to be participating citizens in democratic society. This theory of visual art education paid more attention to community which consisted of people, geographic environment, artificiality and living beings. Teaching in environment and life wastes was in order to improve community environment and allow students to recognize localized visual symbols. In 1994 Education Reform Committee of Taiwan started whole national wide education reformation. Since 2001 it has been implemented Grade 1–9 Curriculum that Learning Areas are main learning contents in Taiwan. All Learning Areas must be designed according to students' interests, needs from communities and special needs for school developments. The curriculums of all Learning Areas also are offered flexible elective courses ready for students to choose and applied principles of Curriculum Integration and Team Teaching (Chen, 1999). The implementation of Life Curriculum is merged art education content for low grade students, Arts and Humanities Learning Areas is merged art education content for grade three to nine students (Wang, 2008). Unknown future of 21st century forces students having more skills, therefore, the curriculum should be designed practically, adaptation to local conditions, flexible and diversification. The curriculum should be designed easy to students to cognize or understand the living environment and encourage themselves be able to change their living environment. It is very important to the art education. The development of Taiwan art education synchronizes with the world now.

3.2.2 *Taiwan art education under post modern art trend*

In 1996 Efland et al. indicated that art is the product from culture in the article of "Postmodern art education: An approach to curriculum". It says culture is no good or bad difference. When people are interested in origins and comprehensive abilities of culture can improve the depth of art understanding. Art should be connected daily life and tended to describe the inside of people's life artistically. The education value in this article is art of mutual respect and diversification and emphasizes to appreciate art from different nations in globe view (Liu, 2002). The thought of post modern art education is becoming popular in the world. Taiwan art education is no longer only focused in training tastes of delicate art or teaching in traditional drawing skills but researches and accepts multiculture. Post modern art education in Taiwan is focus on understanding of culture environment, engaging in communication and interpretation of different cultures, deconstruction and construction of art meaning, valuing sociality of art and combining art with daily life. Finally it is united visual, auditory s and performance art as one of Learning Areas and teaches cross-subjects (Lu, 1999).

4 RESEARCH OF TEACHING ACTIVITIES OF DESIGNING FOR STUDENTS IN FORMAL OPERATIONAL STAGES

Lin (2001) stated that art would spread to all living sections in future Taiwan society. The behavior of art is not only exists for art works but naturally infiltrates into the various aspects of human life. Post modern art education should be adjusted for the changes of multi-culture society. The practical teaching of design education should be a useful tool to cut schools and communities' distances. The writer conducted a teaching activity of T-shirt design for elementary school students and researched aesthetic abilities and numbers of different motive categories designed from high grade students. She also took some individual interviews to comprehend design concepts from students in formal operational stage, experiences and feels which students gained from this learning activity. Finally, the writer wishes this research could enrich the content of elementary design education in Taiwan.

4.1 *Teaching activities of design with school-based curriculum*

Teaching of design education in visual art is concerted with auditory and performance art today and also enriches contents of Arts and Humanities Learning Areas. But the curriculum of design education of elementary school does not match routine activities such as Sports competition, graduation ceremony or traditional festivals like the Dragon Boat Festival, Mother's day or Christmas. It is not school-based and even not matches diversification and practical features of post modern art trend. For example, the high grade students have the tradition of wearing uniform to take part in any school activities in the elementary school where the writer teaching. The uniforms were designed by design studios outside school formerly. The writer used action research to plan a school-based teaching activity of design. The research process and content were finally amended by observation of how students learned and analysis of the research results. The teaching and research contents are listed as following.

4.1.1 *Research content and target*

The research took place in an elementary school in Taichung city, Taiwan. 68 students were involved in this research. The writer designed the whole course, used natural observation to watch the teaching process in the classroom and also analyzed and researched in T-shirt pattern designed by students. The results were concluded from aspects such as design abilities and ideas of pupils, their experiences and feeling from the activity via interviews. The results was verified and analyzed by the effects and the outcomes gained from pupils design activity.

4.1.2 *Questions and participants*

A. *Research and Questions interested:*

1. How many different motive categories of T-shirt pattern were designed? Why students did choose these motives? Did they copy or misappropriate the sample motives?
2. Which color did students like most? Is the color in cool tone or warm tone? Do they use one color or colors assorted by them? Did the colors match the patterns?
3. Did students design patterns independently? Did the pattern express their fully opinions?
4. Did the texts match or enhance the design effects? Did the texts really state what the students wanted to say? Did the whole picture of students' designs particularly reflecting them in formal operational stage?
5. What levels did students keep focus on or involved in the design activity and art produce process?

B. *Research participants*:

Examples in this research were from an elementary school in Taichung City of Taiwan, 68 students from two classes (females 34, males 34). They were in age of 12 years old (41 students) and 13 years old (27 students). They were all in formal operational stage in Piaget's theory. Teacher guide and teaching model should be suit to children's psychological requirements. The teaching activity designs were differed according to those requirements.

4.2 *Outcomes and difficulties*

Using natural observation watched and interviewed students during classes. The writer made assessments of abilities of aesthetic and design at the end of teaching activities through finished copies of students and made further interviews if required. This research essay recorded not only the outcomes but also the difficulties or unexpected situations met in the research process as below:

4.2.1 *Research outcomes*

1. Motive Categories: Most of motives were from cartoons or comics in their daily life such as Songe-Bob SquarePants, Snoopy, and Miffy etc. Students were able to make some necessary changes from imitating original of cartoons or comics.

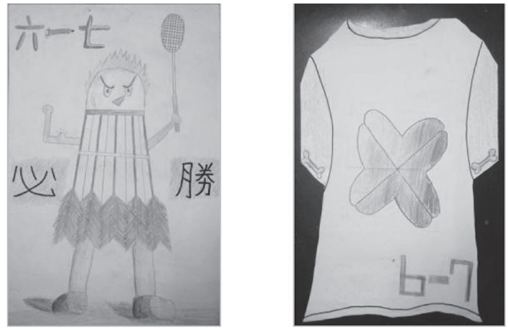


Figure 3. The most of students used pure color or warm colors because children liked vivid color.

2. Colors: Most of students used pure color or warm colors because children liked vivid color. But their color schemes were not applied well enough so that teacher will take more time to teach them how to match colors for a remedial teaching. (Figure 3).
3. Composition Showing: Most students were paid lots attention to their designs and completed by themselves. They also were mastered of aesthetic principles of simple, elegant and theme clear. Their patterns clearly delivered information what they wanted to express. However, they still need help from teachers when they added more or complex patterns in one picture. Their behaviors matched the formal operational stage during learning. They were able to think with abstract thinking, follow formal logic disciplines and solve problems in order.
4. Text Delivery: Texts can match patterns and clearly deliver their opinions and feels. This research also found that: teaching plan must be based on students former learning experiences so that the results can be reached much positive. During the observation of teaching activity of design, the course with daily life experience or the practice would raise students' learning interests and stimulate their creative passion. Using easy or simple materials and techniques can decrease the chance of learning frustration. When the whole work of design was presented confidences of students would get increased as well.

4.2.2 *Difficulties*

Because the writer was not students' home room teacher, she had to spend more time to maintain disciplines in the class. When students finished their design unexpected situation occurred afterwards. For example, finished products can not be compared with original designs because one home room teacher made changing images and the others lost the original copy. Furthermore, the writer hadn't enough time to do a remedial teaching because of other scheduled courses. Therefore this research gets to use other examples for amendment research which will increase the varieties of research results. This is the drawbacks of non-laboratory studies, but also reveals the true state of action research.

5 CONCLUSION

At the beginning of 20th century, Dadaism revealed the performance of traditional and modern art and the rising of Pop art broke the boundary of fine art and commercial design. In the 21st century artists and designers compete at the same stage and post modern art trend influences education theories. There are lots of achievements of art education since the recovery of Taiwan from Japanese but still lots questions may take our attentions and make progresses. Especially under the post modern art trend, art education should response to the era changing which is very necessary. Post modern design education should be a useful tool to cut distances between schools and communities. The writer wishes results of this research can be shared or communicate to other teachers in the same school but also enrich the content of design education of Taiwan primary and increase students' abilities of complying with trend of d post modern art.

REFERENCES

Best, S. & Kellner, D. 1991. *Postmodern theory: Critical interrogations*. London: Macmillan.

Borich, G.D. & Tombari, M.L. 1997. *Educational Psychology- A Contemporary Approach*. New York: Longman.

Chalmers, F.G. 1996. Celebrating pluralism: Art, education, and cultural diversity. Angeles, CA: *The Getty Center for Education in the Arts.* Columbus: Ohio State University Press.

Chang, Chiwen, 1993. Discussion Primary Art Education under Information Era Trends. *Art Education Theory and Practice Forum*: 29–40. Taipei:National School Teacher Discussion of Taiwan Province.

Chen, Chaoping. & Huang, Jenlai. 1995. *Elementary School Art and Craft Teaching Materials and Skills*. Taipei: Wunan Books

Chen, Pochang. 1999. The Theoretical Analysis of the Concept of Grade 1–9 Curriculum. *Grade 1–9 Curriculum Seminar Collection: Towards New Era*:10–18. Taipei: Republic of China Text book Research and Development Institute.

Cheng, Huiwen. 2008. The Depthlessness of Postmodern Art: Downward or Outward? *Mingtao Liberal Forum 4: 103*.

Chou, Lingwen. 2002. *A Study on the Development and Trends of Thoughts in the Visual Arts Education in the Elementary School since the Restoration of Taiwan*. Pingtung: Essay in Master degree from National Pingtung Institute of Visual Art Education.

Efland, A., Freedman, K. & Stuhr, P. 1996. *Postmodern Art Education: An approach to curriculum.* Reston, VA: The National Art Education Association.

Esslinger, H. 1990. Design - The Common Language all over the World. Liu, Huimin edited. *Product Design and Packaging Vol. 40: 43–46*.

Hamblen, K. 1997. The emergence of Neo-DBAE. In J. Hutchens & M. Suggs (Eds), *Art education: Content and practice in a postmodern era: 40–46*. Reston,VA: National Art Education Association.

Hassan, I. 1987. *The Postmodern Turn. Essays in Postmodern Theory and Culture*. Columbus: Ohio State University Press.

Ho, Ching. 1994. *Modern and Postmodern - A Short History of Western Art and Culture*. Zhejiang: China Academy of Fine Arts Publishing House.

Kao, Ichan. 1997. Revelation for Postmodern Philosophy Education to School Education. *Journal of Primary Education of National Tainan Normal Institute 10*: 273–97.

Lin, Chenghuang & Library Bulletin editor team. 2004. *National Chengchi University Library Bulletin* 13. http://www.lib.nccu.edu.tw/Pub/013/index.htm

Lin, Manli. 1999. Theory Development of Visual Art Education in Taiwan. *Information from Visual Arts Education Curriculum Reform Seminar*. Taipei: National Art Education Center.

Liu, Fengjung. 1991. *Eisner Art Education Thought Research*. Taipei: Shui Niu Press.

Lowenfeld, V. & Brittain, W.L. 1982. *Creativeand Mental Growth 7*: 3–129. NYC: Macmillan.

Lu, Jungchih. 2003. *"Breaking" Post Modern Art*. Taipei: Artist Press.

Lu, Yenching. 1999. *Nine-Year National Education Research of Syllabus of Arts and Humanities Learning Areas*. Taipei: Education Department.

Park, M.E. (Hsu, Jungjung translated) 1989. Postmodernism in Art Education. *Modern Art 34*: 48–51.

Rose, M.A. 1991. *The post- modern and the post- industrial: A critical analysis*. London: Cambridge University Press.

Sandler, I. 1996. *Art of the Postmodern Era: From the Late 1960s to the Early 1990s*. New York: Icon Editions.

Wang, Chungpin. 2000. Post Modern Art and the Aesthetic Meaning of Expressions. *National Education Report 12*: 18.

Wang, Liyen. & Cheng, Minghsien (Chief Editor). Visual Art Education -Brief Introducation of Visual Art School Education in Taiwan. *Taiwan arts education history*. Taipei: Taipei City Arts Museum.

Wolctt, A. 1994. Whose Shoes Are They Anyway? A Contemporary Approach to Interpretation. *Art Education 47(5)*:14–20.

Woods, M. 1989. Art of the Western World, *Summit Books:* 323.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Eating habits and the health conditions on middle-aged working women

Yoshika Suzaki
Japanese Red Cross Kyushu International College of Nursing, Munakata, Japan

Hiromi Ariyoshi
Institute of Nursing, Saga Medical school Faculty of Medicine, Japan

Minoru Moriguchi
Hiroshima International University, Higashi Hiroshima City, Japan

Shuhei Ryu
Tokyo Women's Medical University, Tokyo, Japan

ABSTRACT: The survey period was from May to December in the year of 2007, and the survey was based on relative factor searches. We sent the questionnaires to 2,000 middle-aged females in employment with ages ranging from 30s to 60s, who were residents of City A or worked for the municipal government or its affiliated organizations of City A in Prefecture A. According to the preference of the respondents, we divided them into four groups: the home-made food group, the pre-made food group, the eating out group, and mixed group, and examined the correlations of health habits, BMIs, subjective health conditions, ages, and lunch break time.

Keywords: eating habits, middle-aged working women, health conditions

1 INTRODUCTION

1.1 *Introduction*

Eating habits and health are closely related to our lifestyles or life activities. Middle-aged women play a major role to maintain their family's health by providing meals. On the other hand, we are facing a phenomenon of reduced cooking time partly because the food service industry has been developed and a lot of people "eat out" and partly because the so-called "pre-made food" eating style has appeared, in which people will do take-outs of cooked food as it is and eat it at home. The increase of people who "eat out" or "eat pre-made food" in number has brought up health issues due to over consumption of cooked food. Therefore, this study is intended to identify the usage of eating pre-made food and eating out among the middle-aged working women as well as to define the relationships between the patterns of such eating habits and the health conditions/activities.

1.2 *Operational definitions of the terms*

Pre-made food: A style of doing take-outs of cooked food as it is and eating it for a meal at home.

Home-made food: A style of cooking food at home and eating it for a meal.

Eating out: A style of having a meal outside home.

2 STUDY METHOD

2.1 *Survey method*

The survey period was from May to December in the year of 2007, and the survey was based on relative factor searches. We sent the questionnaires to 2,000 middle-aged females in employment with ages ranging from 30s to 60s, who were residents of City A or worked for the municipal government or its affiliated organizations of City A in Prefecture A. Their consent to accept the study was gained on collecting the questionnaires.

2.2 *Analysis method*

According to the preference of the respondents, we divided them into four groups: the home-made food group, the pre-made food group, the eating out group, and mixed group, and examined the correlations of health habits, BMIs, subjective health conditions, ages, and lunch break time.

3 RESULTS

3.1 *Collection of questionnaires*

Questionnaires with any missing entries and illegible questionnaires due to roughly written answers were deemed to be invalid answers. The final valid answers were obtained from 841 respondents. The valid response rate was 79.4%.

3.2 *Proportion of four groups: Home-made food, pre-made food, eating out, and mixed*

The result was as follows: the home-made food group = 11.2% (94 respondents), the pre-made food group = 20.3% (171 respondents), the eating out group = 5.1% (43 respondents), and the mixed group = 63.4% (533 respondents). Thus, the mixed group was the majority among the respondents (Table 1).

3.3 *Comparison of health habits, BMIs and subjective health conditions among the four groups*

Significant differences were observed about subjective health conditions while there were no significant differences of health habits or BMIs (Tables 2, 3, 4).

3.4 *Comparison of working hours, lunch break time and commuting time among the four groups*

Significant differences were observed about lunch break time while there were no significant differences of working hours or commuting time. It was also found that more than 70% respondents in the eating out group have lunch break time more than 60 minutes (Table 5).

Table 1. Proportion of four groups: Home-made food, pre-made food, eating out, and mixed.

	n	%
Home-cooking group	94	11.2
Pre-made food group	171	20.3
Eat-out group	43	5.1
Mixed group	533	63.4
Total	841	100

Table 2. Comparison of health habits among the four groups.

		Health habits medium or better	Poor	P
Home-cooking group	n	62	32	
	%	66.0	3.4.0	
Pre-made food group	n	60	111	
	%	35.1	64.9	
Eat-out group	n	14	29	
	%	32.6	67.4	
Mixed group	n	261	272	
	%	49.0	51.0	0.0001***

+ p <0.10, *p <0.05, **p <0.01, ***p <0.0001.

Table 3. Comparison of BMIs among the four groups.

		BMI <24	>25	
Home-cooking group	n	89	5	
	%	84.7	5.3	
Pre-made food group	n	138	33	
	%	80.7	19.3	
Eat-out group	n	36	7	
	%	83.7	16.3	
Mixed group	n	423	110	
	%	79.4	20.6	0.0019*

+ p <0.10, * p <0.05, ** p <0.01, *** p <0.0001.

Table 4. Comparison of health conditions among the four groups.

		Health conditions relatively		
		Fine	Poor	
Home-cooking group	n	73	21	
	%	77.7	22.3	
Pre-made food group	n	121	50	
	%	70.8	29.2	
Eat-out group	n	28	15	
	%	34.9	65.1	
Mixed group	n	423	110	
	%	79.4	20.6	0.0033*

+ p <0.10, * p <0.05, ** p <0.01, *** p <0.0001.

4 DISCUSSION

Respondents to be included in the pre-made food group were mostly found among the 40s and 50s. Unlike a bachelor or family of only a husband and

Table 5. Comparison of working hours, lunch break time and commuting time among the four groups.

		Lunch break time			
		30 to <45 mins	45 to <60 mins	>60 mins	P
Home-cooking group	n	10	24	60	
	%	10.6	25.5	63.8	
Pre-made food group	n	20	39	112	
	%	11.7	22.8	65.5	
Eat-out group	n	8	4	31	
	%	18.6	9.3	72.1	
Mixed group	n	93	84	356	
	%	17.4	15.8	66.8	0.0037*

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$.

wife, many people in their 40s and 50s would have a family with children, which inevitably results in the increased number of family members. If the number of family members is increased, eating out will cause more financial burden. On the other hand, pre-made food allows all of the family members to eat together and enjoy a happy family atmosphere without causing so much financial burden. In addition, housewives can have more space of time in their daily lives by making use of pre-made food and allocate it to their own freshener, rest, or time to spend with their family. It can be assumed that the use of pre-made food is essential for people at this age who have established socially responsible positions.

Respondents to be included in the eating out group were mostly found among the 30 s in employment. There can be diverse purposes and reasons, such as to spend time with friends or colleagues or enjoy a happy family atmosphere, for the fact that the eating out style consists of one factor, however, it can be assumed that they do not really care about what their food is made of or how it is cooked. As for the nutrition labels, a higher percentage of who "do not care at all" about country of origin or additives were found among the 30s. It can be assumed that this way of thinking also relates and contributes to the fact that many of the 30 s are included in the eating out group. Respondents to be included in the mixed group were a majority of people from 30s to 60s.

Significant differences were observed about subjective health conditions and lunch break time. Although there were no significant differences of health habits or BMIs, some relations might be revealed if we continue the survey. In addition, more than 70% respondents in the eating out group have lunch break time more than 60 minutes. This was probably because many of the eating out group are sales representatives and tended to take lengthy lunch time break.

5 CONCLUSION

People in employment may have no cooking time due to a compelling reason, such as working overtime or bringing their work home; therefore, it can be assumed that they are leading their daily lives balancing their careers and home, while successfully combining home-made food, pre-made food, and eating out. A follow-up survey will probably clarify the relationship with health activities.

REFERENCES

Date, C. 2008. Factors contributing to world's longest active life expectancy why can the Japanese maintain the world's longest active life expectancy? 3 trends in Japanese dietary habits and the present situation. *Journal of the Japan dietetic association* 51(7): 687–692.

Miyoshi, M. & Hayashi, F. 2008. Regional characteristics of secular changes in obesity-related lifestyle behavior in Japan. *Anti Aging Med* 5(2): 30–38.

Mitsui, T. & Shimaoka, K. 2008. Gentle exercise of 40 minutes with dietary counseling is effective in treating metabolic syndrome. *Tohoku J Exp Med* 215(4): 355–361.

Nakamura, Y. 2010. Comparison of the national nutritional survey in japan estimated individual-based nutritional data and nippon data 80 food frequency questionnaires. *J Epidemiol* 20(3): 582–586.

Nakade, M. & Hirota, K. 2009. Improving diet and exercise behavior with a computer-assisted health education program. *Japan society of health education and promotion* 17(2): 97–108.

Part VIII: Performance modeling & simulation

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Fitts' law modeling of remote target acquisition hand movements in a 3D environment

Sue Kai Wu
Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

Fay Gao & Pei Luen Rau
Department of Industrial Engineering, Tsing Hua University, Beijing, China

Yung Hui Lee
Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

ABSTRACT: This study investigates and models the performance of target acquisition hand movements in a free-hand, touchless 3D environment. The targets in this study had different positions, sizes, and distances. Both performance time and moving trajectories were recorded as performance measurements. Total performance time was divided into "primary submovement time" and "secondary submovement time." These differences in performance time can be explained by the psychomotor ability of the postures used, visual cues, and control/display ratio settings. Model primary submovement time was satisfactorily explained by Fitts' model (without depth $R^2 = 0.96$; with depth $R^2 = 0.85$), but this was not the case for secondary submovement time (without depth $R^2 = 0.46$; with depth $R^2 = 0.38$). Different models were then developed for the primary submovement time, secondary submovement time and the total movement time. The R^2 values were improved from 0.38 to 0.977.

Keywords: target acquisition hand movement, three-dimension, depth perception, Fitt's law

1 INTRODUCTION

As an alternative to traditional input devices (mouse, trackball, joysticks, and etc.) "3D target acquisition hand movement" was explored as an input function to various appliances that require hand-free and no–touch interaction with computer. It is a movement of the hand/arm towards a specific direction, position, and/or an object. This activity enable user performs pointing, selecting, and manipulating objects. In comparing with wearable input devices, hand movement seem to be a more natural alternative, can be considered as a part of the environment, and be operates remotely (Nakamura, et al., 2008).

Factors of "target width," "target height," and "target distance" were used to study and model user performance in human computer interaction tasks. Task performance in 2D human-computer interface has been successfully characterized and be modeled by Fiit's law (Fitt's, 1954, MacKenzie & Buxton, 1992; Accot & Zhai, 2003). Among all factors, target width and the formulation used in the calculation of a task's index of difficulty play a critical role in the accuracy of the model (Accot & Zhai, 2003). Developing and evaluating appropriate theoretical performance models using modified versions of Fitt's law can also be found in 3D environments (Murata & Iwase, 2001; Grossman & Balakrishnan, 2004; and Liao & Walter, 2004). Most 3D studies examined target acquisition with a device in hand, our reviewed of literatures failed to find performance and modeling of remote target acquisition with free hand movements.

Total performance time was subdivided into "primary submovement time" and "secondary submovement time" followed the definitions of Walker et al. (1993), so as to realize among of time of fine adjustment for target acquisition, especially for the depth dimension. In this paper, we study and model performance of 3D target acquisition hand movements in a hand-free, no-touch interaction, 3D environment. We proposed models, followed Fitt's law, that appropriately characterize these 3D target acquisition behaviors.

2 METHOD

2.1 *Subjects*

Ten volunteers (six male and four female), aged 22–26 years (23.4 ± 1.5 years old), self-declared right-handed, participated in this study. They were paid voluntary for the study.

2.2 *The 3D control*

A 3D computer vision based systems was used to track free hand, on-touch, control movement in real time. Figure 1 illustrated the hand/head position tracking and the target acquisition hand movement.

The system is capable of recognizing gestures at a speed of 20 Hz. It was then the velocity and the accelerations of the hand movements were calculated. More details and results of system validations can be found in Lee & Ko, (2009).

2.3 *The 3D perspective display*

Targets, cursor, and the visual database were computer generated and were presented on a 42-inch LCD screen (BENQ S-Series). The resolution of the display is 1920 x1080 pixels. The computer used is a Inter (R) Core (TM), a 2 Duo CPU E8300 with 2.83 GHz, 2.00 GB RAM, and running with Microsoft Windows XP. Since targets were presented in a 2D perspective display, shadow of the target on the mesh with linear perspective and a forward-looking schematic was used to provide enhance visual information of target position in 3D (followed Jan & Arjen, 2002).

2.4 *Tasks*

There are a total of 17 targets on this reference framework (17 positions). Eight out of the 17 targets are located in the middle frame (with z = 0, without depth). They are in the X-axis, Y-axis, and XY plan, that is a 2D plan. Acquiring "targets without depth" are common practice while using traditional input device (e.g., mouse, joystick …). Other 9 out of 17 targets were located, with depths, in the z-axis. Eight out of the 17 targets were located in the middle frame (with z = 0, without depth) [eight targets: 0 ("Right"), 45 ("Upper Right"), 90 ("Upper"), 135 ("Upper Left"), 180 ("Left"), 225 ("Lower Left"), 270 ("Lower") and 315 ("Lower Right") degree]. Acquiring "targets without depth" is a common practice when using traditional input devices (e.g. mouse, joystick …). The other 9 of the 17 targets were located with depths [nine targets: 0 ('Forward Right"), 45 ("Forward Upper Right"), 90 ("Forward Upper"), 135 ("Forward Upper Left"), 180 ("Forward Left"), 225 ("Forward Lower Left"), 270 ("Forward Lower") and 315 ("Forward Lower Right") degree and ("Forward"). Please see Figure 2 for illustration.

In addition to different positions, spherical targets are one of three sizes (1.8, 2.8 and 3.8 cm in radius) and are located on the framework in one of the three distances (8, 12, and 16 cm).

2.5 *Performance measurements*

Both moving trajectories and performance time were recorded as performance measurements. Moving trajectory was recorded as a time displacement records of the cursor (free hand) movements. Total performance time included the duration between the point at which movement velocity exceeding 7% of the peak velocity (the disappearance of the home target) and the point at which the acquired target disappeared (end of target acquisition). Total movement time consisted of a "primary submovement time" and a "secondary submovement time," following Jagacinski, et al. (1980), Walker, et al. (1993) Liao, et al. (1997) and Thompson, et al. (2007). Hand movement velocity determined the end of the primary submovement time, following the study of Thompson, et al. (2007).

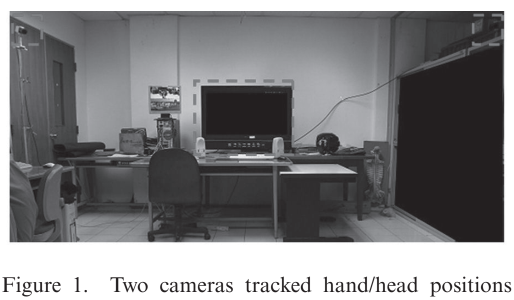


Figure 1. Two cameras tracked hand/head positions and target acquisition hand movements. Subjects sat 3 meters in front of a display screen.

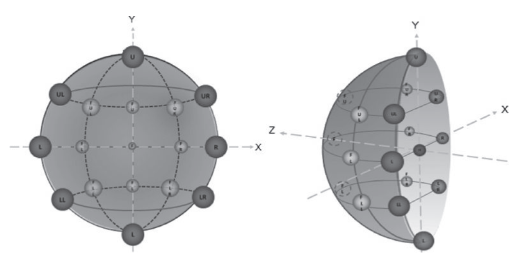


Figure 2. Targets and reference framework in the 3D spaces.

Performance data were interpreted in terms of Fitts' law. The movement time (MT) was expressed as:

$$MT = a + b\ ID,$$

where: MT is Movement Iime (total movement time, primary submovement time, and secondary submovement time), a and b are the empirically determined constants; ID is the index of the difficulty of the task.

2.6 *Experimental procedure*

To avoid mixing the change of target sizes with the change of target sizes as they recede in distance from an observer in 3D spaces, a split plot design was adopted, where the target size was selected as a whole plot. Target acquisition tasks were repeated in three different days for each subject. The experiment involved 3 days of experiment for each subject, each day consisted of three sessions and a total of 153 (3 sizes × 3 distances × 17 positions) targets were acquired.

2.7 *Data analysis*

A split-plot design was used for the data analysis, with target size as a whole plot effect. Factors of target distance and position were randomized within the plot. The means and standard deviations for all test measurements were calculated using standard statistical methods. Analysis of variance (ANOVA) was used to determine the effect of inter-subject variability, target size, target distance, target positions (excluding the "Forward" position, the remaining 16 positions represented 8 locations, with and without depth) and possible interactions. Tukey Test was used for post hoc comparison. An alpha level of 0.05 was selected as the minimum level of significance.

3 RESULT

3.1 *Primary submovement time*

ANOVA results show that "target size" had a significant effect on the primary submovement time ($F(2, 18) = 184.44$, $p < 0.01$). Post hoc analysis indicates that the time for a 3.8 cm target (1.55 ± 0.78 s) was significantly shorter than that for 2.8 cm (1.69 ± 0.84 s), followed by 1.8 cm (2.28 ± 1.31 s) (Tukey HSD critical value = 3.32, $p = 0.05$). "Distance" significantly affected mean primary submovement time ($F(2, 18) = 221.77$, $p<0.01$). The time for a distance of 8 cm was 1.43 ± 0.82 s, which was significantly shorter than that for 12 cm (1.81 ± 0.91 s), followed by 16 cm (2.28 ± 1.22 s) (Tukey HSD = 3.32, $p = 0.05$). Figure 3 illustrate the primary submovement time on size and distance (no depth and depth).

"Position" also significantly affected the mean primary submovement time ($F(7, 63) = 8.19$, $p<0.01$). The middle frame contained 8 targets (without depth, $z = 0$), while 8 other targets had depths. ANOVA results indicate that the "depth" of the targets affected the primary submovement time ($F(1, 9) = 23.95$, $p<0.01$). The primary submovement time for targets without depth was 1.76 ± 1.08 s, which was significantly shorter than that for targets with depth (1.92 ± 1.03 s) (Tukey HSD = 2.77, $p = 0.05$).

In addition, ANOVA results reveal a significant two-way interaction effect between "depth" and "distance" ($F(2, 18) = 7.56$, $p<0.01$). This interaction suggests that targets with and without depth have significantly different primary submovement times when they were located at different "reach distances,".

There was also a significant two-way interaction effect between "depth" and "position" ($F(7, 63) = 13.09$, $p<0.01$). This interaction suggests that the "primary submovement time" for targets without depth and in the lower part of the spherical frame was shorter than that for targets with depth. However, the time for targets with depth and in the upper part of the spherical frame was shorter than that for targets without depth.

3.2 *Secondary submovement time*

"Target size" significantly affected *secondary submovement time* ($F(2, 18) = 211.56$, $p<0.01$). Post

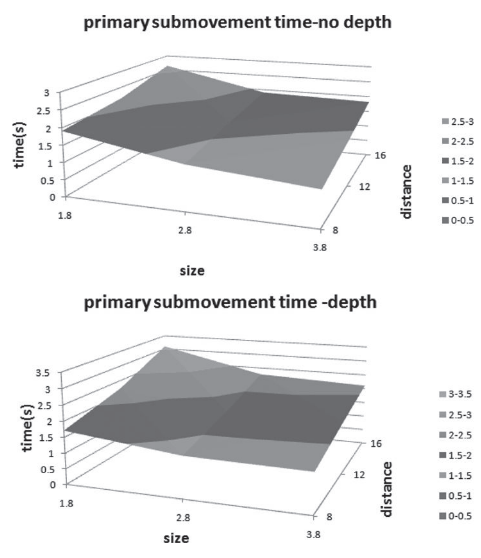


Figure 3. The primary submovement time on size and distance (up: no depth, down: depth).

hoc analysis indicates that the mean time for a 3.8 cm target (0.71 ± 1.42 s) was significantly shorter than that for a target of 2.8 cm (1.00 ± 1.71 s), followed by that for 1.8 cm (2.46 ± 2.72 s) (Tukey HSD critical value = 3.32, p = 0.05). Figure 4 illustrate the secondary submovement time on size and distance (no depth and depth).

3.3 *Total movement time*

"Target size" significantly affected total movement time ($F(2, 18) = 376.17$, $p<0.01$). Post hoc analysis indicates that the mean time for a 3.8 cm target (2.25 ± 1.60 s) was significantly shorter than that for a 2.8 cm target (2.69 ± 1.88 s), followed by that of 1.8 cm (4.74 ± 2.95 s) (Tukey HSD critical value = 3.32, $p = 0.05$).

"Distance" had a significant main effect on total movement time ($F(2, 18) = 30.83$, $p<0.01$). The mean time for a distance of 8 cm (2.90 ± 2.41 s) was significantly shorter than that for a distance of 12 cm (3.13 ± 2.46 s), followed by that of 16 cm (3.65 ± 2.49 s) (Tukey HSD critical value = 3.32, $p = 0.05$). However, the target acquisition times at distances of 12 cm and 8 cm were not significantly different. Figure 5 illustrate the total movement time on size and distance (no depth and depth).

ANOVA results show a significant two-way interaction effect between "depth" and "position" $F(7, 63) = 10.67$, $p<0.01$. This interaction suggests that the "total movement time" for targets without depth in the lower part of the spherical frame was

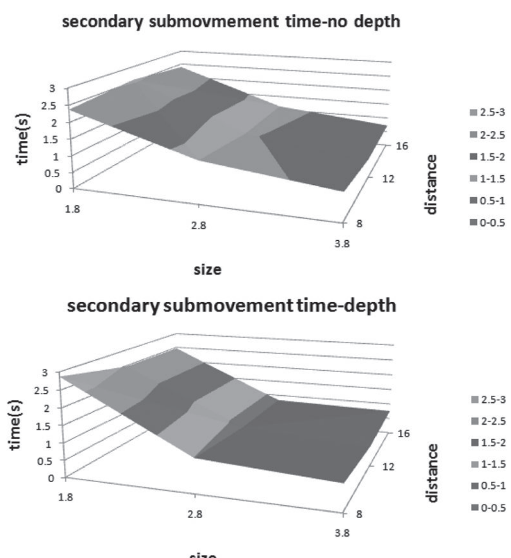


Figure 4. The secondary submovement time on size and distance (up: no depth, down: depth).

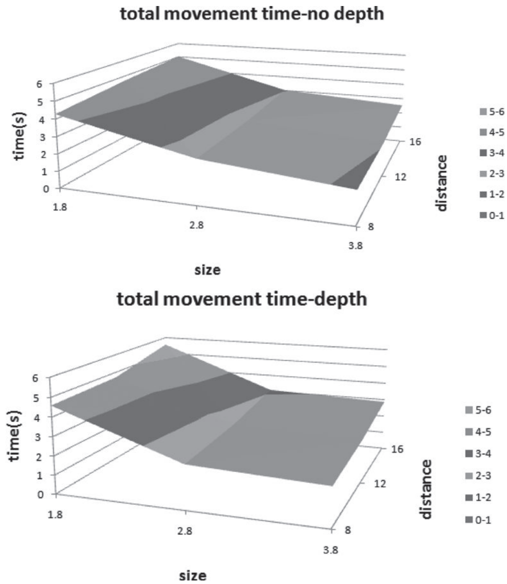


Figure 5. The total movement time on size and distance (up: no depth, down: depth).

shorter than that for targets with depth. However, the time for targets with depth in the upper part of spherical frame was shorter than that for targets without depth.

3.4 *Fitting conventional Fitts' law*

We deduce Fitt's regression models based on moving Distances (D) and target Sizes (S). The calculation of the task difficulty index ($ID=\log_2 2D/S$) showed that they were ranged from 2.07 to 4.15. The regression models were listed followed:

Primary Submovement Time (PST):
without depth $PST = -0.48 + 0.754\ ID$, $R^2 = 0.96$
with depth $PST = -0.459 + 0.814\ ID$, $R^2 = 0.85$

Secondary Submovement Time (SST):
without depth $SST = -1.942 + 1.299\ ID$, $R^2 = 0.44$
with depth $SST = -2.391 + 1.377\ ID$, $R^2 = 0.38$

Total movement time (TMT):
without depth $TMT = -2.422 + 2.052\ ID$, $R^2 = 0.72$
with depth $TMT = -2.85 + 2.191\ ID$, $R^2 = 0.72$

Based on the result of R2, it seems that on "primary submovement time" was explained well by the calculation of task difficulty index since it was significantly affected by factors of "target size" and "target distance." However, the calculation of ID did not explain well of the secondary submovement time as much as that of the primary submovement time. Based on different hand movements for primary submovement time and for fine secondary submovement time. a further effort was done to develop models for primary submovement time, secondary submovement time, and the total time, independently. The results are presented in Table 1.

4 DISCUSSION

The amplitude of Fitts's index of difficulty was originally directly related to the amplitude of the arm/hand movement (target distance) and visual input (target size). Our modeling results of the combined data successfully describe movement duration by using only these two variables. However, models for secondary submovement time, and total time were developed independently and the R-squared values are improved from 0.38 to 0.72. Basically, target acquisition movement consisted of two phases: an initial impulse toward the target (primary submovement time) and a deceleration phase under current control to home in on the target (secondary submovement time). The adjustment of cursor position requires hand-eye control and a back-forth closed-loop cycle to arrive a fine adjustment of cursor position. As a result, modeling of primary submovement time, secondary submovement time, and the total movement time should be different. The best model for the 3D, remote, free hand, target acquisition models are:

The ANOVA result could found that "target size", "target distance" and "target position" had a significant effect on Primary Submovment Time (PST) and Total Movement Time (TMT). Results also indicate that the PST and TMT for targets in the upper part of the spherical framework were shorter than targets in the lower part with depth. The targets in the lower part of the spherical framework were shorter than targets in the upper part without depth. Therefore, we judged that the movement direction to the eight target positions could be taken into account by incorporating θ into the ID in Eq. (1) and (2). Based on the discussion, the ID was revised using the following formula:

$$PST = MST\ a + b(1/S) + c(D) + d(\sin\theta) \tag{1}$$

$$PST = MST\ a + b(1/S) + c(D) + d(\cos\theta) \tag{2}$$

The Secondary Submovement Time (SST) only had a significant effect on "target size". Therefore,θ was not calculated due to the lack of a significant position, size x position or distance x position interaction as discussed earlier. The study suggest a Eq. (3) for secondary submovement time. A further effort was done to develop models for primary submovement time, secondary submovement time, and the total movement time, independently. The results are presented in Table 1 and Table 2.

$$SST = a + b(1/S) \tag{3}$$

From Table 1 and Table 2 could detect R2 of Eq. (1) was higher than Eq. (2). Therefore, the study support Eq. (2) for Primary Submovment Time (PST) and Total Movement Time (TMT).

$$PST = MST = a{+}b(1/S) + c(D) + d(\cos\theta) \tag{2}$$

The amplitude of Fitts's index of difficulty was originally directly related to the amplitude of the arm/hand movement (target distance) and visual input (target size). Our modeling results of the combined data successfully describe movement duration by using target size, distance and position. However, models for secondary submovement time, and total movement time were developed independently and the R-squared values are improved from 0.38 to 0.977. Basically, target

Table 1. Summary of without depth model fitting results.

Model without depth	Parameter estimates						
	a	(2D/S)	D	1/S	sinθ	cosθ	R^2
PST-2D/S	0.787	1.318					0.96
PST-(1)	−0.364		2.674	0.088	0.002		0.849
PST-(2)	−0.364		2.674	0.088		0.002	0.849
SST-2D/S	2.301	0.561					0.46
SST-(3)	−0.895			5.949			0.977
TMT-2D/S	1.621	0.456					0.72
TMT-(1)	−1.261		8.804	0.081	−0.008		0.89
TMT-(2)	−1.261		8.804	0.088		0.081	0.894

Note: PST-Primary Submovement Time; SST-Secondary Submovement Time; TMT-Total Movement Time.

Table 2. Summary of with depth model fitting results.

	Parameter estimates						
Model-with depth	a	(2D/S)	D	1/S	sinθ	cosθ	R^2
PST-2D/S	1.05	1.057					0.85
PST-(1)	–0.508		2.338	0.127	0.173		0.843
PST-(2)	–0.497		2.398	0.117		0.126	0.878
SST-2D/S	2.506	0.431					0.38
SST-(3)	–1.273			6.808			0.923
TMT-2D/S	1.687	0.426					0.72
TMT-(1)	–1.591		9.151	0.108	–0.229		0.853
TMT-(2)	–1.399		8.881	–1.181		0.101	0.869

Note: PST-Primary Submovement Time; SST-Secondary Submovement Time; TMT-Total Movement Time.

acquisition movement consisted of two phases: an initial impulse toward the target (primary submovement time) and a deceleration phase under current control to home in on the target (secondary submovement time). The adjustment of cursor position requires hand-eye control and a back-forth closed-loop cycle to arrive a fine adjustment of cursor position. As a result, modeling of primary submovement time, secondary submovement time, and the total movement time should be different. The best model for the 3D, remote, free hand, target acquisition models are.

$$PST = MST = a + b(1/S) + c(D) + d(\cos\theta) \qquad (2)$$

REFERENCES

Accot, J. & Zhai, S. 2003. Refining Fitts' law models for bivariate pointing. In *Proceedings of the CHI 2003, ACM Conference on Human Factors in Computing Systems*:193–200.

Fitts, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47: 381–391.

Grossman, T. & Balakrishnan, R. 2004. Pointing at trivariate targets in 3D environment. In *Proceedings of the CHI 2004, ACM Conference on Human Factors in Computing Systems*: 447–454.

Jagacinski, R.J., Repperger, D.W., Moran, M.S., Ward, S.L. & Glass, B. 1980. Fitts' law and the microstructure of rapid discrete movements. *Journal of Experimental Psychology: Human Perception and Performance* 6: 309–320.

Lee, Y.H. & Ko, C.H. 2009. A remote, hand-free, 3-dimensional finger trajectory tracking system and its application. In *Proceedings of the 17th International Congress of International Ergonomics Association*: 10–16.

Liao, M., Jagacinski, R.J. & Greenberg, N. 1997. Quantifying the performance limitations of older and younger adults in a target acquisition task. *Journal of Experimental Psychology: Human Perception and Performance* 23: 1644–1664.

Liao, M.J. & Walter, W.J. 2004. Characterizing the effects of droplines on target acquisition performance on a 3-D perspective display. *Human Factors* 46(3): 476–496.

MacKenzie, I.S. & Buxton, W. 1992. Extending Fitt's law to two-dimensional tasks. In *Proceedings of ACM CHI 2003 Conference on Human Factors in Computing Systems*: 219–226.

Nakamura, T., Takahashi, S. & Tanaka, J. 2008. Double-crossing: A new interaction technique for hand gesture interfaces. In *Computer Human Interaction 8th Asia Pacific Conference*: 292–300.

Thompson, S.G., McConnell, D.S., Slocum, J.S. & Bohan, M. 2007. Kinematic analysis of multiple constraints on a pointing task. *Hum Movement Science* 26: 11–26.

Walker, N., Meyer, D.E. & Smelcer, J.B. 1993. Spatial and temporal characteristics of rapid cursor-positioning movements with electromechanical mice in human-computer interaction. *Human Factors* 35: 431–458.

Zhai, S., Milgram, P. & Rastogi, A. 1997. Anisotropic human performance in six degree-of-freedom tracking: An evaluation of three-dimensional display and control interfaces. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 27(4): 518–528.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Modeling Fitts' law

Jui-Feng Lin
Department of Industrial Engineering and Management, Yuan Ze University, Chung-Li, Taiwan

Colin G. Drury
Department of Industrial and Systems Engineering, State University of New York at Buffalo, Buffalo, New York, USA

ABSTRACT: The model proposed in Lin, Drury, Karwan, & Paquet (2009) was only tested against published data, limiting model validation. The purpose of the current pilot study was to validate the application of the general model for modeling Fitts' law in three designed experiments. Four graduate students participated in the experiments to measure their (1) ballistic movement time and variability, (2) the relationship between movement time and index of difficulty, and (3) the relationship between the number of ballistic movements and index of difficulty. The motor properties measured in the first experiment were utilized by our proposed general model to predict the individuals' relationships measured in the last two experiments. The comparisons of the experimental and the predicted relationships showed that the designed experiments were a feasible basis for further model validation. Some experimental modifications will be required for future research.

Keywords: Fitts' law, intermittent correction servo, aiming movement, goal-directed movement

1 INTRODUCTION

Using "Fitts' law" as keywords, one can easily get more than 13,600 relevant research papers using Google Scholar. The popularity of Fitts' law (1954) is mainly due to its promised results for many different types of movements, manipulations, environments, and participant populations [see (Lin, 2009) for review].

Fitts' law, as shown (Eq. 1), describes the speed-accuracy tradeoff relationship while performing self-paced aiming movements in which a human controls an object to reach a target by moving a certain distance according to his/her own determined speed.

$$MT = a + b \times \log_2 \frac{2A}{W} \tag{1}$$

where MT is movement time; a and b are experimentally determined constants; the logarithmic term is called "Index of Difficulty (ID)" where A is movement amplitude and W is target width.

Although Fitts' (1954) law was originally developed based on information theory concepts, some researchers consider that the feedback concepts of control theory might explain Fitts' law better. According to Craik, (1947, 1948) and Vince, (1947, 1948), while performing movements the human behaves as an intermittent correction servo that completes a movement by intermittently generating several sub-movements. The concept of intermittent correction servo was further applied in several studies to explain the rationale of Fitts' law.

The studies of Crossman & Goodeve, (1963/1983) and Keele, (1968) together have been accepted as viable accounts of Fitts' law. Their deterministic iterative-corrections model states that movements are made in rapid succession. Each sub-movement is assumed to travel a constant proportion of the distance and to the target in a fixed period of time (i.e., corrective reaction time denoted as t_r). With these assumptions, their model demonstrates that the total MT is a result of the product of t_r and the number of sub-movements required for completing an aiming movement. The model was further enhanced by Keele, (1968) who used an experimentally measured t_r of 200 ms and the assumed fixed proportion value of 1/7. Although the deterministic iterative-corrections model were developed with several doubtful assumptions (e.g., invariability of sub-movements and the fixed proportion value), the model shows the potential of applying control theory concepts in modeling Fitts' law.

Another explanation of Fitts' law was made by Meyer and his colleagues (Meyer, Abrams, Kornblum, Wright & Smith 1988, Meyer, Smith, Kornblum, Abrams & Wright 1990) who proposed stochastic optimized sub-movements models. Meyer and his colleagues also agreed on the intermittent feature and stated that an aiming movement was made with two or more sub-movements. However, they disagreed about the deterministic feature stated by Crossman & Goodeve (1963/1983) and suggested the existence of motor variability. To account for motor variability, they assumed that the endpoints of a sub-movement formed a normal distribution and could be predicted by the impulse-variability model (Meyer, Smith & Wright 1982, Schmidt, Zelaznik, Hawkins, Frank & Quinn, 1979). By conceptualizing individuals' strategy for coping with the motor variability of sub-movements to Minimize the Total *MT*, their multiple-sub-movement model (Meyer et al. 1990) predicts well the speed-accuracy tradeoffs relationships predicted by Fitts' law as the number of sub-movements increases towards infinity. Although Meyer and his colleagues' studies didn't explain how the corrective reaction time plays a role in our motor control system, their studies made contributions by involving motor variability while modeling Fitts' law.

More recently, Lin, Drury, Karwan & Paquet, (2009) proposed a general model that enhanced the concepts of the intermittent correction servo with four specified motor properties: corrective reaction time (t_r), ballistic movement time, ballistic movement variability, and moving behavior and strategy. In the general model, the sub-movement mentioned above was defined as the "ballistic movement" that is executed by a single movement impulse. Once it is executed, it cannot be autonomously modified until it is completed or the next ballistic movement is ready for executing. Similar to the concepts used by Crossman & Goodeve (1963/1983), the length of t_r would affect the execution of ballistic movement. However, the time required for performing a ballistic movement, called the "ballistic movement time ($t_{ballistic}$)", does not equal the length of t_r. Lin et al. (2009) hypothesized that Gan & Hoffmann's (1988) model, shown as Equation 2, could be utilized to predict $t_{ballistic}$.

$$t_{ballistic} = a + b \times \sqrt{d_u} \tag{2}$$

where a and b are experimentally determined constants.

In line with Meyer and his colleagues' motor variability concept, Lin et al. (2009) stated that the total *MT* is affected by ballistic movement variability. However, instead of the impulse-variability model, Lin et al. (2009) hypothesized that Howarth, Beggs & Bowden's, (1971) model, shown as Equation 3, could predict ballistic movement variability.

$$\sigma^2 = a + b \times d_u^2 \tag{3}$$

where σ is the standard deviation of the endpoint distribution measured in the movement direction; a and b are experimentally determined constants.

The last motor property is called the "moving behavior and strategy" that describes how a movement is composed of ballistic movements. While performing a self-paced aiming movement (i.e., Fitts'-type movement), the moving behavior and strategy can be explained by Figure 1.

As shown in Figure 1a, the aiming movement begins with the first ballistic movement that was assumed to move with d_u equal to the movement amplitude. Endpoints of the first ballistic movement as well as the others were determined by the ballistic movement variability model (i.e., Eq. 3). If the first ballistic movement's endpoints are inside the target [Region 1 in Fig. 1a], the movement ends with the first ballistic movement. If the endpoints are in Region 2 [Fig. 1b], two ballistic movements

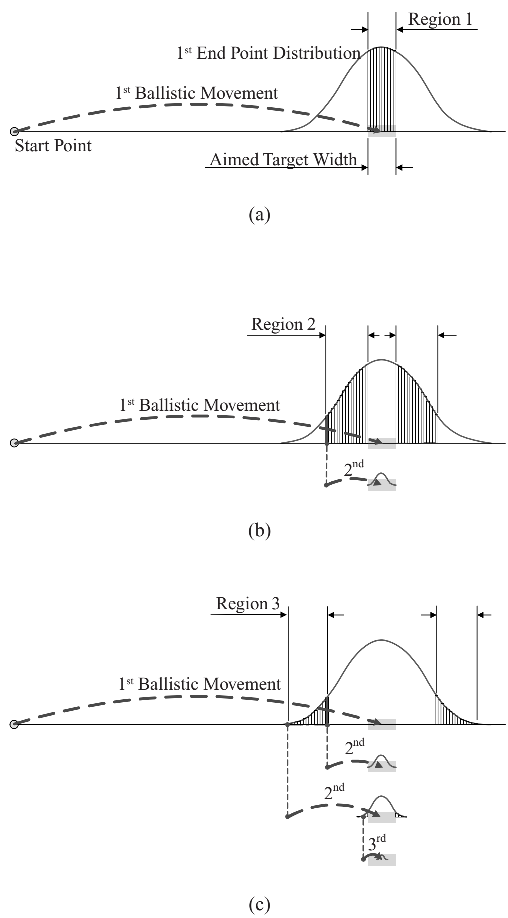


Figure 1. Moving behavior and strategy of the self-paced aiming movements (Lin et al. 2009).

are required to finish the movement. Note that the Region 2 is defined such that all the ballistic movements that start from this region can end inside the target region. And if the endpoints are in Region 3 [Fig. 1c], the movement needs either two or three ballistic movements to finish. Based on this simplified concept, we know that the endpoint distribution magnitudes and the target width together would determine the number of ballistic movements ($n_{ballistic}$) required for completing the aiming movement; the expected $n_{ballistic}$ can be obtained by multiplying every possible combination of ballistic movements for finishing the aiming movement with their associated probabilities. Furthermore, the expected total MT can be obtained by taking $t_{ballistic}$ and t_r into account. Lin et al. (2009) assumed that $t_{ballistic}$ could be predicted by Equation 2 and t_r has a reasonable range from 190 to 290 ms. They also postulated that if a ballistic movement is not the last one to finish the aiming movement and its $t_{ballistic}$ is shorter than t_r, there is a "compensatory delay" of $t_r - t_{ballistic}$ added to that ballistic movement, resulting in the same t_r as Crossman & Goodeve (1963/1983) proposed. Contrary to Crossman & Goodeve's (1963/1983) concept, however, $t_{ballistic}$ can be longer than t_r, which occurs mainly for the first ballistic movements. Furthermore, Lin et al. (2009) asked one research question: whether or not there is a "reaction delay" of $t_r/2$ between the first and the second ballistic movements, indicating the average time required to wait for the next available ballistic movement.

The general model with the moving behavior and strategy introduced above, called the "the self-paced aiming movement model", was only tested against published data in Lin et al. (2009). Due to data limitations, Lin et al. (2009) only demonstrated that the general model can predict the linear speed-accuracy tradeoffs relationship described by Fitts' law.

To further validate this general model, three experiments were designed and tested in this study, comprising (1) the experiment of ballistic movement time and variability, (2) the experiment of normal aiming movement, and (3) the experiment of ballistic aiming movement. The first experiment was designed to measure each individuals' ballistic movement time and variability and to further validate the applications of Gan & Hoffmann's (1988) model and Howarth, et al.'s (1971) model. Due to space limitations, the details of the first experiment will be discussed elsewhere. In this article, the two measured motor properties and the reasonable range of t_r from 190 to 290 ms were treated as inputs of the simulated model programmed based on our self-paced aiming movement model. The outputs of the simulated model were used to predict of the individual participants' actual performance while conducting the two types of aiming movements measured in the last two experiments.

2 METHOD

2.1 *Participants and apparatus*

Two male and two female graduate students, aged from 25–30 years, participated in this pilot study. All the participants were right-handed with normal or corrected-to-normal vision.

A Personal Computer (PC) with a 17" (432 mm) LCD monitor of 1280×1024 pixels resolution and an Intous3 305×483 mm drawing tablet was used. The PC ran Visual Basic (VB) using three experimental programs that displayed experimental tasks and measured task performance. The drawing tablet was utilized as the input device through all the three experiments. The movement distance ratio between the tablet and the screen was set as 1:1, ensuring equal visual and physical movement distances on the screen and the tablet.

2.2 *Experimental setup and procedures*

While conducting the three experiments the participants sat alongside a dual surface adjustable table on which the monitor and the tablet were placed on the rear and the front surfaces, respectively. To eliminate undesired sources of movement variation other than motor system noise, three strategies were applied. Firstly, while performing movements, the participants wore a nylon half-finger glove and lightly rested their hands on the tablet surface to keep the friction between moving hand and the tablet surface small and constant. Secondly, they were asked to move the stylus tip by moving their whole forearm and by avoiding extending/contracting fingers or wrists to make sure that the measured motor variability was generated from the same sources. Finally, a cardboard screen was placed between their eyes and the tablet to hide the visual feedback from their moving hands so that the only feedback was from the monitor screen.

After informed consent procedures, the participants conducted the three experiments in the following order: (1) the experiment of ballistic movement time and variability, (2) the experiment of ballistic aiming movement, and (3) the experiment of normal aiming movement. Each experiment started with a one-hour practice followed by one formal measurement lasting from 20 to 60 minutes. The participants individually completed all the experiments across three or four appointments within three days. The measured data of ballistic movement time and variability and the two validation experiments are presented in turn.

2.3 *Experiment of ballistic movement time and variability*

This experiment was designed to measure the participants' two motor properties: ballistic movement time and ballistic movement variability. As mentioned above, only the results are presented here. Table 1 shows the participants' ballistic movement time and Table 2 shows ballistic movement variability. Note that instead of Equation 3, it was found that Equation 4, which utilizes ballistic movement distance (d_u) as the predictor, can better predict the ballistic movement variability measured in the movement direction.

$$\sigma^2 = a + b \times d_u \quad (4)$$

2.4 *Experiment of normal aiming movement*

The purpose of this experiment was to measure the participants' speed-accuracy tradeoffs relationships while performing Fitts-type movements. The measured results were treated as "ground truth" for validating the self-paced aiming movement model. As shown in Figure 2 below, this experiment required the participants to draw lines horizontally from a start point to end within a target line. The independent variables were six *ID*s (2, 2.5, 3, 3.5, 4, and 4.5 bits) and four start point locations, used to diminish the learning of kinesthetic feedback. Each *ID* value included four combinations of target width (*W*) and movement amplitude (*A*). The four values of *W* were 8, 16, 24, and 32 pixels (1 pixel ≅ 0.266 mm), while the values of *A* were determined by Fitts' law. All experimental combinations were replicated 12 times, resulting in a total of 288 trials.

2.5 *Experiment of ballistic aiming movement*

This experiment was designed to measure $n_{ballistic}$ for completing aiming movements according to different *ID* values. Since the self-paced aiming movement model predicts the total *MT* based on the $n_{ballistic}$, the measured $n_{ballistic}$ could be utilized to validate the moving behavior and strategy shown in Figure 1. The Fitts-type movements in this experiment were designed to be performed ballistically; an aiming movement was completed by performing sequential ballistic movements. The task started by pressing down on the pen cursor on the start point. Once the cursor was moved away from the start point toward the target, the visual information disappeared and only reappeared when the ballistic movement stopped. If the endpoint of the ballistic movement was outside the target line, the participants continuously performed ballistic movements from the previous endpoints until the target region was reached. Except for the ballistic movement feature, all the other experimental details were as the same as those in the experiment of normal aiming movement.

Table 1. The measured ballistic movement times represented with Gan & Hoffmann's (1988) model.

Participant	Intercept (ms)	Slope (ms pixel2)	r^2
All	58.32	17.38	0.981
1	27.27	23.16	0.992
2	73.58	16.34	0.952
3	60.99	15.87	0.966
4	71.44	14.15	0.962

Table 2. The measured ballistic movement variability represented with equation 4.

Participant	Intercept (pixel2)	Slope (pixel)	r^2
All	−59.44	2.981	0.985
1	−113.3	3.649	0.899
2	15.22	1.684	0.819
3	−106.7	2.546	0.902
4	−79.28	4.022	0.918

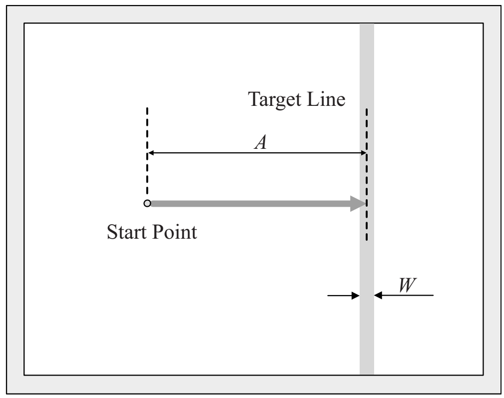


Figure 2. The movement tasks in the experiment of normal aiming movement shown on a monitor screen.

3 RESULTS

3.1 *The experiment of normal aiming movement*

The means of *MT* were regressed on to *ID* to give the slopes and intercepts shown in Table 3. Fitts' law predicted both the overall and individual participants' *MT* data very well; it accounted for 98.6% variance of the overall participants' data and at least 95.1% variance of the individual participants' data.

Table 3. Regressions of *MT* on to *ID*.

Participant	Intercept (ms)	Slope (ms/bit)	r^2
All	−71.39	111.1	0.986
1	−181.1	156.6	0.988
2	−66.68	97.96	0.955
3	0.81	90.68	0.951
4	−38.62	99.15	0.974

3.2 *The experiment of ballistic aiming movement*

The means of $n_{ballistic}$ were regressed on to *ID* to give the slopes and intercepts. As shown in Table 4, Fitts' law also predicted both the overall and individual participants' $n_{ballistic}$ data very well; it accounted for 97.8% variance of the overall participants' data and at least 87.2% variance of the individual participants' data.

3.3 *Model testing*

To test the self-paced aiming movement model, the measured motor properties shown in Tables 1 and 2, the t_r values of 190 and 290 ms, and the reaction delay of 0 and $t_r/2$ were treated as inputs of the simulated model. The outputs of the simulation were the predicted *MT* and $n_{ballistic}$ corresponding to the *ID* values measured in the two aiming movement experiments. The simulated model predicted well the linear relationships between $n_{ballistic}$ and *ID* as well as the linear relationships between $n_{ballistic}$ and *ID*. Fitts' law accounted for more than 98.5% and 97.4% variance of the simulated *MT* data and $n_{ballistic}$ data, respectively. Because both the relationships between *MT* and *ID* and the relationships between $n_{ballistic}$ and *ID* can be well accounted for by Fitts' law no matter whether the data were predicted or measured, the validation of the self-paced aiming movement model could be tested by statistically comparing the predicted and the measured linear regression lines.

Tables 5 and 6 show the comparison results of *MT* regression lines when the reaction delay was set as 0 and $t_r/2$, respectively. The highlighted values in the tables indicate no significant difference ($p > 0.05$) between the model predictions and the experimental measurements. No matter whether the reaction delay was set as 0 or $t_r/2$, only two out of 10 comparisons shows no significant difference. However, it seems that when the reaction delay was set as $t_r/2$ the model could predict better, since there are more highlighted values in Table 6.

Graphic representations of the comparisons made for all participants' data are shown in Figure 3 below. As shown in the figure, no matter what settings of t_r and the reaction delay, the model predict longer *MT*s than experimental ones.

Table 7 shows the comparison results of $n_{ballistic}$ regression lines. The model predict $n_{ballistic}$ better than *MT*. The only significant difference of regression lines was found in Participant 3's data. Graphic representation of comparisons made for

Table 4. Regressions of $n_{ballistic}$ on to *ID*.

Participant	Intercept (time)	Slope (time/bit)	r^2
All	0.7247	0.2063	0.978
1	0.5369	0.2643	0.917
2	0.7198	0.2155	0.977
3	0.8125	0.1667	0.872
4	0.8294	0.1786	0.907

Table 5. Comparisons of predicted and experimental regression lines of *MT* data when reaction delay = 0.

		Intercept		Slope	
Participant	(ms)	*t*	*p*	*t*	*p*
All	190	4.71	<0.001	−4.49	< 0.001
	290	6.19	<0.001	−4.23	< 0.001
1	190	3.39	0.0008	−3.97	0.0001
	290	3.63	0.0003	−3.80	0.0002
2	190	4.42	<0.001	−2.23	0.0267
	290	5.62	<0.001	−2.32	0.0209
3	190	−1.36	0.1753	−0.16	0.8712
	290	−1.81	0.0719	0.89	0.3726
4	190	2.18	0.0297	−2.06	0.0405
	290	3.01	0.0028	−1.53	0.1273

Table 6. Comparisons of predicted and experimental regression lines of *MT* data when reaction delay = $t_r/2$.

		Intercept		Slope	
Participant	(ms)	*t*	*p*	*t*	*p*
All	190	2.34	0.0195	1.04	0.2979
	290	1.04	0.2967	4.17	<0.001
1	190	2.07	0.0395	−1.38	0.1675
	290	1.32	0.1868	0.05	0.9606
2	190	3.29	0.0011	0.91	0.3637
	290	2.67	0.0079	2.68	0.0077
3	190	−3.61	0.0004	3.42	0.0007
	290	−4.79	<0.001	5.30	<0.001
4	190	1.08	0.2830	0.86	0.3891
	290	0.41	0.6815	2.68	0.0078

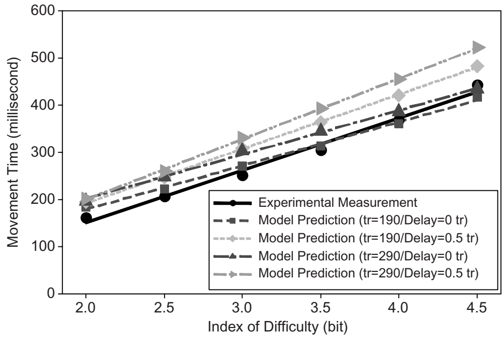


Figure 3. Comparisons of *MT* regression lines made for all participants' data.

Table 7. Regressions of $n_{ballistic}$ on to *ID*.

	Intercept		Slope	
Participant	*t*	*p*	*t*	*p*
All	–0.7407	0.4591	1.9147	0.0558
1	–0.1665	0.8679	0.4214	0.6737
2	0.4009	0.6888	–0.3021	0.7628
3	–3.5524	0.0004	3.2658	0.0012
4	–0.7007	0.4841	1.6139	0.1076

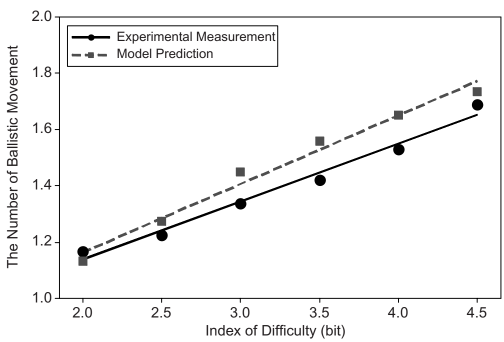


Figure 4. Comparisons of $n_{ballistic}$ regression lines made for all participants' data.

all participants' data is shown in Figure 4 above. Although there is no significant difference between the two regression lines, the model tends to predict more $n_{ballistic}$ than found experimentally.

4 DISCUSSION

The three experiments in this study were feasible for testing the self-paced aiming movement model proposed by Lin et al. (2009). Although detailed contents of the experiment of ballistic movement time and variability are not presented in this article, the experiment successfully captured the participants' two motor properties and demonstrated that they can be described by Gan & Hoffmann's model (1988) model and Equation 4, a modification of Howarth, et al.'s (1971) model. The experiment of normal aiming movement also captured the speed-accuracy tradeoffs relationship described by Fitts' law, which again shows the robustness of Fitts' law. Further, the experiment of ballistic aiming movement successfully measured the number of ballistic movements ($n_{ballistic}$) required for completing the Fitts-type movements. Surprisingly, $n_{ballistic}$ was also linearly related to *ID*. Based on strong linear relationships, the self-paced aiming movement model could be tested by statistical comparisons of the model predictions and the experimental measurements. Although the model did not precisely predict the relationships of *MT* and $n_{ballistic}$, the results showed the feasible application of the designed experiments. The comparisons of *MT* and $n_{ballistic}$ relationships showed that the model predicted longer *MT* and more $n_{ballistic}$. The reason could due to any residual learning effect—the *MT* and $n_{ballistic}$ were measured after the two motor properties. Hence, more practice or multiple measurements of the two motor properties before/after the two types of aiming movement experiments are suggested for future research. Of course, more participants should be recruited.

5 CONCLUSION

This pilot study used three experiments for validating the Lin et al. (2009) self-paced aiming movement model developed to model Fitts' law. The motor properties of ballistic movement time and ballistic movement variability measured in the first experiment were utilized as inputs of the model. The statistical comparisons of the model outputs and the experimental measurements obtained in the last two experiments showed that the designed experiments were feasible for further testing. Some modifications of the experiments were suggested for future research.

REFERENCES

Craik, K.J.W. 1947. Theory of the human operator in control systems I: The operator as an engineering system. *British Journal of Psychology* 38: 56–61.

Craik, K.J.W. 1948. Theory of the human operator in control systems II: Man as an element in a control system. *British Journal of Psychology* 38: 142–148.

Crossman, E.R.F.W. & Goodeve, P.J. 1963/1983. Feedback control of hand-movement and Fitts' law. *Quarterly Journal of Experimental Psychology* 35: 251–278.
Fitts, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47: 381–391.
Gan, K.-C. & Hoffmann, E.R. 1988. Geometrical conditions for ballistic and visually controlled movements. *Ergonomics* 31: 829–839.
Howarth, C.I., Beggs, W.D.A. & Bowden, J.M. 1971. The relationship between speed and accuracy of movement aimed at a target. *Acta Psychologica* 35: 207–218.
Keele, S.W. 1968. Movement control in skilled motor performance. *Psychological Bulletin* 70(6): 387–403.
Lin, J.-F. 2009. *A unified model for self-paced movements, industrial & systems engineering.* The State University of New York at Buffalo, (unpublished dissertation).
Lin, J.-F., Drury, C., Karwan, M. & Paquet, V. 2009. A general model that accounts for Fitts' law and Drury's model. In *Proceedings of the 17th Congress of the International Ergonomics Association*, August 9–14, Beijing, China.
Meyer, D.E., Abrams, R.A., Kornblum, S., Wright, C.E. & Smith, J.E.K. 1988. Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review* 95: 340–370.
Meyer, D.E., Smith, J.E.K., Kornblum, S., Abrams, R.A. & Wright, C.E. 1990. Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In *Attention and performance XIII: Motor representation and control*, edited by M. Jeannerod. N.J.: Lawrence Erlbaum.
Meyer, D.E., Smith, J.E.K. & Wright, C.E. 1982. Models for the speed and accuracy of aimed movements. *Psychological Review* 89 (5): 449–482.
Schmidt, R.A., Zelaznik, H.N., Hawkins, B., Frank, J.S. & Quinn, J.T. 1979. Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review* 86: 415–451.
Vince, M.A. 1947. The intermittency of control movement and the psychological refractory period. *British Journal of Psychology* 38: 149–157.
Vince, M.A. 1948. Corrective movements in a pursuit task. *Quarterly Journal of Experimental Psychology* 1: 85–103.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Optimum strategies for urgent evacuation in aircraft accident using AAMAS simulation

Tetsuya Miyoshi
Toyohashi Sozo University, Toyohashi, Japan

Hidetoshi Nakayasu & Yuki Ueno
Konan University, Kobe, Japan

Masao Nakagawa
Shiga University, Hikone, Japan

ABSTRACT: In this paper, a new optimization method is proposed to minimize the evacuation time of passenger in an aircraft accident. In order to minimize the total evacuation time of the passenger, an optimization technique is examined as a multi-agent problem using effective evacuation guidance by cabin attendant or the urgent safety system in an aircraft accident. The urgent evacuation behavior is formulated as an Autonomous Agent and Multi-Agent System (AAMAS) model evolving over a two-dimensional grid cell that represents the aircraft cabin and passenger. The effect of the proposed method was illustrated with some numerical simulations.

Keywords: safety engineering, AAMAS model, optimization, aircraft evacuation

1 INTRODUCTION

It is well known that the development of effective evacuation system is the primary task for aircraft safety, while it is so critical for aircraft safety in case of emergency accident that the skill of flight crews and cabin internal arrangement must concentrate to save human lives. From these points of view, Federal Aviation Administration (FAA) makes its attention on these issues, therefore the new aircraft must satisfy several rules as Federal Aviation Regulation (FAR) Part 25.803 (FAA, 1990) where one of them is called "90 seconds rule." This rule means that the maximum seating capacity including the number of crewmembers can be evacuated from the airplane to the ground under simulated emergency conditions within 90 seconds. However, it is very difficult because the egress time necessary for evacuation is concerned with so many factors such as airframe (number, size and location of emergency exit, seat and aisle arrangement), passengers (age, health, sex, interrelationship and degree of panic) and flight crew (skill and training level).

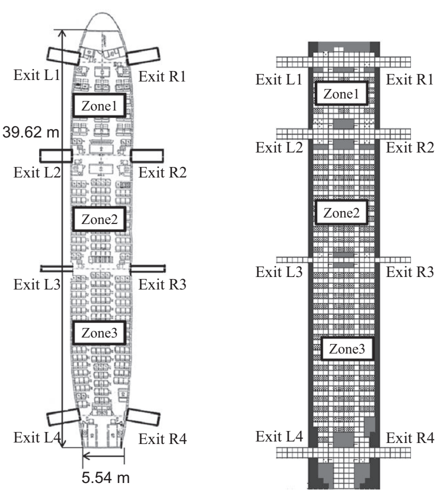
The practical and traditional approach for evacuation experiments with participants are executed by airplane manufacturers (Movtevalli, et al. 2008). However, these experiments are dangerous and expensive and not easily repeatable. For these reasons, a dynamic model for aircraft evacuation system by Ceruti and Manzini (Ceruti, 2003) is proposed as a visual interactive simulation tool in order to develop and evaluate the several evacuation techniques. In general, it is too difficult to solve the optimization problems for minimum evacuation time exactly since these problems are not only able to formulate analytically but also to solve the problem under chaotic uncertainty of incomplete or dynamic stages of urgent evacuation.

In this paper, it is tried to provide a new optimization method for the flight crews and the cabin attendant how to minimize the evacuation time of passenger using AAMAS simulation and recursive procedure in an aircraft accident. The main idea of optimization dues to use and update the quasi-solution of optimization problem obtained sequential AAMAS simulation and search the satisfactory solution among them by a recursive algorithm. Therefore the quasi-solution obtained by the proposed procedure can make the passengers possible to select the best route to emergency exit in an aircraft accident. In order to investigate the optimum way for guiding route to emergency exit from aircraft cabin, it is assumed that the agent moves to the emergency exit under the constraint of previously specified local rules. The individual behavior of an agent controlled by the local rules makes a macro behavior of group as an evacuation flow in an aircraft cabin automatically synchronized to the results of the movement of an agent. From the viewpoints of the above, it is recommended in this paper that the passengers have not only the information of avoiding the traffic jam of evacuation but also the information of the degree of length of waiting queue in order to find the effective and satisfactory evacuation guide. Some comparative studies with and without using the proposed optimization are also illustrated for the verification of the optimization.

2 SIMULATION MODEL BY AAMAS

2.1 *Model as analogies of passengers and equipments*

In this paper an Autonomous And Multi-Agents System (AAMAS) model is applied to construct the simulation model where a two-dimensional grid cell model shown in Figure 1(b) is introduced to represent the passenger flow of the emergency evacuation in the aircraft. Figure 1(a) also represents a layout and allocation of equipments in cabin of the aircraft DC-10-30, where it is seen that there are the three passenger cabins, zone1, zone2 and zone3. There are 24, 98 and 155 passenger seats in each zone, respectively. The grid cell model in the Figure 1(b) is used as an analogy of the seat of economy class in the cabin whose size is approximately 0.43 m square. The equipments such as exit doors, exit signs, lavatories, galleries, counters, aisles and seats are also identified as approximately analogies in two-dimension model so as the location and dimension of an aircraft. Passengers are represented as multi-agents that are initially placed in seat squares and move around the grid by one's objectives whose behavior is controlled by the autonomous algorithm based on local rules previously specified. The generic passenger's run speeds in evacuation were investigated by Galea (Galea, 2007), and the computerized evacuation simulations were performed using these evacuation speeds by Ceruti (Ceruti, 2003). The time step of computer simulation are determined based on the average 1.14 (m/s) of the evacuation speed, i.e., simulation step interval is set to 0.37 sec/step. A time variable, T, is set to 0, and incremented by 1 every 0.43 s. It is assumed in the previously specified rules that the multi-agents move to the exit each other in order to evacuate from an aircraft. It is noted that though the individual behaviors of multi-agents are controlled by the local and autonomous algorithm, the final results of the movement of individuals by the autonomous and multi-agents will be seemed to yield a macro behavior controlled by group dynamics as an evacuation flow in an aircraft cabin.



(a) Cabin interior of DC-10 (b) Two-dimensional cell model

Figure 1. Cabin interior and cell model.

2.2 *Algorithm for an autonomous movement of multi-agent*

The flow of passengers towards the exits is determined through four steps. Each agent is then instructed to move towards the nearest exit. They gather the information on the location of emergency exits or exits signs at the first step. At the second step, they memorize the nearest direction into emergency exits or exits signs as an own action object if they could recognize the emergency exits or exit signs, otherwise they forecast the direction to the nearest exit route as a candidate object. In this step the direction is determined among the 16 directions as shown in Figure 2. At the third

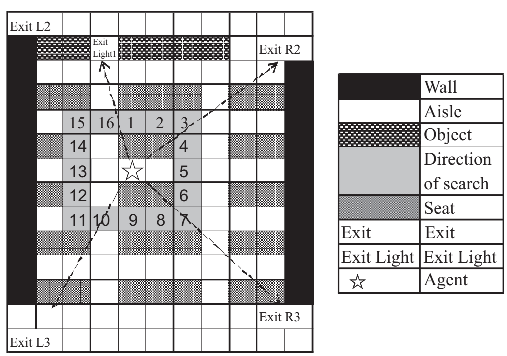


Figure 2. Field of view of agent and movement rule.

step, they determine the adjacent cell to move by considering both the memorized direction and current location itself, which is detected as a critical path so as to take the shorter time to go for exit subject to the physical limitation such as seats and equipments allocation. Adjacency is defined by the Moore neighborhood (i.e., the eight surrounding patches). At the forth step, confirming whether the cell to move is occupied or not, they moved to the cell if it is not occupied, otherwise they must wait until the cell is empty. When a lot of passengers would go for the same exit over the capacity of exits, the queuing of passengers happened.

3 OPTIMIZATION ALGORITHM

3.1 *Efficient guidance system*

The goal of the evacuation simulation is to understand and recognize the property of urgent evacuation in the aircraft so as to consider the several conditions such as allocation of passengers, efficiency of exits, a method to instruct passengers and so on. In this paper the goal of urgent evacuation simulation analysis is to consider the safety evacuation plan satisfying "90 seconds rule" by FAR Part 25.803 (FAA, 1990). The evacuation guidance to passengers by crewmembers is necessary to achieve this goal. Therefore the instruction for the guidance must be planed to minimize the total egress time of all passengers and crewmembers from the cabin to ground on the air port. Subject to the agent moving rule, the queue of passengers to some exits happen since all the passengers tend to go out through the nearest emergency exit and the concentrate to a specific some exits memorized as an own action object. Based on the above consideration it is set up as a optimization discipline that the allocation of passenger to the emergency exits will be determined so as that the number of passengers who would go out through each emergency exits should be shared in order to balance the equally-divided number of passengers into the each emergency exits.

If the sizes of seat capacities in cabins are different with each others, the numbers of passengers to the emergency exits are unbalanced for the other emergency exit. Figure 3 shows the numbers of passengers to the emergency exits with and without optimization. It is seen from the figures that a larger number of passengers would go out the exit equipped near the larger cabin (zone 3). In Figure 3(a) the ellipses illustrate the area in which the passenger would go to the corresponding exit, and the number of passengers who would chose each exit is also represented in the figure without optimization when the passenger could see wide area in aircraft with long view length. The waiting queues of passenger to the emergency exit3 and exit7 happen and it was also estimated that the complete egress time would be longer.

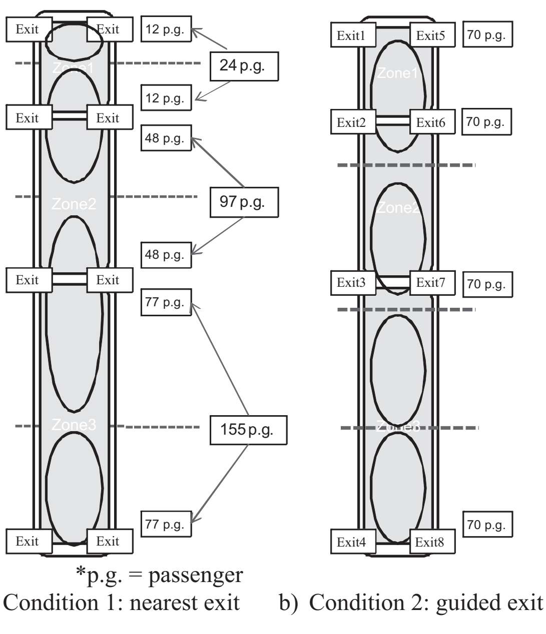


Figure 3. The number of passengers out of exits.

Figure 3(a) shows 125 passengers concentrated to the emergency exit3 and exit7, and a lot of queuing passengers would happen if the passengers to each exit can be share to balance the queuing passengers to the specified emergency exit, the total egress time would be reduced. In order to induce the optimum solution, it is attempted that the optimum allocation of passengers to each exits are determined considering the moving length to the exit, but it is difficult to evaluate the length previously because of the complex allocation of cabin equipments. Thus, it is noted that it is impossible to obtain the optimum allocation of passengers to the emergency exits previously analytically. Instead of analytical optimization, it is tried as a quasi-optimization to obtain the solution of satisfactory allocation of passengers from the viewpoint of reducing the moving distance considering the running speed of the passenger and the queuing length of passengers the emergency exit.

3.2 *Initial allocation of passenger to the exits*

The number of passengers who would go out through the exit should be determined equally corresponding to the evacuation efficiency of the exit in order to minimize the complete evacuation time. If the exit is large and the more passengers could go through it in unit time, more passengers should be induced to evacuate through it, i.e., cabin attendants or crews should instruct their aim is set to the exit. Figure 3(b) shows the initial setting of the passenger's aim when the efficiency of all exits is same level with each other, in which exit1 and exit5, exit2 and exit6, exit3 and exit7, exit4 and exit8 are set to 70 passengers equally with each other. If the passengers whose aims are set the same exit are located separately, the passenger's evacuation flow to the exit is crossed and the jams of passenger flow are raised and delay the complete evacuation time. In order to avoid the separate allocation of passengers whose aims are set to the same exit, the passengers' aims are set in order at location of them sequentially. The aims of passengers in front of the aircraft are set to the first exit or the fifth exit up to the predetermined limitation. After that the aims of passengers are determined continuously.

The number of passengers in the aircraft is denoted by N, and the ID of passenger is set from the passenger allocated in front seat sequentially. The number of exits is denoted by D and the ID of left exits are set to 1, 2, 3, 4 and the ID of right ones to 5, 6, 7, 8. The efficiency of the *i* th exit is denoted by the e_i, where the efficiency of exit mean the rate of the number of out-going passengers per unit time. The number of passengers whose aims are set to the *i* th exit, p_i, is determined by

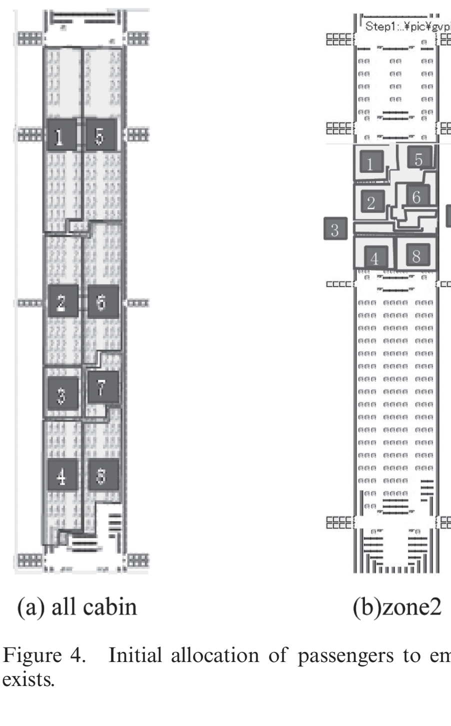
$$p_i = N\frac{e_i}{\sum e_i} \quad (i = 1, 2, \ldots, 8) \tag{1}$$

If the exit is closed by something reasons, such as the fire or some trouble, e_i are set to 0. Then, in the configuration where the 4 exits are closed and the efficiency of the 4 exits is equal with each other, the number of passengers to each exit is set to $N/4$.

Next, the aim of passenger whose ID is smaller are set to first or fifth exit up to $p_1 + p_5$. After the first exit and fifth exit are allocated to passengers, the aims of passengers are set to the second or sixth exit. According the above mentioned procedure, the aims of all passengers are set to exits, i.e., they are induced to evacuate the setting exit. Figure 4 shows the initial setting of some numerical examples, in which the numbers rounded with square denote the exit number which the passengers of the area would be set to.

3.3 *Optimization by recursive procedure*

The initial allocations of the aim of passengers to the exits are determined according to the exit efficiency, however, the evacuation time would not necessarily be minimized because the moving time is not considered. It is difficult to formula and solves the optimization problem to minimize the complete evacuation time. Therefore, the optimization is calculated by the recursive procedure using multi-agent evacuation



(a) all cabin (b)zone2

Figure 4. Initial allocation of passengers to emergent exists.

simulation. The latest time that an agent went out are recorded for each exit. A small number of passengers are allocated to the exit whose latest time is small. In other words, the when the difference between maximum and minimum of latest time of all exits are large, the passengers allocation to the exit are unbalanced, and consequently the egress time became long. Considering the difference between maximum and minimum of latest time, the initial allocation of passengers are determined by heuristics. In the heuristics the number of passengers to the exit whose latest time is minimum is added with c_1 and one whose latest time is maximum is reduced with c_2. c_1 and c_2 is the determined by the try and error.

4 SIMULATION RESULTS FOR AIRCRAFT EVACUATION

4.1 *Comparison in the configuration with load factor*

In this paper the evacuation simulation for the DC-10-30 aircraft were performed. The DC-10-30 has 8 exits. The third and seventh exits are narrower than other exits actually. The capacity of them is half of the one of the others. Therefore, even though the numbers of passenger allocations to each exit were equal, the queue of passengers allocated to the third or seventh exits would happen, and then the latest time of complete evacuation at them would become large. The allocation of passengers to exits should be determined with considering the capacity of the exit.

The several evacuation simulations with some experimental conditions are performed to compare between the results by the agent moving independently and by the proposed optimization method. In this section the recursive algorithm will not be applied in order to discuss the effect of the initial setting of the aim of passenger.

Load factor, which is the rate of passenger for the seats in an aircraft, are set to100%, 75% and 50%, and the passenger's initial locations are set randomly, and the passenger could see over all with the long visual length. In the simulator, the visual length is set to 20 cells. The initial locations of passengers in three conditions are shown in Figure 5, in which the passengers are marked with △. Figure 6 shows the comparison with the rate of remaining passengers in the results simulated by the proposed method and by the agent moving rule. The evacuating flows are shown at the 20 steps before the evacuation have been completed. The evacuation complete time became shorter in the results simulated by the proposed method than the one by only agent local rule in all condition with load factors. The passengers concentrated to the eixt3 or exit7 in evacuation by the only agent local rule, on the other hand, the passengers went for some exits by the proposed initial setting in all cases. These results show that the proper allocation of passenger to the exit has been performed

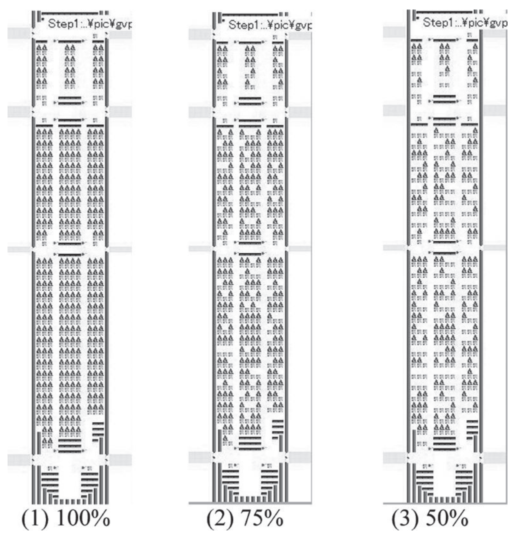


Figure 5. Initial allocation of passengers in each condition.

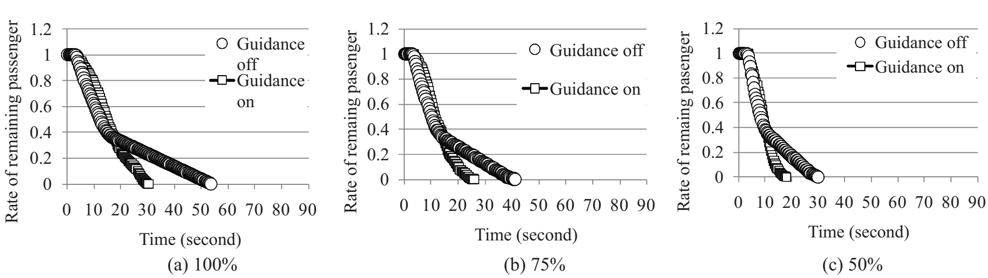


Figure 6. Simulation results in case that all exits are open.

so as to balance the number of passengers to the exits and the allocation could minimize the total evacuation time.

4.2 *Evacuation simulation by optimization using recursive procedure*

In this paper the objective exits of passengers are determined based on the efficiencies of exits and location of passengers firstly. Considering the result of evacuation simulator with the initial condition, the condition, e.g., the number of passengers who should be guided to the exit is determined in next iteration. These procedures are repeated iteratively. Table 1 shows the iterative procedure for the simulation, in which the load factor is 50% (passenger 138) and 8 exits are opened. Firstly, the numbers of the passengers allocated to the emergency exits are determined by equation (1), which are shown at the first table in Table 1. The exits which each passenger should go out from are determined in order from the front side so as not to exceed the predetermined number of passengers and the evacuation simulation is performed.

1st iteration

allocation number				result(latest time)				complete step
exit1	20	exit5	20	exit1	48	exit5	30	51
exit2	20	exit6	20	exit2	46	exit6	43	
exit3	10	exit7	10	exit3	35	exit7	33	
exit4	20	exit8	18	exit4	29	exit8	34	

2nd iteration

allocation number				result(latest time)				
exit1	18	exit5	20	exit1	48	exit5	30	51
exit2	20	exit6	20	exit2	46	exit6	45	
exit3	10	exit7	10	exit3	35	exit7	31	
exit4	22	exit8	18	exit4	31	exit8	36	

3rd iteration

allocation number				result(latest time)				
exit1	16	exit5	22	exit1	48	exit5	32	51
exit2	20	exit6	20	exit2	46	exit6	45	
exit3	10	exit7	10	exit3	32	exit7		
exit4	22	exit8	18	exit4	31	exit8	36	

4th iteration

allocation number				result(latest time)				48
exit1	12	exit5	20	exit1	45	exit5	37	
exit2	20	exit6	20	exit2	43	exit6	42	
exit3	10	exit7	12	exit3	30	exit7	34	
exit4	24	exit8	18	exit4	33	exit8	36	

5th iteration

allocation number				result(latest time)				48
exit1	10	exit5	22	exit1	45	exit5	38	
exit2	20	exit6	20	exit2	42	exit6	38	
exit3	12	exit7	12	exit3	34	exit7	33	
exit4	24	exit8	18	exit4	33	exit8	36	

6th iteration

allocation number				result(latest time)				49
exit1	8	exit5	22	exit1	48	exit5	30	
exit2	20	exit6	20	exit2	46	exit6	43	
exit3	12	exit7	14	exit3	35	exit7	33	
exit4	26	exit8	18	exit4	29	exit8	34	

Table 1. Quai-optimization using recursive procedure.

The latest time of passenger's evacuation complete were measured as the simulation result, which are shown in the second table in Table 1. The evacuation simulations would be repeated while the evacuation complete time is reduced. When the complete time is increased, the recursive procedure is finished and the exits which the passengers are guided to are determined as the optimization solution. The change of initial allocation of passengers by the proposed optimizaion method is illustrated in Figure 7.

4.3 *Discussion*

The passengers go for the near exit based on the collecting information about the exit and interior equipments if they were not guided in aircraft. Since the exits are not necessarily designed evenly in the aircraft, the queue of passenger to exits would happen, especially the queue to the exits located in center of aircraft would tend to become longer. In DC-10-30 the above situation is remarkable since efficiency of the third exit and the seventh exit is half of one of the other exits. The proposed guidance system is a powerful method to avoid/reduce the queue of passengers to exit by making the number of passengers for the exits equal. In order to improve the performance of evacuation, the exit which each passenger would be guided to is changed with iterative procedure base on the simulation result.

Figure 6 represents the residual rate of passenger evacuation. The performance of evacuation in the latter half was lower than one in the first half since the passengers concentrate the third and seventh exits. Thus, the performance of evacuation is evaluated by the rate of residual passenger in the aircraft. In our proposed optimization method the exit which each passenger should be guided to is determined by the previous simulation result. In order to guide the passenger corresponding to the situation in the aircraft accidents, not only the previous simulation results but also the performance of the passenger evacuation shown in Figure 6 should be used necessarily.

5 CONCLUSION AND FUTURE WORKS

In the proper guidance the exit which the passenger should go for is informed to the passenger, which is determined by the proposed iterative procedure using the evacuation simulation system. The effect of the proposed method was illustrated with some numerical simulations. The future work is to improve the proposed method to change the guidance to passenger according to the situation.

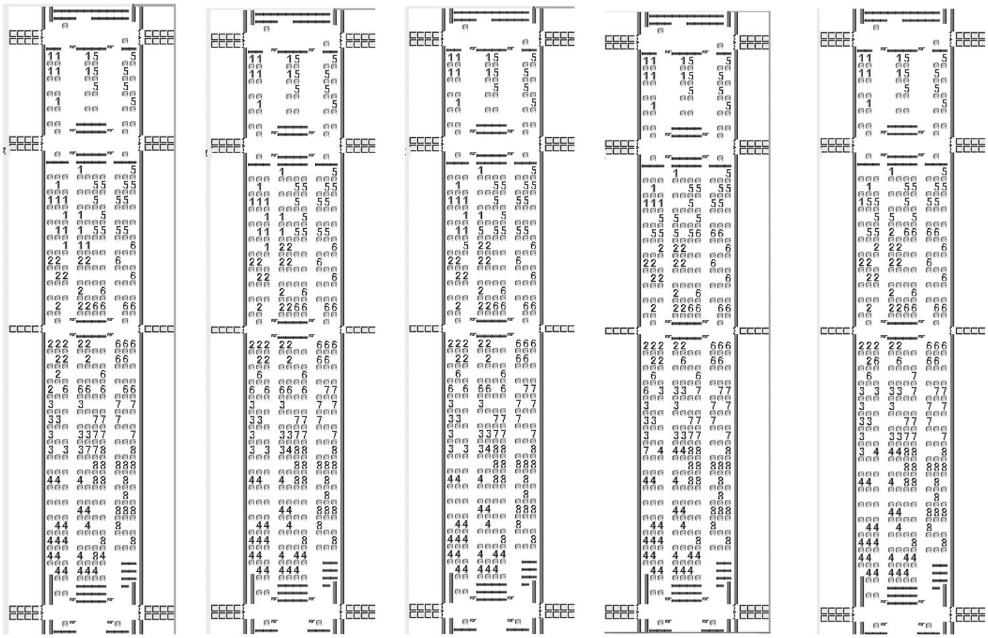


Figure 7. The change of initial allocation of passengers by the proposed optimization method.

REFERENCES

Ceruti, A. & Manzini, R. 2003. Aircraft evacuation dynamic model based on simulation tool, *Proc. Advanced Simulation Technologies Conference 2003*, Orland.

FAA 1990. *Federal Aviation Regulation, Code of Federal Regulations, Part 25.803 Emergency Evacuation, Part 25 Appendix J. Part 121.391 Flight assistants*. Federal Aviation Administration.

Galea, E.R. 1999. Predicting the evacuation performance of mass transport vehicles using computer models. *Focus on Commercial Aviation Safety* 34: 14–15.

Galea, E.R., Wang, Z., Togher, M., Jia, F. & Lawrence, P. 2007. Predicting the likely impact of aircraft postcrash fire on aircraft evacuation using fire and evacuation simulation In *Proc. the Fifth Triennial International Fire & Cabin Safety Research Conference, 29 October –1 November, Atlantic City 2007.*

Movtevalli, V., Monajemi, L. & Rassi, M. 2008. *ACRP Report 2 -Evaluation and mitigation of aircraft slide evacuation injuries*. Transportation Research Board of the National Academy.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Research on the script language as a fundamental coordination mechanism for agents of spatial objects

Chiung-Hui Chen
Department Media and Design, Asia University, Wufong, Taichung County, Taiwan

ABSTRACT: The agent theory and technology is currently an important research issue in artificial intelligence. The agent-oriented design is a new generation method of programming design that mainly uses the script language as a basis for development and that is suitable for the design of multi-agent application system. In the urban design environment, suppose design objects such as pavements, lamps and plazas can look after themselves, where would a design object acquire the knowledge that allows it to interact intelligently? This paper investigates the question through the notion of objects as agents in design. In order to explore the relationships between the design pattern and the agent model, this research has developed an interactive agent model with the design pattern as an orientation to express the interaction of each individual with the multiagent-based model. The preliminary research adopts the way of the script language to describe the agent behavior, converting a design atlas into an interactive agent pattern; that is, to construct the mechanism and behavior of heterogeneous agents.

Keywords: script, coordination, multi-agent, behavior, pattern

1 INTRODUCTION AND PURPOSE

In recent years, the applied theory of agent-based modeling has been rich in researches on economy, ecology, environmental planning, population prediction, etc. In the aspect of studying pedestrians in the field of spatial movement, the agent-based model is an important research method. The biggest advantage of using the agent-based model to simulate the pedestrians is that when the problem about aggregates is considered, the way of thinking which is constructed from down to up corresponds to the real phenomenon of pedestrian movement. The agent-based model regards each individual as an actor who will generate perception and behavioral reaction toward the surroundings. It is constructed by means of the interactions among these actors as well as the reactions toward the characteristics of the surroundings. This way, the pedestrians' movement condition which is closer to the real situation can be more reflected and the future change can be further predicted. With different application needs, there are also obvious differences in its agent program and framework. The construction mechanism of the agent design system and the agent program related with the field of design explore mostly the problems such as how to own the design information, how to present knowledge, how to process automatically and how to solve communication. However, the agent model which is pedestrian-based only is still insufficient for the environmental design. The activities in a spatial environment and the design planning of substantial objects also require a mechanism to determine heterogeneous agent behaviors.

"A Pattern Language" is written by Christopher Alexander in 1977. The implied key thought is to presume that certain structured features will not change with time during the construction process in an artificial environment. And these features will dominate mankind's living behavior along with the generative ways of a spatial environment. In other words, the internal mechanism which can be systemized by humans' thoughts and the pattern which can be offered for general operation are hidden behind these structured features. Similar design questions can be solved in similar ways. A pattern, whose content consists of the description of problems and solutions as well as the relevant situations or contextualized information, is used to record a general solution corresponding to a problem. On the other hand, due to the agent theory is suitable to obtain and analyze the planning requirements of interactive coordination, an agent behavior which is basically the same as the operational pattern of human society is autonomous. The agent system will divide the jobs which are to be carried out into numerous sub-jobs according to the complexity of problems, and then will allow these program modules to take responsibility and complete the jobs individually. Sub-jobs can also be divided into smaller sub-jobs and let the granulized modules inside the modules to complete the job by means of cooperation. Consequently, the whole system becomes a multi-agent environment formed by various agents who base on functions along with missions and who follow certain specific subordinate, parallel or independent relationships. In order to explore the relationships between the design pattern and the agent model, this research has developed an interactive agent model with the "design pattern" as an orientation to express the interaction of each individual with the multiagent-based model. The preliminary research adopts the way of the script language to describe the agent behavior, converting a design atlas into an interactive agent pattern; that is, to construct the mechanism and behavior of heterogeneous agents.

2 LITERATURE REVIEW

Based on the goal required in this research, all the important literature collected at this moment will be further illustrated according to different issues. The comments and descriptions are summarized as follows.

2.1 *Design patterns for parametric modeling*

The knowledge database of parametric design pattern is developed and constructed from the organization of the parametric model assiduously promoted by Canadian Design Research Network (CDRN) and Smart Geometry Group (Woodbury, 2007) both of which have co-hosted plenty of seminars and workshops. CDRN even further pass on the developed knowledge, accumulate the results, disclose them on website platforms for sharing and extend more contents of the knowledge database of design pattern.

The research on this model emphasizes to explore the influences that the fluctuation of design development procedure has on the design result. For instance, a grammar-based design system can understand the designer's behavior, derive design, explore the patterns of specific areas, or explore the composition and the expression of form, etc. In this research, it is considered that the knowledge must be organized and recorded so that it is favorable for re-utilization. The content of knowledge based on the parametric design must contain the operational process which is used to solve the design problems. And the concept of Christopher Alexander's design pattern must be adopted as a way to record knowledge. Besides, the sharing record of the pattern is divided into six parts to be recorded: *Example, How, Use When, Why, Intent & Related patterns*, and *Sample*. The establishment of shared website platforms for the knowledge database of pattern along with the experience of CDRN and the ways they do are all the references for this research.

2.2 *Interactive script mechanism*

In recent years, many scholars have proposed the interactive script mechanisms to control agent behaviors. "*Q-Language*", an interactive script mechanism proposed by Professor Toru Ishida, Department of Social Informatics, Kyoto University, Japan, in 2003, is similar to the humans' natural language pattern. Its purpose is to target the web-based agent system whose internal agent behavior can interact with users by means of the files driven by the script description, and further to assist the application of simulation of crowd movement management in disasters in the urban public space of the real world.

In the *Q-Language*, three major elements are required; they are *Cue and Action*, *Guarded Commands*, and *Scenarios*. A script file written with the *Q-Language* can propel the agent behavior after it is equipped with the above three elements. However, these elements and the definitions of signs are not easy to understand. Therefore, for the convenience of users to write the script file, the *Q-Language* designs an editing interface in the tabulated form. All the designers have to do is simply to fill the content in the table, and then the content in the table can be converted to the standard *Q-Language* description file by means of the implemented compiler. Therefore, through the interaction mechanism inside the script, pedestrians will respond and trigger specific event behaviors at the most of the time when they detect other objects or collisions in an environment. The motivation to trigger the behavior can be roughly divided into two: the event-driven random behavior and the scripted behavior. To let the pedestrians show different kinds of decision-making power in the crowd simulation, users can only have the partial control right although they can manipulate the pedestrians' activities. The pedestrians will maintain part of rule-based behavioral autonomy. Even the degree of behavioral intensity and feebleness can be shown with a parameterizable behavioral model. Varied behavioral patterns are adjusted in terms of different targets. And the generation of random numbers or the direct modification in a simulated status is also included.

3 RESEARCH PROCESS

Jan Gehl, an urban design scholar, indicated in his book "*Life between buildings: Using public space*" that the urban life can be divided into two categories: one is the public urban life, social and interrelated, which includes streets, squares, parks, and active commercial areas where the activities are intense. These lives mostly appear in the open space of a city. People gather here and participate in diversified urban activities; that is, the so-called "social life", in which the open-air cafes along the sidewalks and watching plays at night as well as the performances at the squares and so on are included. Such a life is busy and vivid. The other is the private and individual urban life. It is a self-oriented life, in which one can find the sense of belonging and the needed privacy. Such a personal activity requires another form of space, and to stay away from the jam-packed crowds for peace. Consequently, the city should be able to reflect simultaneously these two kinds of needs and activities. The observation in this research is based on the first category which is the social open space in a city.

Figure 1. The good place on the street.

3.1 *Exploration of patterns*

The analysis on behaviors in the urban space is important for urban designs and urban researches. Pedestrian-only shopping streets are characteristic of unique street designs. In this research, by observing a series of processes in which the pedestrians experience in the space of pedestrian-only shopping streets, pedestrian behaviors and the interactions of environmental information are recorded and analyzed. Useful and interesting issues are extracted from activities and converted into a general pattern content for reference so that they can be a basis for the examination of the case where the walking space of a street is studied. Then, through the deduction of the case implementation, a creational pattern of good quality regarding the space of pedestrian-only shopping streets is established. As shown in Figure 1, the preliminary exploration into famous pedestrian-only shopping streets in Taiwan is made in this research. It is found that the space of streets generates the cluster points of activities because of the accumulation of the crowds. Hence, the novelty of elements, the way the objects present in a space, and the relative relationships which foreground in the environments such as a substantial environment, are all the significant constitutional elements.

Figure 1 The good place on the street was taken as examples here. Their similarity is the business behaviors and performances which often take place on the sidewalks in front of the shops, and this generates crowds. Therefore, interesting street performances which attract pedestrians' attention could possibly increase the attraction of the locations where performances are made. It is easier for shops with outdoor coffee tables and seats to attract pedestrians and form a stationary point. In addition to the original functions of the streets, diversified advertisement signs and different types of products displayed in arcades can easily build direct conversational relationships with pedestrians, who tend to pay attention to novel furniture; as a result, colorful, bright, large furniture with effects such as unique shapes, textures, surrounding environments and dynamic presentations can easily attract people's attention and generate crowds.

3.2 *Record of objects on the street*

Objects are the facility elements used when pedestrians generate the events and activities in an environment of streets. They record with which of the objects the pedestrians generate the events and activities in an environment of streets. These objects are the furniture on the street, windows of the shops, plants, railings, etc. They are recorded in the forms of photographs, drawings, video recording, and description of the written language. The qualities and categories of the objects are categorized. And the locations of these environmental objects are recorded. What the original functions of environmental objects are, and how the functions generated by environmental objects in the generation of events and activities are used, are also recorded.

3.3 *Induction of patterns*

By observing real cases, a representative design pattern is extracted and analyzed. Then the format of the pattern is set. After the organization and compilation, a universal and fundamental content of urban space design is indicated. This research quotes Christopher Alexander's concept of *Design Pattern* as a way to record knowledge. A pattern typically comprises a name, a problem description, an abstract solution and a discussion of consequences (Alexander, 1977). Take the observed case in Figure 1 as an example. The patterns used include: the bag-shaped venue, the location of open-air cafes, soft paving, street lamps in warm colors, and the areas equipped with lighting. By taking them as an example, we can put the patterns into a linear sequence according to the requirement order of patterns. This sequence can clearly express the subordinate relationships among the patterns. Analyze the ways in which the objects present in a space: (1) Bag-shaped venue, (2) Location of open-air cafes, (3) Soft pavement, (4) Street lamps in warm colors and (5) Areas equipped with lighting.

3.4 *A design pattern converted into an agent of spatial objects*

Accordingly, Multi Agent Based Simulation is based on the idea that it is possible to represent the global behaviour of a dynamic system as the result of interactions occurring among an assembly of agents with their own operational autonomy (Wooldridge, 2002). Each of the spatial objects contained in the design pattern is endowed with features and behavioral properties to establish the dependent relationships of one-to-multiple, one-to-one, and multiple-to-multiple spatial objects. This way, once the status of a spatial object changes, other dependent objects will be automatically notified in order to make the corresponding updates. Then, the relationships among the categories of objects are converted to the agent concepts. The contents that each agent of the object must define contain six items: *Agent Type, Goal, Environment, Percept, Knowledge,* and *Action* as shown in Figure 2. Take the pattern of "Take a rest" as an example. In an urban open space, compile and design an outdoor venue which can afford at least 5 people to take a rest after they take a stroll and do the shopping. The goal is to form an area which is equipped with lighting and which allows pedestrians to stop and take a rest in an urban space at night.

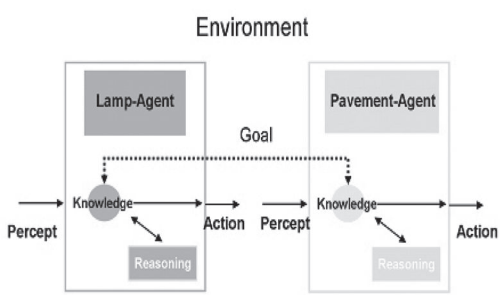


Figure 2. Interactive patterns among the agents.

4 DISCUSSIONS

After the above research process, we propose several important suggestions regarding the principles. The discussions are as follows.

4.1 *A design pattern can be regarded as a similar concept of script function*

The idea of "patterns" reflects the perceived regularities in an environment. Alexander's patterns for architectural features are at different levels of abstraction. Each pattern expresses a relation between a certain context and system of factor which occurs repeatedly in that context and a solution which allows these factors to resolve themselves. Therefore, design scenarios can be derived from empirical studies, such as to have greenery space and decorated pavement surround a coffee shop in a street corner where people can sit outside because it creates a pleasant "atmosphere". To define agents and their roles in play, we develop scenarios, suppose design objects such as pavements and lamps can look after themselves. A scenario consists of some scenes. A scene is used for describing state transitions. Scenarios include agents and actors. Every scenario involves at least one agent and one goal; they include sequences of actions, events and characteristic elements. Each scenario is described as state transitions, and states semantic is depending on cues and actions.

For example, the Lamp-Agent and Pavement-Agent: agent's action is defined by state and based on the distance and location plot, the Lamp-Agent can communicate with the Pavement-Agent or another agent. The "(Dxi,Dyj)" is referring to Pavement-Agent's Direction; the "(Hi, Hj)" is referring to Lamp-Agent's Location, the "T" is referring to Lamp-Agent's Type, the "C & M" is referring to one Lamp-Agent and another's distance. The Coordination-Agent's State and Action-Script, as shown in Figure 3.

Symbol: "***A***" is refer to Agent,
"***PA***" is refer to Pavement-Agent,
"***LA***" is refer to Lamp-Agent,
"***SA***" is refer to Sub-Agents and the
"**Ri**" is refer to applied rule.

Function: if (***PA***) *cue* (***LA***)
then (***A***) *action* (***SA***)

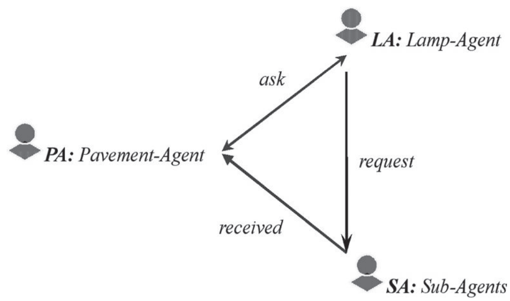


Figure 3. Behavioral function of agents.

An example of cue and action is as follows:

Scene_1- if (***PA***) *ask* (***LA***)
Scene_2- then (***A***)
request [***SA***: *Sub-Agents*]
goto
Scene_3-1 [**SA:** Sub-Agents *received* R1.]
else
Scene_3-2 [**SA:** Sub-Agents *received* R2.]
else
Scene_3-n [**SA:** Sub-Agents *received* Rn.]
else
goto next Scene

4.2 *Coordinative agents*

When a script function is carried out, it means a goal to be completed. Therefore, the action dependency of each agent of the object with one another needs to increase coordinative agents to take responsibility for the behavioral function of coordination if a good coordination is desired to be generated. Coordinative agents responsible for coordinative function of behavior, such as ask, receive, request, tell, and inform, need to be added if excellent coordination is required for action dependencies among the agents when a script function is executed; that is, to complete a goal. In the operational mechanism of coordinative agents in this research, each individual agent which is equipped with only one major function; that is, the independent variable and rule, is adopted, in order to respond to the missions among the groups. And, each of the spatial objects contained in the design pattern is endowed with features and behavioral properties to establish the dependent relationships of one-to-multiple, one-to-one, and multiple-to-multiple spatial objects. This way, once the status of a spatial object changes, other dependent objects will be automatically notified in order to make the corresponding updates. Then, the relationships among the categories of objects are converted to the agent concepts.

5 CONCLUSION AND SUBSEQUENT STUDY

Currently, this study aims to discuss the interactions between the "good place" of city streets and pedestrians. Rather than pedestrian flows, elements which easily catch pedestrians' attention are the most essential in street environments. Such elements, which clearly indicate or imply different activity behaviors that an environment is able to support and increase the time pedestrians halt and the diversities of the area, make that particular environment a good place in city life. This research has developed an interactive agent model with the "design pattern" as an orientation to express the interaction of each individual with the multiagent-based model. Agents can presume all the different individuals, such as the space element, furniture on the street, plants, and lamps, to be regarded as single individual of behavior. How these different agent individuals interact with one another is the research point in this research. This model will be able to assist experts to express the knowledge of spatial configuration in the future, and to provide students or beginners to conduct the exploration of spatial configuration. In the subsequent study, the completed pattern script will be input into the platform system for use. It can be a knowledge database of net groups for design agents to continue its growth. What's more, in order to keep the openness of the platform, the open source and the free software will be adopted in the operational system and the server.

REFERENCES

Alexander, C., Sara, I. & Murray, S. 1977. *A Pattern Language*. New York: Oxford University Press.

CDRN 2006. Canadian Design Research Network. Retrieved December 19, 2007, from http://www.cdrn.ca/

Gehl, J. 1996. Life between buildings: Using public space. Copenhagen: Arkitektens Forlag.

Ishida, T. & Nakanishi, H. 2003. Designing scenarios for social agents. In: Zhong, N., Liu, J., Yao, Y. (Eds.), *Web Intelligence*. Springer, Berlin. 59–76.

Ishida, T. 2002. Q: A scenario description language for interactive agents. *Computer* 35: 42–47.

Woodbury, R. 2007. Design patterns for parametric modeling. Retrieved December 19 from http://www.designpatterns.ca/

Wooldridge, M. 2002. *An Introduction to Multi-Agent System*. Chichester, England, John Wiley and Sons.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Relationship between personnel, structure design factors and team effectiveness in virtual teams

Tzai-Zang Lee
Kun Shan University of Technology, Tainan, Taiwan

Yu-Yuen Show
National Cheng Kung University, Tainan, Taiwan

ABSTRACT: This study analyzed the characteristics of virtual teams based on work system subsystem model, and focused on the personnel factors and their relationships with structure design factors; also, this study surveyed the impact of personnel factors on team effectiveness. The research was conducted within the staffs of the member libraries of Nationwide Document Delivery Service (NDDS) and Consortium on Core Electronic Resources in Taiwan (CONCERT). Using canonical correlation analysis, positive relationships between one personnel subsystem factor (professionalism) and two organizational subsystem factors (formalization and decentralization) were found. Furthermore, the professionalism was positively related to team effectiveness. The range of findings do support the general statement that personnel subsystem is an important construct to consider in studying virtual teams. It is not sufficient for managers to simply develop some structural characteristics of a virtual team while ignoring personnel factors. According to the findings, the library consortia in Taiwan have lots of employees with high professionalism, and these institutes might aim to lower the formalization.

Keywords: virtual team, work system subsystem model, team effectiveness, library consortia

1 INTRODUCTION

Virtual team is important by degrees in organizations, but we have not caught on it deeply all along the line, especially about personnel factors. Thus, this study concentrates on the concerning of personnel factors with structural design factors and team effectiveness in virtual teams to get a more complete understanding about it.

1.1 *Research background and motivations*

Virtual teams are becoming accustomed in organizations and critical for companies to survive. But these teams also increase the complexity of a work system. The success of virtual teams relies not merely on the introduction and adoption of Information and Communication Technologies (ICTs), but also on the underlying social and organizational aspects. However, little is known about the management of virtual teams and the human resources within these teams. Dani, Burns, Backhouse and Kochhar (2006) also considered that we need extensive researches to understand the design characteristics of successful virtual teams.

SocioTechnical Systems (STSs) theory is the foundation of macroergonomics (Hendrick, 2000, 2007; Kleiner, 2002, 2004; Kleiner & Shewchuk, 2001; Wilson, 2000). The principle of STSs theory is that any work system requires technological and personnel factors links with each other, however, using virtual teams tend to be driven by technical considerations (Cramton & Webber, 2005). Complex sociotechnical systems, like virtual teams, require a true macroergonomics approach to design work system (Lee, 2005). Work system subsystem model which grounds on sociotechnical system involves technological subsystem, personnel subsystem, external environment, and a further defined organizational subsystem. STSs theory conveys that understanding each subsystem is important. Immediately, work system subsystem model indicates that the more important thing is we should *design* the system according to these subsystems. Hendrick (2007) described that the characteristics of the personnel subsystem are more important than those of the technological subsystem. Further, the subject of this research is about virtual team, which should be paid more attention to related non-technical issues as discussed earlier, so this study aims at the relationship between personnel subsystem and the organizational subsystem. Hendrick (2007) also argued that all the factors among personnel subsystem impact the effectiveness of a work system, thus, this study also wants to know how these factors affect the effectiveness of a work system.

Library consortia are a comparatively new phenomenon in Taiwan, but in the times of rapid societal change, technological change, and economic constraints, participation in a library consortium is essential for an individual library to provide a better and greater variety of information services (Ching, Poon, & Huang, 2003). There are lots of contributions of library consortia such as stimulating the need for electronic communication, such as e-mail, e-mail discussion lists, and Web sites, the facilitation of the use of common library systems, resource sharing, e.g., through the Ariel system, and improving access especially to electronic resources (Thomas & Fourie, 2006). Judging from the above, regarding library consortia as virtual teams is rational.

1.2 *Team effectiveness*

During the past 15 to 20 years, organizations have increasingly used work teams, that is, groups of individuals with mutual accountability working interdependently to solve problems or work (Kirkman & Mathieu, 2005). Organizations have experimented with team-based structure to become more flexible (Jackson & Ruderman, 1995). Teams have become hot topics. Organizations have come to dependent on teams to improve quality, productivity, problem solving, customer service and the experience at work for their members (Lira, Ripoll, Peiró, & González, 2007). The increasing use of teams in organizations (Mohrman, Cohen, & Mohrman, 1995) has motivated researchers to survey the factors determining team effectiveness.

The term effectiveness refers to how well a team accomplishes its purpose or mission (Lira et al., 2007). Commonly, team effectiveness is analyzed in terms of work outcomes (González, Burke, Santuzzi, & Bradley, 2003). However, effectiveness also includes other effects helping perpetuate work outcomes over time. For example, the extent to which the team experience enhance the capability of the members to work together in the future, and the extent to which team members' experience on the team is satisfying (Lira et al., 2007). Lurey and Raisinghani (2001) explored the effectiveness within virtual teams, and in an effort to determine the factors that contribute to or inhibit the success of a virtual team. Beyond the traditional strategies used to enhance a team's effectiveness, they try to determine which practices led to the success of the participating virtual teams and two separate measures of team effectiveness are established in this survey. The first scale relates to the teams' abilities to perform their work tasks and the second focus on the team members' levels of satisfaction while working with their virtual teams. Doolen, Hacker and Van Aken (2003) explored the relationships between nine organizational context variables, team processes and three measures of team effectiveness, and highlighted the need to include multiple types of effectiveness measures. Besides, this study notices that teams in the workplace are often formed according to the technical knowledge of its members. Little is known, however, regarding the nontechnical factors that determine team effectiveness (Varvel et al., 2004).

1.3 *Virtual team*

Virtual team, a new business model, is empowered by latest advancement in ICTs and becoming a key ingredient for organizations to success (Rezgui, 2007). Virtual teams are capable of combing positive aspects of conventional teamwork with ICTs, accomplish works with much greater independence from time or space constraints, and make expert knowledge from remote sites to be integrated (Hertel, Konradt, & Orlikowski, 2004). In such teams, members have to work across differences in assumptions, motivations, knowledge bases, and working styles (Shapiro, Furst, Spreitzer, & Von Glinow, 2002), thus, increasing the complexity of a organization.

Hertel et al. (2005) indicated that as a minimal consensus, virtual teams consist of (a) two or more persons who (b) collaborate to achieve common goals, while (c) at least one of the team members works at a different location, organization, or at a different time so that (d) communication and coordination is mainly depending on electronic communication media.

Researchers should investigate this new type of organization from all aspects to better make use of it. Rezgui (2007) pointed out that we must continue to explore how to make virtual teams work effectively. There are some researches focus on the virtual team and its effectiveness (e.g., Lurey & Raisinghani, 2001; Rezgui, 2007). To sum up, the term 'virtual team' is used to cover a wide range of activities and forms of technology-supported working, inevitable, lots of researches focus on the technology-related issues (e.g. Anderson, McEwan, Bal, & Carletta, 2007; Huang, Wei, Watson, & Tan, 2002; Wiesenfeld, Raghuram, & Garud, 1999; Zammuto, Griffith, Majchrzak, Dougherty, & Faraj, 2007), and pay less intention on the staffs using these ICTs, the personnel factors.

For reasons of simplicity, and this study follows the definition of Hertel et al. (2005) about virtual team. Virtual teams consist of two or more persons who collaborate to achieve common goals, while at least one of the team members works at a different location, organization, or at a different time so that communication and coordination is mainly depending on electronic communication media. That is, this study uses "virtual team" throughout the text as a label of teams with high degrees of virtuality.

1.4 *Work system subsystem model*

Work system subsystem model is building on the sociotechnical theoretical foundation (Kleiner & Drury, 1999). The term "sociotechnical system" was first coined by Emery and Trist to better convey the nature of complex human–machine systems, and STSs theory is the foundation of macroergonomics. There are various models that defined elements of sociotechnical system at the present day, these models present different ways of going to sociotechnical systems, either vertically, functionally, or by domain, and work system subsystem model is one of them (Carayon, 2006). According to work system subsystem model, the design of a work system involves consideration of the key elements of three major sociotechnical system components. Each of these three major sociotechnical system components has been studied in relation to its effect on the fourth component, organizational subsystem which focuses on structure designs (Hendrick, 2007). One of the most consistent findings in related researches is that these basic sociotechnical system elements are mutually interdependent (Hendrick & Kleiner, 2002). But in this study, we focus on the relationship between personnel subsystem and organizational subsystem, thus we only introduce these two subsystems in the following:

- Personnel subsystem

The most important factor in personnel subsystem is the degree of professionalism of the work force, refers to the education and training requirements of a given job (Hendrick, 2007). Increasing cultural diversity in the workforce also poses one of the most challenging human resource and organizational issues of this times (Richard, 2000).

Hall (1968) identified five attitudinal attributes that define professionalism as follows, using the professional organization as a major referent, belief in public service, belief in self-regulation, sense of calling to the field, and autonomy. Professional organization can reinforce the values, beliefs, and identity of the profession, and can be both formal and informal. Once the practitioner acquires such consciousness, he/she is thought to be strongly influenced by the standards of his/her profession. Belief in public service means that a practitioner believes his/her profession is both indispensible and beneficial to society. Developing this kind of belief is slow, since the general public has difficulty believing in the indispensibility of services performed by some occupations. A belief in self-regulation represents a profession's endorsement of the notion of colleague control. Individuals with belief in self-regulation argue that given the state of specialized knowledge required in their occupation, only colleagues, not "outsiders," are qualified to judge their work. Sense of calling to the field conveys that a professional is totally devoted to his/her work, and has the belief that the development and exercise of expertise is worthy of the devotion of a lifetime and carries its own reward. In one word, he/she performs his/her services primarily for the psychic satisfactions and secondarily for the monetary compensation. Autonomy, as a professional attitude, consists of the practitioner's desire to be free to make decisions about his/her work. These decisions should be made without the threat of external pressures, such as pressures at the organizational level.

Cultural diversity means the representation, in one social system, of people with different group affiliations of cultural significance (Richard, 2000). Watson, BarNir and Pavur (2005) defined a cultural model by individuals' gender, age, ethnicity, tendency to help others, and tendency to work as an individual to extend typical definitions of cultural diversity and investigated the effects on desired team process outcomes.

- Organizational subsystem

Organizational subsystem is focused upon the design of three core dimensions: complexity, formalization, and centralization. Understanding of these sociotechnical characteristics of a given work system will guide us to optimize the work system's organizational structure (Hendrick & Kleiner, 2002).

Complexity refers to the degree of differentiation and integration exiting in an organization. Differentiation is focused on the segmentation of the organization. Three major kinds of differentiation are found in an organization's structure: vertical, horizontal, and spatial dispersion. Vertical differentiation is operationally defined as the number of hierarchical levels separating the chief executive's position from the jobs directly involved with the systems output. Horizontal differentiation refers to the degree of departmentalization and job specialization that are designed into the organization. Division of labor creates groups of specialists. The optimal degree of specialization to design into the system depends upon various sociotechnical system factors. Spatial dispersion may be defined operationally as the degree to which an organization's facilities and personnel are dispersed geographically from the main headquarters. Increasing any one of the dimensions described above increases a work system's complexity. Integration is focused on linking the segments together with coordinating mechanisms. Some of the more common integrating mechanisms are formal rules and procedures, liaison positions, committees, task teams, system integration offices, and information and decision support systems.

Formalization is defined in terms of the degree of standardization. Organizations with low formalization allow employees more freedom to exercise discretion. Employee behavior is relatively unprogrammed, and workers are able to make greater use of their mental capacities.

Centralization is concerned with decision-making and the extent to which authority is concentrated within a few individuals. In general, when a comprehensive perspective is required, such as in strategic decision making, or when operating in a highly stable and predictive environment, centralization is desirable.

Hacker and Kleiner (1996) highlighted that the current understanding of team effectiveness will not greatly enhance our understanding of virtual teams. They also commented that researchers can better study the virtual teams based on STSs theory.

2 RESEARCH METHODOLOGY

First, this study hypothesizes that personnel subsystem and organizational subsystem would have correlative relationship. Secondly, this study also hypothesizes that personnel subsystem would positively influence the team effectiveness.

Given the main objectives of this research, the instrument was developed to measure the characteristics in the personnel subsystem, organizational subsystem, team effectiveness, and team members' demographics. The instruments were measured on 5-point Likert scales except the demographical data.

- Personnel subsystem

There were two characteristics in personnel subsystem in this study, professionalism and cultural diversity. Professionalism was measured with a 25-item instrument (Snizek, 1972). Cultural diversity was measured by four factors: gender, age, tendency to work as an individual, and tendency to help others (Watson et al., 2005). Gender and age data were gathered on the demographic form. Type-A personality estimated the tendency to work as an individual and was measured by 12 items (Wang, 1994). Communal orientation scale estimated the tendency to help others and was measured by 14 items (Clark, Powell, Ouellette, & Milberg, 1987).

- Organizational subsystem

The organizational subsystem had three dimensions in this study: complexity, formalization, and decentralization. Complexity, that was vertical differentiation in this study, was measured with a 3-item instrument (Koufteros, Nahm, Cheng, & Lai, 2007). Formalization was measured with a 4-item instrument (Koufteros et al., 2007). Decentralization was measured with a 6-item instrument (Olson, Slater, & Hult, 2005).

- Team effectiveness

Team effectiveness was measured by 9 items which were developed by Lurey and Raisinghani (2001).

- Sample selection

Kirkman and Mathieu (2005) encouraged researchers to study virtual teams with differing degrees of virtuality and explore the influences of such differences. Therefore, this study chose some virtual teams in Taiwan with similar degrees of virtuality to verify the conceptual framework, and the staffs of member libraries of Consortium on Core Electronic Resources in Taiwan (CONCERT) and Nationwide Document Delivery Service (NDDS) were the subjects.

This study needed at least 150 efficient responses with a margin of error of 8% at a 95% confidence, thus we used convenience sampling and sent out 200 questionnaires with response rate of 75%.

3 RESULTS

The analysis results, including pilot test and questionnaire retrieval, descriptive statistics analysis, reliability analysis, structural equation modeling, canonical correlation analysis, and multiple regression analysis were done. For the propose of limiting the length of this article. Results of canonical correlation analysis, multiple repression analysis, are presented.

3.1 *Correlation analysis*

A canonical correlation analysis was conducted, and two canonical functions were produced. These functions captured all the correlations between the two groups of variables. The Wilks' lamd revealed that only the significance of the first canonical function exceeds the critical value of 0.05, and the canonical correlation of the first canonical function was 0.401. The criteria for determining the significance of canonical loadings are the same as with factor loadings in factor analysis, that is, variables with canonical loadings of 0.5 or greater should be included in the interpretation (Hair, Tatham, Anderson, & Black, 1998). The function indicated that professionalism was significantly and positively correlated with formalization and decentralization. The canonical correlation results also suggested that cultural diversity and complexity did not significantly contribute to the relationship

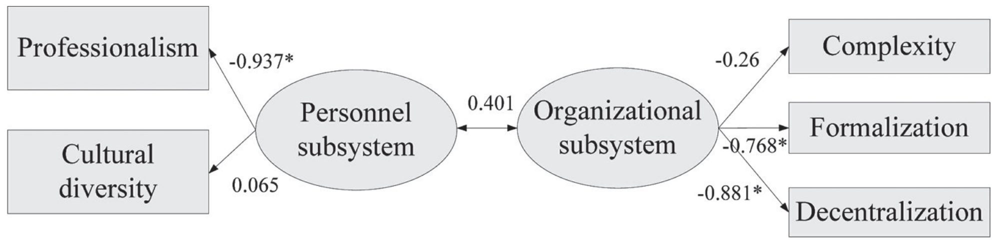


Figure 1. Canonical function 1.

Table 1. Multivariate test of significance for canonical correlation analysis.

Canonical function	Canonical correlation	Wilks' lamda	Chi-square	df	Sig.
1	0.401	0.835	23.989	6	0.001
2	0.068	0.995	0.617	2	0.735

between personnel subsystem and organizational subsystem.

3.2 *Multiple regression analysis*

To test hypothesis 2, multiple regression analysis was conducted. The multiple regression analysis showed that team effectiveness was positive related to professionalism, and professionalism explained approximately 20% of the variance in team effectiveness. As a practical matter, for typical data found in the social sciences, values of R^2 as low as 0.25 are often considered useful; for data in the physical and life sciences, R^2 values of 0.6 or greater are often found; in business applications, R^2 values vary greatly, depending on the unique characteristics of each application (Anderson, Sweeney, & Williams, 2002). In general, correlation coefficient ranging from 0.4 to 0.6, that is, R^2 ranging from 0.16 to 0.36, is considered as some relationship, thus R^2 value of 0.216 and adjusted R^2 value of 0.205 in this study were considered acceptable.

4 DISCUSSION

This research indicates that certain personnel subsystem factors do positively affect certain organizational subsystem factors. It also supports that certain personnel subsystem factors do enhance the team effectiveness. Before explaining the analysis results, there are some points should be notice.

- Personnel subsystem

Although there are five attributes defining professionalism, only three of them are remained after reliability test in the pilot test. That is, belief in public service, belief in self-regulation, and autonomy. Extant research indicates professions share a number of attributes; professionalism is the attitudinal component, or behavioral dimension, that conditions how individuals think about, believe in and behave toward their occupation or profession. On the other hand, structural attributes lead to the creation of the professional and include entrance requirements, formal education, mandatory skill development and licensing or certification (Hampton & Hampton, 2004). According to the descriptive statistics data, more than 90% respondents' education levels are above college/university, which also show professional in structural attributes.

This study adapts the cultural model of Watson et al. (2005). There are two important attributes in this model, tendency to work as individual and tendency to help others, which are considered suitable for this study. That is to say, tendency to work as individual and tendency to help others tend to relate to organization structural design factors, and affect team effectiveness. But after the analyzing, cultural diversity didn't show any significant relationship with other factors. It might due to the proposed cultural model itself. This cultural model might not have enough explanatory power or not proper for this study.

- Organizational subsystem

Complexity is a complex concept, but only vertical differentiation is concluded in this study. This may help account for the results that complexity didn't show any significant relationship with other factors.

Formalization implies an internal environment with rules and procedures that encourage creative, autonomous work and learning; formalization does not mean an environment absent of written policies and procedures, but rather policies and procedures that enable the firm to capture, organize, and share knowledge (Nahm, Vonderembse, & Koufteros, 2003). Although Robbins (1983) noted that formalization can take place either on the job or through the process of professionalization that is there is a trade-off between formalization and professionalization of the job. The analysis showed that professionalism and formalization have significant positive relationship. This may due to the sample organizations are school libraries and regional

Table 2. Regression analysis results.

Dependent variable	Significance of regression model (p)	R^2	Adjusted R^2	Independent variable	Standard regression coefficient	Significance of coefficient (p)
Team effectiveness	<0.001	0.216	0.205	Professionalism	0.448	<0.001
				Cultural diversity	0.050	0.532

libraries. These educational institutes and official associations tend to have high formalization. Professionals may do more quality improvement or experiment their jobs, at the mean time, they have to follow the rules or procedures. Thus there are reasons of thinking that professionalism and formalization have positive relationship in this study.

According to Hendrick and Kleiner (2002), in highly decentralized work system, decisions are authorized to the lowest level having the necessary expertise, thus highly decentralized work systems require lower-level employees to have a relatively higher level of professionalism. Accordingly, it is for this reason that professionalism and decentralization have significant relationship.

- Team effectiveness

The analysis shows that professionalism has positive influence on team effectiveness. And it makes sense to see a trained workforce with belief and autonomy having better effectiveness.

5 CONCLUSIONS AND SUGGESTIONS

The increasing complexity of sociotechnical systems poses unique challenges; further integrating the different dimensions and elements of sociotechnical systems is necessary to anticipate the implications of working across organizational, geographical, cultural and temporal boundaries (Carayon, 2006). Rezgui (2007) also indicated that the migration path to successful virtual team working is grounded in human and cultural elements. The purpose of this study is to deepen insight into the relationship between personnel subsystem and organizational subsystem, personnel subsystem and team effectiveness. In this study, personnel subsystem includes professionalism and cultural diversity, and organizational subsystem includes three structure design factors, complexity, formalization, and decentralization. The results demonstrate that professionalism and formalization, professionalism and decentralization have positive relationship; also, professionalism has positive influence on team effectiveness. The range of findings do support the general statement that personnel subsystem is an important construct to consider in studying virtual teams. It is not sufficient for managers to simply develop some structural characteristics of a virtual team while ignoring personnel factors. According to the findings, the library consortia in Taiwan have lots of employees with high professionalism, and these institutes might aim to lower the formalization.

Although Lurey and Raisinghani (2001) indicated that many of the issues that affect virtual teams are similar in nature to those that affect co-located teams, Zigures (2003) emphasized that virtual team members cannot rely on simple transferring their behavior in traditional teams and expect to be successful in virtual environments. Given current limitations of virtual team research, the paper contributes to existing knowledge by (a) inquiring personnel subsystem to designing organizational subsystem, (b) specifying personnel subsystem to support team working effectively; and (c) researching into factors that facilitate virtual team adoption and use in the library consortia in Taiwan.

There are a number of issues that have come to light as a result of this research. Additional studies of various types of virtual teams will help confirm the distinctness of personnel subsystem as well as the proposed structure.

In the second place, the different level of virtuality may influence on the relationship between personnel subsystem, organizational subsystem, and team effectiveness. Future studies including some explicit measures of team virtuality may give a deeper understanding of these interrelationships and may make comparisons between different types of teams.

Furthermore, this study focused on professionalism and cultural diversity in personnel subsystem. The relationships between these factors and organizational subsystem, and team effectiveness were explored. It is clear that additional studies with further fine-tuning items used to assess the various aspects of personnel subsystem would provide additional insight into the set of relationships.

REFERENCE

References will be provided when requested.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Exploring factors of team workload for the advanced main control room

Tsung-Ling Hsieh, Chiuhsiang Joe Lin & Ping-Jung Tsai
Department of Industrial Engineering, Chung-Yuan Christian University, Taiwan, Republic of China

Chih-Wei Yang & Tzu-Chung Yenn
Institute of Nuclear Energy Research, Taipei, Japan

ABSTRACT: An operating team of advanced main control room is comprised of three operators. Some study pointed out that optimizing the allocation of workload to operators could reduce human errors, improve systems safety, and increase productivity and operator satisfaction. Oppositely, both excessive and low mental workloads could degrade operators' performance. Therefore, to evaluate the workload of team operators in the advanced MCR would be one valuable issues of human factor engineering research. This study first describes not only the operating environment but also the tasks and interaction of the operators in the advanced MCR. Then, the literature about teamwork and workload would be reviewed and explored. Finally, this study concludes the potential factors of team workload which including "coordination", "communication", "supporting and leadership", and "time sharing" according to the literature review, and suggests the method to measure and calculate the team workload.

Keywords: advanced main control room, teamwork, workload

1 INTRODUCTION

There is a developing consciousness among the international nuclear power community regarding the importance of considering human factors. This consciousness has been heightened by a realization that incidents are seldom the result of merely technical failures. Rather, incidents arise from the interaction between the human and the technical elements of a system (Chuang & Chou, 2006). Therefore, in the last decade there has been a tremendous wave of interest in the relationship between computer system and human capability such as the computerization, digitization, alarm system of interface design (e.g. Carvalho et al., 2008; Hwang et al., 2008; Jou et al., 2008), and workload, situation awareness, performance of personnel (Hwang et al., 2008; Kim et al., 2006; Mumaw et al., 2000; Park et al., 2004) in the Advanced Nuclear Power Plant (ANPP).

One of the most widely studied areas of human capability has been concerned with the performance effects of operator workload (Bowers, Braun & Morgan, 1997). For example, it has been reported that more than 500 workload papers have been written since the early 1960s (Hancock et al., 1988). The mental workload has been generally defined as the amount of resource difference between task demands and capacity provision by an individual (O'Donnell & Eggemeier, 1986; Sanders & McCormick, 1993; Veltman & Gaillard, 1996; Xie & Salvendy, 2000). It is one of the major concerns within complex systems (O'Donnell & Eggemeier, 1986; Xie & Salvendy, 2000), such as the Main Control Room (MCR) of ANPP (O'Hara & Hall, 1992; Sebok, 2000; Ha et al., 2007), because of its effect on human error and performance.

Some study pointed out that optimizing the allocation of workload to operators could reduce human errors, improve systems safety, and increase productivity and operator satisfaction (Moray, 1988; Sebok, 2000). Oppositely, both excessive and low mental workloads could degrade operators' performance (Nachreiner et al., 2006; Xie & Salvendy, 2000; Mo, Lee & Seong, 2007). Therefore, demand for more precisely evaluated and predictive workload of operators is valued.

Various mental workload measurements have been proposed, and these measurements could be divided into three categories: performance measures, physiological measures and subjective measures (Rubio et al., 2004; Hwang et al., 2008). Their pros and cons have been extensive discussions in many studies (Veltman & Gaillard, 1996; Charlton, 2002; Farmer & Brownson, 2003). The subjective measures can be defined as subject's direct estimate or comparative judgment of the mental or cognitive workload experienced at a given moment (Ameersing & Ravindra, 2001). They do not disrupt primary-task performance, and derive more easily. Moreover, their costs relate to the uncertainty with which an operator's verbal statement diagnostically reflects the investment of or demand for processing resources and is not influenced by other biases (Wickens & Hollands, 2002).

The most famous measurement tool is the NASA task load index (NASA-TLX), developed by Hart and Staveland in 1988, is a conventional subjective measures for assessing workload, because it is applicable when detailed and diagnostic data are needed (Hill et al., 1992) and it can be used to support the general prediction model for experienced workload (Nygren, 1991).

Increased technology of computer hardware and software has contributed to the complexity of many tasks performed in the workplace, and tasks are usually accomplished by teamwork because they are too difficult for a single operator, especially during emergency operating procedures (Urban et al., 1996; Sebok; 2000; Mathieu et al., 2000; Huang & Hwang, 2009), so teams have become an increasingly important unit of performance in the workplace (Cannon-Bowers, Oser & Flanagan, 1992). Hence, the focus of human machine interaction has shift from an emphasis on the evaluation of individual performance to attention to evaluation of team performance.

A team can be defined as a distinguishable set of two or more individuals who interact dynamically, interdependently, and adaptively toward a common, specified, shared, and valued goal/objective/mission. The individual team members are each assigned specific roles or functions to perform (Morgen et al., 1986; Salas et al., 1992). Team workload can be characterized as the relationship between the finite performance capacities of a team and the demands placed on the team by its performance environment (Norman & Bobrow, 1975).

In some respects, team performance environments can be viewed as being similar to that of a dual-task paradigm that was dealt by individual (Bowers, Braun & Morgan, 1997). As deduced from the above, the characteristics of teamwork are not entirely different from personal work. Hence, some studies evaluated the workload of the team was assessed through the NASA-TLX, with individual scores combined into a team workload score (Thornton et al., 1992; Bowers et al., 1992; Urban et al., 1995; Sebok, 2000; MacMillan, Entin & Serfaty, 2004). However, it should be noted, some studies revealed the there was no significant effect of team workload measured with NASA-TLX on team performance. In other words, the sensitivity of assessing team workload with NASA-TLX seems to be verified or reconfirmed. This may suggest some characteristics of teamwork have been not considered by NASA-TLX.

In order to compensate for this defect, this study attempted to review the literature about teamwork to explore the potential factors of team workload, and then to suggest the method to measure and calculate the team workload.

2 TEAMWORK IN THE MCR OF ANPPS

The MCR of ANPP can use graphical interfaces of human-machine systems that display the operating rules of the reactor system and the relationship between the sub-systems and offer operators with direct control (Vicente, Mumaw & Roth, 1997). The human-machine systems of the nuclear power plant are numerous and the interdependence among these systems is complicated (Jones et al., 2007).

Nevertheless, the information presented by advanced interfaces does not often support team performance (O'Hara & Hall, 1992; Sebok, 2000). When designing for a team is considered, a number of studies have noted that teamwork assistance systems can provide efficient communication strategies, reduce human error, support higher team performance and situation awareness, and organize the operator crew (Plott et al., 1995; Hallbert et al., 1997; Furuta et al., 1999; Huang & Hwang, 2009).

According to NUREG/CR-6838 (2004) and NUREG-0711 (2004), operating teams of ANPP MCRs are comprised of one Reactor Operator (RO), one Assistant Reactor Operator (ARO), and one Shift Supervisor (SS). The safety-related system is taken care of by the RO, while the ARO is responsible for operations aside from the safety-related system. The SS deals with conditions of the nuclear power plant, gives commands, and coordinates operations in the entire control room. Furthermore, the operators are required to monitor system information from nearly 1000 monitors on 45 Video Display Units (VDUs) in the operation environment of the MCR, and the operators can touch the screen to select the ones they want to see, use the information on the Wide Display Panel (WDP), and thus determine system conditions and the necessary operational steps (Chuang & Chou, 2006).

Under normal operational conditions, operators only need to follow general operation procedures to maintain normal operations of the MCR. However, when incidents and accidents occur, not only would the system make alarm sounds, but also the wide display panel shows warning information.

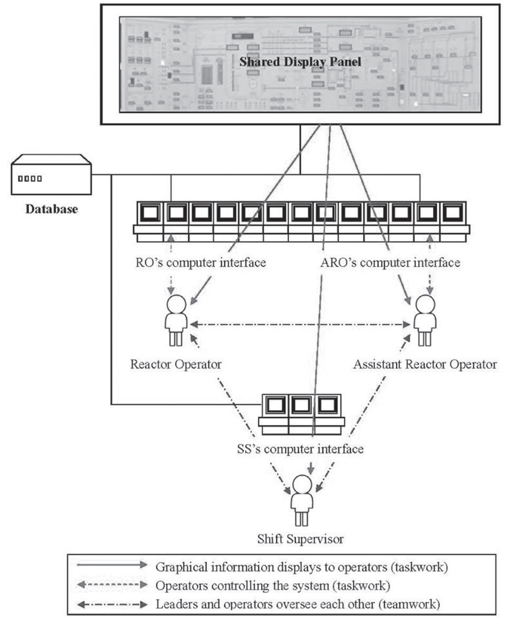


Figure 1. The framework of the teamwork system (adapted from Huang & Hwang, 2009).

In such instances, the RO and ARO must first verify the alarm information, and closely monitor the relevant VDUs for information about the reactor system. They must then report abnormal conditions to the SS. Following this, the SS watches information on the VDUs to select the Emergency Operating Procedures (EOPs) to be consulted, and finally follows the steps in the EOPs to guide the RO and ARO to activate backup or rescue systems to stabilize the reactor system and avoid worsening of the situation. Then the SS contacts departments outside the control room for assistance in restoring normal reactor functions (Lin et al., 2009). To sum up, the interaction of operators in advanced MCR is shown in Figure 1.

3 THE FACTORS OF TEAM WORKLOAD

Team performance requires team members to engage concurrently in two broad categories of activities; namely, taskwork and teamwork activities (Glickman et al., 1987; Morgan, Salas & Glickman, 1994). Taskwork refers to a team's interaction with tasks, tools, machines, and systems, such as the RO and ARO control the reactor system in the MCR. Teamwork refers to the interpersonal interactions among individuals that are necessary for exchanging information, developing and maintaining communication patterns, coordinating actions, maintaining social order, such as the RO reply the information of shared display panel to the SS. Effective teamwork has been further defined in terms of behaviors that contribute to "error identification and resolution, coordinated information exchange, and intermember reinforcement" (Oser et al., 1989).

Team members must time-share the requirements for interaction and coordination with the demands imposed by performing their specifically assigned tasks (Bowers, Braun & Morgan, 1997). Teamwork skills are needed for effective interactions among team members (Johnston, Smith-Jentsch & Cannon-Bowers, 1997) and are largely dependent of the task to be performed. Such skills assist team members in their development of a shared mental model and help develop team leadership skills (Glickman et al., 1987; Cannon-Bowers et al., 1995; Ford & Schmidt, 2000). According to literature review, the factors may interact with one another and potentially affect team performance were as follows.

3.1 *Coordination*

Coordination is one of the broad categories of team behaviors in the absence of widely accepted team taxonomies (Bowers, Braun & Morgan, 1997). Coordination that is the act of performing organized operations to fulfill common criteria (Suchman, 1987; Hoc, 1988) lies at the heart of effective team performance (Kleinman & Serfaty, 1989; Orasanu & Salas, 1993). Some studies highlighted there was a significant effect of coordination on the team performance (Morgan et al., 1986; Glickman et al., 1987; Dickinson & McIntyre, 1997; Chiocchio, 2007). The coordinative demands of a team task require greater use of team members' cognitive resources (Kontogiannis & Kossiavelou, 1999; Ford & Schmidt, 2000; Kidd, 1961). The nature of coordinated behavior is the distinction between implicit and explicit coordination (Entin & Serfaty, 1999; Stout et al., 1999). Explicit coordination requires that team members communicate to articulate their plans, actions, and responsibilities, whereas implicit coordination describes the ability of team members to act in concert without the need for overt communication. For implicit coordination to be successful, team members must have a shared understanding of the situation and an accurate understanding of each other's tasks and responsibilities.

3.2 *Communication*

Communication and team operation are inseparable (Bowers, Braun & Morgan, 1997). The focus of the communication dimension is on how that information is exchanged. Communication between team members is critical for the successful completion of team tasks (Scholtes, 1988; Pinto & Pinto, 1991; Weil, 2004). Communication frequency can be seen as an indicator of the information processing activities of team members (Roth, 1995; Reich & Benbasat, 2000). High-communication frequency can result in more information being exchanged between team members with the greater exchange requiring more information processing. At some point, such exchanges may overload the capabilities of team members and inhibit their performance (Goodman, Ravlin & Argote, 1986). Similarly, low-communication frequencies may not supply enough information to team members and may not facilitate the innovative combination of information and expertise, required for high performance (Denison, Hart & Kahn, 1996). Thus, communication "efficiency" seems to be more meaningful to discuss. The efficiency of communication has been shown to be positively related to team performance (MacMillan, Entin & Serfaty, 2004; Entin & Serfaty, 1999; Serfaty, Entin & Volpe, 1993). To perform effectively, all team members must understand how their actions affect and are affected by each of the other team members (Ford & Schmidt, 2000).

3.3 *Leadership and supporting*

Leadership is defined as providing needed guidance to other team members, helping team members focus their activities appropriately, anticipating tasks that should be performed, and providing instruction to other team members to enable them to perform or complete their tasks (Smith-Jentsch, 1995; Johnston et al., 1997). Some studies indicated that Leadership is crucial to effective and safe team performance (Mearns et al., 2001; O'Dea & Flin, 2001), and the qualities of leadership influence the performance of all teams, whether they are exhibited by formally appointed leaders, or by non-appointed emerging leaders (Klimoski & Jones 1995). Another similar helping behavior between team members is supporting that could be defined as monitoring the activities of other team members, taking action to correct errors, giving and receiving feedback in a non defensive manner, and providing and seeking assistance or backup when needed (Smith-Jentsch, 1995; Johnston et al., 1997; Weil et al., 2004). In brief, both leadership and supporting would involve assisting the performance of other team members, so this study would integrate them to one aspect.

3.4 *Time sharing*

As suggested by the findings of Gopher and Braune (1984), the workload imposed by combinations of teamwork and taskwork requirements should also be expected to exceed the workload associated with teamwork or taskwork alone. In turn, this could lead to diminished performance in complex team performance situations, an expectation consistent with Steiner's (1966, 1972) concept of "process loss" and Navon and Gopher's (1979) notion of "concurrence cost." This latter point is particularly important because teams are almost always required to engage in teamwork and taskwork concurrently. In most team performance situations, team members must time-share the requirements for interaction and coordination (i.e., teamwork) with the demands imposed by performing their specifically assigned task (taskwork). Therefore, it is reasonable to conclude that team workload consists of workload demands associated with teamwork, taskwork, and the requirements to time-share task performance with team interactions (Bowers, Braun & Morgan, 1997).

4 CONCLUSIONS

The purpose of the teamwork is enhancement of the performance of the operators who deal with the complex tasks in the digital environment such as the MCR of ANPP. However, the measurement of the team performance criteria still seems to need to be developed. One of these reasons is due to NASA-TLX would be unsuitable to evaluate the team workload. Thus, this study reviewed the literature about teamwork to explore the potential factors of team workload. The potential factors includes "coordination", "communication", "leadership and supporting", and "time sharing" that be evolved from theoretical base outlined in the published literature concerned with the characteristics of teamwork.

This study suggests developing a preliminary questionnaire of those factors containing some items to measure the teamwork demand. Those items are classified according to "coordination", "communication", "leadership and supporting", and "time sharing" with effects on team performance and rated on a scale of 0–10. Furthermore, an experiment design might be suggested to conduct to collect the data for validating the suitability of the questionnaire. Then, the validity and reliability could be examined by Confirmatory Factor Analysis and Cronbach's α coefficient, respectively. Finally, the scores of the questionnaire could be compared with the NASA-TLX to examine the sensitivity.

REFERENCES

Ameersing, L. & Ravindra, S. 2001. Simplified subjective workload assessment technique. *Ergonomics* 44(3): 229–243.

Bagozzi, R.P. & Yi, Y. 1988. On the evaluation for structural equation models. *Journal of the Academy of Marketing Science*, 16(1): 74–94.

Bowers, C.A., Braun, C.C. & Morgan, B.B. 1997. Team workload: Its meaning and measurement. In: Brannick, MT, Salas, E, Prince, C, editors. *Team performance assessment and measurement: theory, methods, and applications*. Mahwah, NJ: Erlbaum, 85–108.

Bowers, C.A., Urban, J.M. & Morgan, B.B., Jr. 1992. *The study of crew coordination and performance in hierarchical team decision making* (Team Performance Laboratory Tech. Rep. No. 92–1). Orlando, FL: University of Central Florida.

Cannon-Bowers, J.A., Oser, R. & Flanagan, D. 1992. Work teams in industry: A selected review and proposed framework. In R.W. Swezy & E. Salas (Eds.), *Teams: Their training and performance*. Norwood, NJ: Ablex.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E. & Volpe, C.E. 1995. Defining team competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). San Francisco: Jossey-Bass.

Carvalho, P.V.R., Dos Santos, I.L., Gomes, J.O., Borges, M.R.S. & Guerlain, S. 2008. Human factors approach for evaluation and redesign of human–system interfaces of a nuclear power plant simulator. *Displays* 29: 273–284.

Charlton, S.G. 2002. Measurement of cognitive states in test and evaluation. In: Charlton, S.G., O'Brien, T.G. (Eds.), *Handbook of Human Factors Testing and Evaluation*. Lawrence Erlbraum Associates, Mahwah, NJ, 97–126.

Chiocchio, F. 2007. Project team performance: A study of electronic task and coordination communication. *Project Management Journal* 38: 97–109.

Chuang, C.F. & Chou, H.P. 2006. Investigation on the design of human-system interface for advanced nuclear plant control room, *5th International Topical Meeting on Nuclear Plant Instrumentation Control and Human Machine Interface Technology (NPIC & HMIT 2006)*, Nov. 12–16, Albuquerque, NM, USA.

Denison, D.R., Hart, S.L. & Kahn, J.A. 1996. From chimneys to cross-functional teams: Developing and validating a diagnostic model. *The Academy of Management Journal* 39: 1005–1026.

DeVellis, R.F. 1991. *Scale development: Theory and application*. Newbury Park, CA: Sage.

Dickinson, T.L. & McIntyre, R.M. 1997. A conceptual framework for teamwork measurement. In M.T. Brannick, E. Salas & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods and applications* (pp. 19–45). Mahwah, NJ: Erlbaum.

Entin, E.E. & Serfaty, D. 1999. Adaptive team coordination. *Human Factors* 41: 312–325.

Farmer, E. & Brownson, A. 2003. *Review of workload measurement, analysis and interpretation methods* (No. CARE-integra-TRS-130-02-WP2).

Ford, J.K. & Schmidt, A.M. 2000. Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials* 75: 195–215.

Fornell, C. & Larcker, D.F. 1981. Evaluating structural equation models with unobservable and measurement errors. *Journal of Marketing Research* 18(1): 39–50.

Furuta, K., Shimada, T. & Kondo, S. 1999. Behavioral simulation of a nuclear power plant operator crew for human-machine system design. *Nuclear Engineering and Design* 188: 97–109.

Gay, L.R. 1992. *Educational research: Competencies for analysis and application (4th Ed.)*. New York: Macmillan.

Glickman, A.S., Zimmer, S., Montero, R.C., Guerette, P.J., Campbell, W.J., Morgan, B.B. & Salas, E. 1987. *The evolution of teamwork skills: An empirical assessment with implications for training* (NTSC Tech. Rep. No. TR-87-016). Orlando, FL: Naval Training Systems Center.

Goodman, P.S., Ravlin, E.C. & Argote, L. 1986. Current thinking about groups: setting the stage for new ideas. In: Goodman, P.S. (Ed.), *Designing Effective Work Groups*. Jossey-Bass, San Francisco, CA, 1–33.

Gopher, D. & Braune, R. 1984. On the psychophysics of workload: Why bother with subjective measures? *Human Factors* 26: 519–532.

Ha, J.S., Seong, P.H., Lee, M.S. & Hong, J.H. 2007. Development of human performance measures for human factors validation in the advanced MCR of APR-1400. *IEEE Transactions on Nuclear Science* 54(6): 2687–2700.

Hallbert, B.P., Sebok, A., Morisseau, D.S. & Persensky, J.J. 1997. Effects of advanced plant design features and control room staffing on operator and plant performance. *IEEE Conference on Human Factors and Power Plants* 5.7–5.12.

Hancock, P.A., Mihaly, T., Rahimi, M. & Meshkati, N. 1988. A bibliographic listing of mental workload research. In P.A. Hancock & N. Meshkati (Eds). *Human mental workload*. New York: North-Holland.

Hart, S.G. & Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload*, Amsterdam: North Holland.

Hill, S.G., Iavecchia, H.P., Byers, J.C., Bittner, Jr., A.C., Zaklad, A.L. & Christ, R.E. 1992. Comparison of four subjective workload rating scales. *Human Factors* 34(4): 429–439.

Hoc, J.M. 1988. *Cognitive psychology of planning*. London: Academic Press.

Huang, F.H. & Hwang, S.L. 2009. Experimental studies of computerized procedures and team size in nuclear power plant operations. *Nuclear Engineering and Design* 239: 373–380.

Hwang, S.L., Yau, Y.J., Lin, Y.T., Chen, J.H., Huang, T.H., Yenn, T.C. & Hsu, C.C. 2008. Predicting work performance in nuclear power plants. *Safety Science* 46: 1115–1124.

Johnston, J.A., Smith-Jentsch, K.A. & Cannon-Bowers, J.A. 1997. Performance measurement tools for enhancing team decision making. In M.T. Brannick, E. Salas, & C. Prince (Eds.). Team performance assessment and measurement: Theory, methods and applications (pp. 311–327). Mahwah, New Jersey, Lawrence Erlbaum Associates.

Jones, J.M., Ma, R., Starkey, R.L. & Ma, Z. 2007. Information complexity and appropriate interface design in nuclear power plant control rooms, *proceeding of Human Factors and Power Plants and HPRCT 13th Annual Meeting*, 45–49.

Kidd, J.S. 1961. A comparison of one-, two-, and three-man work units under various conditions of workload. *Journal of Applied Psychology* 45: 195–200.

Kim, S.C., Chang, S.H. & Heo, G. 2006. Team crystallization (SIO2): Dynamic model of team effectiveness evaluation under the dynamic and tactical environment at nuclear installation. *Safety Science* 44: 701–721.

Kleinman, D.L. & Serfaty, D. 1989. Team performance assessment in distributed decisionmaking. *Simulation and Training Research Symposium on Interactive Networked Simulation for Training*, University of Central Florida, Orlando, FL.

Klimoski, R. & Jones, R.G. 1995. Staffing for effective group decision-making: key issues in atching people and teams. In: Guzzo, R.A., Salas, E., Associates, (Eds.), *Team Effectiveness and decision Making in Organizations* (San Francisco: Jossey- Bass) 291–332.

Kontogiannis, T. & Kossiavelou, Z. 1999. Stress and team performance: principles and challenges for intelligent decision aids. *Safety Science* 33: 103–128.

Lin, C.J., Jou, Y.T., Yenn, T.C., Hsieh, T.L. & Yang, C.W. 2009. A study of control room staffing and workload from the human information processing perspective, *2009 IEEE International Conference on Networking, Sensing and Control*, Okayama City, Japan.

MacMillan, J., Entin, E.E. & Serfaty, D. 2004. Communication overhead: The hidden cost of team cognition. In: Salas, E., Fiore, S.M. (Eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association, Washington, DC, 61–82.

Mathieu, J.E., Heffner, T.S., Goodwin, G.F., Salas, E. & Cannon-Bowers, J.A. 2000. The influence of shared mental models on team process and performance. *Journal of Applied Psychology* 85: 273–283.

Mearns, K., Flin, R., Gordon, R. & Fleming, M. 2001. Human and organisational factors in offshore safety. *Work and Stress* 15(2): 144–160.

Mo, K., Lee, S.J. & Seong, P.H. 2007. A neural network based operation guidance system for procedure presentation and operation validation in nuclear power plants. *Annals of Nuclear Energy* 34: 813–823

Moray, N. 1988. Mental workload since 1979. *International Review of Ergonomics* 2: 123–150.

Morgan, B.B., Jr., Glickman, A.S., Woodard, E.A., Blaiwes, A.S. & Salas, E. 1986. *Measurement of team behaviors in a Navy environment* (NTSC Tech. Rep. No. TR-86-014). Orlando, FL: Naval Training Systems Center.

Morgan, B.B., Jr., Salas, E. & Glickman, A.S. 1994. An analysis of team evolution and maturation. *Journal of General Psychology* 120: 277–291.

Mumaw, R.J., Roth, E.M., Vicente, K.J. & Burns, C.M. 2000. There is more to monitoring a nuclear power plant than meets the eye. *Human Factors* 42(1): 36–55.

Nachreiner, F., Nickel, P. & Meyer, I. 2006. Human factors in process control systems: The design of human-machine interfaces. *Safety Science* 44: 5–26.

Navon, D. & Gopher, D. 1979. On the economy of the human processing system. *Psychological Review* 91: 216–234.

Norman, D.A. & Bobrow, D.G. 1975. On data-limited and resource-limited processes. *Cognitive Psychology* 7: 44–64.

Nunnally, J.C. 1978. *Psychometric Theory*, NY: McGraw-Hill.

Nygren, T.E. 1991. Psychometric properties of subjective workload measurement techniques: implications for their use in the assessment of perceived mental workload. *Human Factors* 33(1): 17–33.

O'Donnell, R.D. & Eggemeier, F.T. 1986. Workload assessment methodology. In: Boff, K., Kaufman, L., Thomas, J.P. (Eds.), *Handbook of Perception and Human Performance*. Wiley-Interscience, New York, 42.41–42.49.

O'Dea, A. & Flin, R. 2001. Site managers and safety leadership in the offshore oil and gas industry. *Safety Science* 37: 39–57.

O'Hara, J.M. & Hall, M.W. 1992. Advanced control rooms and crew performance issues: Implications for human reliability, *IEEE transactions on nuclear science* 39: 919–923.

O'Hara, J.M., Higgins, J.C., Persensky, J.J., Lewis, P.M. & Bongarra, J.P. 2004. Human Factors Engineering Program Review Model (NUREG-0711, Rev.2), U.S. Nuclear Regulatory Commission, Washington, D.C.

Orasanu, J. & Salas, E. 1993. Team decision making in complex environments. In Klein, G.A., Orasanu, J. & Calderwood, R. (Eds.), *Decision making in action: Models and methods* (pp. 327–345). Norwood, NJ: Ablex.

Oser, R.L., Mccallum, G.A., Salas, E., & Morgan, B. Jr. 1989. *Toward a Definition of Teamwork: An Analysis of Critical Team Behaviour*. NTSC technical report no. 90-009 (Orlando: Naval Training Systems Centre).

Park, J.K., Jung, W.D., Ha, J.J. & Shin, Y.H. 2004. Analysis of operators' performance under emergencies using a training simulator of the nuclear power plant. *Reliability Engineering and System Safety* 83: 179–186.

Pinto, M.B. & Pinto, J.K. 1991. Determinants of cross-functional cooperation in the project implementation process. *Project Management Journal* 22(2): 13–21.

Plott, B.M., Scott-Nash, S., Hallbert, B.P. & Sebok, A.L. 1995. Computer modeling of a nuclear power plant operating crew to aid in analysis of crew size issues. *Proceedings of the Human Factors and Ergonomics Society*, 1214–1218.

Plott, C., Engh, T. & Barnes, V. 2003. Technical basis for regulatory guidance for assessing exemption requests from the nuclear power plant licensed operator staffing requirements specified in 10 CFR 50.54 (m). US Nuclear Regulatory Commission, Washington DC, NUREG/CR-6838.

Reich, B.H. & Benbasat, I. 2000. Factors that influence the social dimension of alignment between business and information technology objectives. *MIS Quarterly* 24(1): 81–103.

Roth, K. 1995. Managing international interdependence: CEO characteristics in a resource-based framework, *Academy of Management Journal* 38(1): 200–231.

Rubio, S., Diaz, E., Martin, J. & Puente, J.M. 2004. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile methods. *Applied Psychology: An International Review* 53(1): 61–86.

Salas, E., Dickinson, T.L., Converse, S.A. & Tannenbaum, S.I. 1992. Toward an understanding of team performance and training. In R.W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Norwood, NJ: Ablex.

Sanders, M.S. & McCormick, E.J. 1993. *Human Factors in Engineering and Design*. McGraw Hill, New York.

Scholtes, P.R. 1988. *The team training handbook*. Madison. WI: Joiner Associates.
Sebok, A. 2000. Team performance in process control: influences of interface design and staffing levels. *Ergonomics* 43(8): 1210–1236.
Serfaty, D., Entin, E.E. & Volpe, C. 1993. Adaptation to stress in team decision-making and coordination. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.
Smith-Jentsch, F., Bowers, C.A. & Holmes 1995. The acquisition and decay of aircrew coordination skills. In R.S. Jensen & L.A. Rakovan (Eds.), *Proceedings of the 8th International Symposium on Aviation Psychology* (pp. 1063–1068). OH: The Ohio State University.
Steiner, I.D. 1966. Models for inferring relationships between group size and potential group productivity, *Behavioral Science* 11: 273–283.
Steiner, I.D. 1972. *Group process and productivity.* New York: Academic Press.
Stout, R.J., Cannon-Bowers, J.A., Salas, E. & Milanovich, D.M. 1999. Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors* 41(1): 61–71.
Suchman, L.A. 1987. *Plans and situated actions*. Cambridge, MA: Cambridge University Press.
Thornton, C., Braun, C., Bowers, C. & Morgan, B. 1992. Automation effects in the cockpit: A low-fidelity investigation. *Proceedings of the Human Factors Society, 36th Annual Meeting*, 1, 30–34, Atlanta, GA Human Factors Society.
Urban, J.M, Bowers, C.A., Monday, S.D. & Morgan, B.B., Jr. 1995. Workload, team structure, and communication in team performance. *Military Psychology*, 7, 123–139.
Urban, J.M., Weaver, J.L., Bowers, C.A. & Rhodenizer, L. 1996. Effects of workload and structure on team processes and Performance: Implications for Complex Team Decision making. *Human Factors* 38: 300–310.
Veltman, J.A. & Gaillard, A.W.K. 1996. Physiological indices of workload in a simulated flight task. *Biological Psychology* 42: 323–342.
Vicente, K., Mumaw, R. & Roth, E. 1997. Cognitive functioning of control room operators—Final phase. *AECB*, 96–175, Atomic Energy Canadian Bureau, Ottawa, Ontario, Canada.
Weil, S.A., Hussain, T., Diedrich, F., Ferguson, W. & MacMillan, J. 2004. Assessing distributed team performance in DARWARS training: Challenges and methods. *Proceedings of the 2004 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*.
Wickens, C.D. & Hollands, J.G. 2002. *Engineering Psychology and Human Performance*, Prentice Hall.
Xie, B. & Salvendy, G. 2000. Review and reappraisal of modeling and predicting mental workload in single- and multi-task environments. *Work and Stress* 14(1): 74–99.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Eye-tracking analysis for the information reception of breast self-examination with text and graphical image format

Fang Suey Lin, Chih Lin Tseng & Ching Ying Lee
Department and Graduate School of Visual Communication Design, National Yunlin University of Science & Technology, Yunlin, Taiwan

ABSTRACT: Healthcare participation requires considerable health literacy, which entails understanding the diagnosis and treatment, as well as communication and decision-making skills. The aim of this study is to evaluate how people perceive and comprehensive breast self-examination processes information in text and graphical format. Experiment method was applied to observe the participants' visual perception, include scan paths, fixations by using eye-tracking system, through objective observation of empirical research to explore the procedures and evaluation of questionnaire after experiment, for understanding the situation how people contrast with healthcare information of the message reception.

In expectation of healthcare education content applied as reference in healthcare education content performance, and valuable for improved interface design, as well as for more accurate interpretations of implicit feedback for learning. The results were as follows:

1. Participants' generally believe that by measuring the form of graphic and text can easily achieve to help understand information content.
2. Graphical images help users to remember firmly the content of healthcare education information.
3. Both of time on message reading and a point of view of time to stay in the message are influenced on the level of message receiving and understanding.
4. Participants stay for the main title time shortly, spent longer with the text browser and graphic image contrast.
5. Participants' stay in graph message longer than stay in the text message.
6. Healthcare information content of the graphic image with the directional signs that will be assisted understanding, reading and memory on action.

Keywords: eye-tracking analysis, breast self-examination, healthcare education information

1 INTRODUCTION

1.1 *Motives and purposes*

Providing health information is a major task in health care since it impacts on health behavior, treatment and decision-making (Kreps, 1988). According to the Department of Health's "Report of Registered Cancer Cases" (Department of Health, 2007), cancer was the first cause of death in Taiwan, and breast cancer ranked the number one of female cancer. Thus, in order to reduce mortality, promoting breast examination is an important strategy for national health and disease prevention. Breast cancer can be detected early when treatment is available. Therefore, the real-time control, conveyance, and application of health information are ignorable infrastructures. Internet has become an important source of breast cancer information, and allows patients to discuss their conditions with their physicians more effectively (Eysenbach G, 2003). The effective health information not only can help promote health information but also is closely related to health concepts and health-seeking behavior, which can be used to achieve the health concept of "prevention is better than cure."

Based on the Breast Self-Examination (BSE) presented by health information, the study used eye-tracking method to record the scan-path of people when watching visual stimuli to understand people's ocular response when receiving the online graphical-text combination information. The purposes of this study are:

1. To combine objective psychological analysis with sentimental one: the eye-tracker was used as the experimental method for measuring eye movement. The objective result obtained was used to compensate for the error caused by

sentimental judgment to increase the reliability of the results and the value as reference.
2. To provide the content of health information with more accurate design basis: correspond to human factor of visual design concern were used as the samples the connection between visual attention and information reception could be understood through this study, and the results obtained could serve as reference to the design of the content of online health information.

1.2 *Research scope and limitations*

Because the age of onset of breast cancer in women in Taiwan has gradually declined year by year, both domestic and foreign medical institutions have been promoting relevant information on breast cancer prevention for young women in recent years to further put the health concept of "prevention is better than cure" into practice. The scope and limitation of the study are as follows:

1. The participants are women age of 30–39, and all had the experience of using internet.
2. There is a lot of breast cancer-related health information; however, only the health education promotion of BSE for prevention was used as test content.
3. The common form presented by online health information content, graphical-text combinations were used as experimental samples. The variable of color was eliminated and health information was exhibited in an achromatic manner.

2 LITERATURE REVIEW

2.1 *Eye movement and information receiving forms*

Atkinson and Shiffrin suggested that the memory system of people is to receive the information from external environment, convert it, store it, and utilize it, suggesting that information is conveyed from external environment to the sensory register (Atkinson & Shiffrin, 1971). Visual communication relies on the two-way communication platform between the information producer and the information receiver and is generate by the interaction between visual sense and visual perception, which is the reaction equilibrium between rationality and intuition. When visual organ receives external stimulus, it will further induce sensorium to change whereas visual perception is generated after visual sense. When sensorium is stimulated, it will make a meaningful interpretation to the stimulus. As for visual communication, after the visual sense is stimulated, visual communication will make it to make a meaningful interpretation to the stimulus. Psychologists (Hochberg & Brooks, 1962; Messaris, 1994; Rieber, 1995) also suggested that people's sense of image stimulation is an innate instinct. However, the comprehension of text information is the result of acquired learning. Many studies also indicated that people's memory of images is stronger than that of texts (Mandler & Ritchey, 1977; Anderson, 1990).

Human visual system has to keep the image of an object located at the fovea of the retina to obtain clear visual image. Therefore, both eyes have to focus on the object observed. The behavior of focusing the eyes on the object is called "fixation." The movement of eyeballs of the visual system to fulfill the fixation among different objects is called "saccade (Rayner, 1998)." The repeated process contributes to the ocular movement. Eye movement is the most important source of sensory information in cognitive process because more than 80% of information is obtained from the visual cognitive process during human information processing (Sanders & McCormick, 1992), Eye tracking has also recently been applied to help answer questions such as how exactly different learning goals (Sweller & Levine, 1982), and it will affect visual attention and the learning of the processing of image (Brunye' & Taylor, 2009).

2.2 *The human factors of online health education information*

Rapid development of Internet technology has significantly influenced health care globally. The communication of health information changes from passive report in the past to active monitoring and multi-interaction online, so that the latest evidence can be reported quickly (National Health Insurance Bureau, 2007). Therefore, in addition to the functions of communication and education, another major function of health education information is to arouse the awareness and raise the alarm of the public on their own health care issues, and the effective and comprehensible health information content can help people memorize and learn more. At present, the online health information forms of breast self-examination process, is mainly presented through graphical-text combinations (Figure 1). Three human factor criterions of visual design visibility are legibility and readability. Eastman Kodak Company (1983) mentions front to use on interface design will be divided into comprehensibility, legibility and readability. Comprehensibility is the meaningfulness of text information to receptor, and also relevant to user background, knowledge and language ability. The rapid growth and advancement of modem

Figure 1. Online health education information (Source: http://www.breastcancer.org/symptoms/testing/types/self_exam/bse_steps.jsp' 2009/01).

Table 1.

	Front size	Visual distant
Literature	mm	m
Kodak (1983)	2~5	0.7
	3~7	0.9
Sato (1992)	2	0.3
	4	0.6
	10	1.5

desktop publishing, Web-based publishing, and multimedia presentation continue to compound the confusion with increasing font and layout capabilities, display and print options, and the need to effectively integrate with other media. The guidelines including size, typeface, contrast, text blocks, spacing are address common issues regarding text legibility (William et al., 2003). Readability addresses the question of whether something is easy to read. The term applies both to how text and graphics are laid out on a page, and the style and structure of writing (Kathy Haramundanis, 2001). The related research indicated the viewing distance and Font size are depended on different from circumstances (Table 1), such common expression forms used in this study as the reference for designing the samples of stimulus materials in the experiment to understand participants' extent of visual attention and information reception for graphical-text combinations to propose the suggestions on the principle for online health education information.

3 METHOD

3.1 *Participants, experimental environment and instruments*

Six female participants aged of 30–39 and the average age was 33.5 years old. The eyesight of the participants had to be normal after vision correction. All the participants had normal color discrimination ability, and neither did they have a visual disability nor an eye-related disease.

The study was conducted in the laboratory of eye-tracker in the Graduate School of Industrial Design of National Yunlin University of Science and Technology in Taiwan.

1. Environmental design: the lamplight in the laboratory was the 6500 K white light. Before the formal experiment was initiated, the participants were requested to sit at approximately 120 cm in front of a 40 inch LCD screen where a straight line would be formed between the center of the screen and the subjects (Figure 2).
2. Experimental Equipments: the participants were requested to take the seat and the wear the modulated Eye Link 1000 eye-tracker which could detect the pupils and retina of the participants. The frequency was set at 250 Hz for recording the scan path of their right eyes.
3. Experimental calibration: before using the eye-tracker for collecting data, nine calibration points were used for calibration.

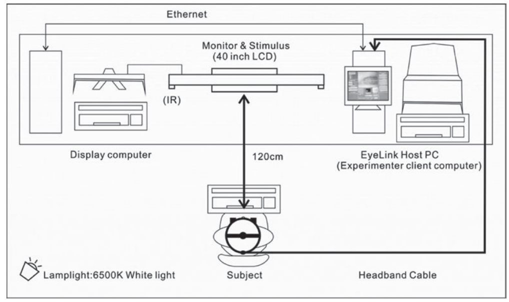


Figure 2. Eye-tracking environmental setting.

3.2 *Stimulus design and experimental procedures*

In this study, the eye tracker was used to conduct the eye-tracking experiment. The participants were requested to wear the eye-tracker to watch the testing pictures on the screen. Meanwhile, their eye movement behaviors were recorded, including "fixation," "saccade," and "scan path." Afterwards, the questionnaires were distributed to the participants to investigate their extent of understanding and preference to understand their visual attention and information reception of health education information content.

The research procedures are as follows:

1. The information content of the seven procedures of BSE, including three visual examination procedures and four palpation procedures. The graphical-text information form corresponds to

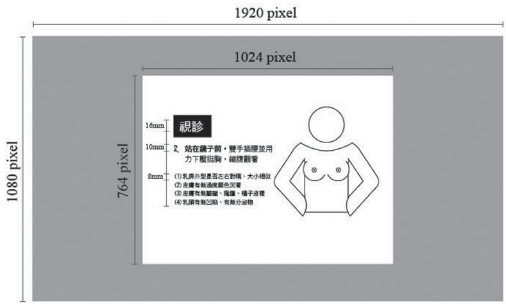


Figure 3. Stimulus design.

human factor of visual design concern was used as the samples (Figure 3).

2. The watching duration was not regulated in the operational model of the experiment. Participant proceeded viewing and clicking into online information as the custom model. Therefore, the participants could press the space bar on the keyboard to continue browsing next testing sample after they watched each procedure.
3. After completing the measurement of the experiment, a questionnaire survey would be performed on each participant. The questionnaire was divided into three parts: Part one, the basic information of the participant; Part two, survey on the comprehension of BSE information content; Part three, the feeling of the participant after watching the information content. Each participant will be inquired to provide their opinion after filling out questionnaire. It took approximately 30 minutes to complete the entire experiment.

4 RESULTS AND DISCUSSION

4.1 *Analysis and discussion on the results of fixation duration and visual attention*

The focus of the observation was to understand the participants' information receiving circumstances of the expression forms of graphical images and texts. The eye-tracker was used to record the participants' fixation positions when watching experimental materials. The six participants were encoded as A, B, C, D, E, and F. The center of the screen, 960 pixel (X) × 540 pixel (Y), was used as the baseline to divide the fixation position of the participants to the left (text) and the right (graphical). The time the participants spent on gazing at the images and texts was further investigated. In terms of the fixation time for reading the left texts, participant E spent the longest time (approximately 115 seconds) whereas participant D spent the shortest time (approximately 87 seconds) likewise. However, as for the time spent on reading texts, participant E spend 69% of the total fixation duration on reading texts whereas subject D spent nearly 90% of time on it. As for the time spent on watching the right graph, participant E still spent the longest time (approximately 51 seconds) whereas participant F spent the shortest time (approximately 11 seconds). In addition, among all the participants, participant E also spent the longest time on watching graph. On average, the participants spent 77% of time gazing at the left text content.

The research result indicated that it indeed took more time to receive text information. Besides, because the stimulus materials of the testing graphs were the graphs that the general public familiar with, while the testing texts were the Medical terms and explanations that the general public seldom read, which might lead to the difference in the fixation time for graphs and that for texts.

As for the average fixation time of the participants spent on each fixation point, which was the time that the viewpoint staying at the graph and text area, as shown in Table 3. In general, the time that participants' viewpoint staying at texts was shorter than that at the image. Such a phenomenon conformed to the relevant research results above. The fixation time of the participants for graphs and texts was further investigated and it was found that the reading model of participant D was opposite to that of participant E. The fixation duration of participant D was the shortest whereas that of participant E was the longest. However, as for the time that the viewpoint stayed at the graph or text, that of D was longer than E. It's indicated that the reading model of participant D helped understand the overall information. However, it did not help memorize information. The circumstance of participant E was opposite. Such a phenomenon seemed to reflect that the time for receiving overall information and the time staying at the information had difference influences on the extent of information content received. It could also be inferred that the expression form of testing materials, the graphs and

Table 2. Total (TFD), and Left text (LTFD), Right graph (RGFD) fixation duration and ratio.

Sub	TFD	Sequence	LTFD	Ratio	RGFD	Ratio
A	157124	2	114392	73%	42732	27%
B	132548	4	102644	77%	29904	23%
C	141024	3	102296	73%	38728	27%
D	100332	6	86844	87%	13488	13%
E	166004	1	114524	69%	51480	31%
F	101428	5	90496	89%	10932	11%

Table 3. Total (TFP), and Left text (LTFP), Right graph (RGFP) fixation point and Sequence.

Sub	TFP	Sequence	LTFP	Sequence	RGFP	Sequence
A	253	3	255	1	247	5
B	260	1	239	3	379	1
C	244	4	224	5	317	2
D	257	2	254	2	281	3
E	241	5	227	4	277	4
F	226	6	224	5	243	6

texts, had an influence on the information received and understood by an individual.

According to the scan path of the participants (Figure 4), when reading text information, the eyeballs will move up and down or from the left to the right along the texts line by line and fixation and saccade will be used in turn. When reading texts, the range of eye movement and average fixation time was smaller and shorter than watching graph. Besides, the movement of eyes was in a regular pattern, which was from the left to the right or from the top to the bottom. The direction of visual attention was from the title on the top left to the graph on the right, and the participants would scan the graph and text repeatedly. However, when the fixation figure (Figure 5) of the participants was investigated, it was found that once the figure was marked with a specific symbol or detection area, the participants' attention to the specific area would increase. Therefore, the scan paths of the six participants recorded during the experiment were further converted into visual attention map and overlapped. In addition, the map was further divided into nine visual attention areas (A1~C3) to understand the participants' visual attentions expression area when reading health education information content. It was found that when the participants read the first page of the health education information content, visual examination (Figure 6), most of their visual attention was located at A2 area, followed by B2 area, while the line of sight would temporarily stay at A1 and C2 areas. When the participants watched the fifth page—palpation (Figure 7), the visual attention was significantly located at C2 areas, followed by A3 and B3 areas. It could be concluded that the health education information content graph with directional symbol could arouse the attention, and people would stay at the graphical information significantly for a longer time than text information.

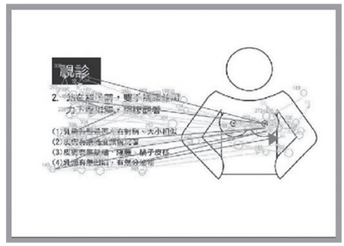


Figure 4. Scan-path figure.

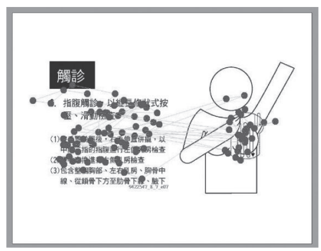


Figure 5. Fixation figure.

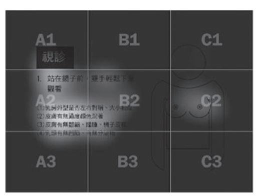


Figure 6. Visual exam.

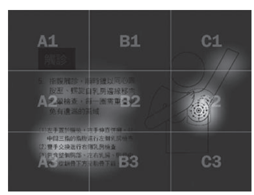


Figure 7. Palpation.

4.2 *The result of analysis on the information reception and learning effectiveness*

After completing the eye-tracker experiment, all the participants had to fill out a questionnaire relevant to the experimental content to investigate whether time had an influence on the correct receiving of information (e.g. whether the participants who spent more time had a higher accuracy rate in filling out the questionnaire).

In general, the participants all accurately understood the information content of the seven procedures of BSE. The accuracy rate of four participants' answers to the five questions was 80%, while that of other two participants was 67%. All the participants answered correctly the question the seven procedures. However, only one

Table 4. Part two questionnaire correct answer rate.

Items	A	B	C	D	E	F
2-1 BSE Procedures	O	O	O	O	O	O
2-2 BSE Method	X	X	X	O	X	X
2-3 The content visual exam can't check out	O	X	X	O	O	O
2-4 Draw the direction of palpation	O	O	O	O	O	O
2-5 Circling the major palpation area	O	O	O	X	O	O
Correct answer rate	0.80	0.67	0.67	0.80	0.80	0.80

participant answered correctly that there are two BSE methods, suggesting that most of the participants would skip reading the title and focused on the content and graphs. As for the graphs, 6 participants correctly drew the direction of palpation. Among them, participants B, D, E, and F could fill out two directions of palpation while C could fill out three directions of palpation. Participants C and D further wrote supplementary description about the information content. As for the question of circling the major examination area for palpation, only participant D did not answer the question whereas participants A, B, C, E, and F could correctly circle the scope of palpation examination. Between them, participants C and F could draw the dashed line as the presented in the information content, suggesting that the graphical-text information presentation step by step would help participants understand and memorize the content.

4.3 *The result of analysis on extent of preference*

Based on the participants' reading and watching of the health education information content, the experience of BSE procedures, the extent information received was rated to present the sentimental psychological analysis according to the extent of approval. It was found that the six participants extremely approved that the presence of image in the health education information content would help grasp the focus of the content (M = 8.33). They all approved that the presence of graphs in the health education information content could attract the browsing more (M = 7.67). They all approved that the health education information content was clear for blowing (M = 7.17). However, the participants disagreed that the title of the BSE could be easily distinguished from the texts in the information content (M = 3.17). Among the six participants, 2 of them disagreed whereas four of them slightly approved it. It could be inferred from the accuracy rate of Part two questionnaire that the title of the information content and the texts failed to be easily distinguished from one another and failed to be effectively memorized. The participants disagreed with the idea that only texts are provided can the information content be easily understood (M = 2.67). 4 participants disagreed that only graphs are provide can the information content be easily understood (M = 3.50), suggesting that in general, the participants suggested that the graphical-text combinations could more help understand Health education information content.

Finally, the participants filled completely out questionnaires and they provide fitly opinions of experiment contents and samplings

Table 5. The recording of interviews.

Sub	The point of opinions
A	1. Text edition will influence reading. 2. Guidance appearing in the graphical can assist to read. 3. It can recall the contents through watching graphs when fills out the questionnaire.
B	The participants once took screening of breast, but they didn't pay attention to the relevant information, however they could recall the contents through this testing.
C	1. Participants could remember clearly contents and could write down the key point. 2. To suggest the sign of arrow enlargeable.
D	1. There were difficult medical terms in the pre-three pictures. Participants would make clearer about the contents, if provided graphs comparing with. 2. It will be batter, if can click online.
E	Participants can learn definitely relevant BSE through this research.
F	Participants can learn definitely relevant BSE through this research.

after brief interviewing (Table 5). The layout of graphical-text combinations and symbols was suggested by A&C, D pointed out the medical and nursing-related terms that would influence text comprehensive, however entire participants agreed graphical-text form to assist health education information content understanding, reading, and memorizing.

5 CONCLUSION

As regards the eye-tracker experiment and the questionnaire survey, the following results are found:

1. Both the time spent on reading information and the time the view point staying at the information had an influence on people's information receiving and extent of understanding.
2. In general, the Participants suggested that the graphical-text combinations could assist to understand the information content more whereas graphs could help users memorize health education information content more. The information received understood by people may vary with the difference in an individual himself and graphical-text expression of information.
3. People spend longer time on graphical information than text information. The presence of directional symbols in the graph of health education information content can help understanding, reading, and memorizing the actions.
4. The testing items of health education information content, medical and nursing-related terms, and relevant descriptive the symptoms on BSE terms were incomprehensible to the general public. Therefore, the graphs had to be compared with the texts.

According to the interviews with the participants after the questionnaire survey, the following opinions were proposed as future study directions:

1. The grey-scale graphs were presented in the health education information content. Thus, it is suggested to use color pictures and operational factors in future studies to reduce the learning effects and conduct relevant differential study on the extent of information received.
2. This study correspond to human factor of visual design concern are used as the samples only, the further interaction form of click online will be discussed on the issues of usability in future studies.

REFERENCES

Atkinson, R.C. & Shiffrin, R.M. 1971. *The control processes of short-term memory*. Technical Report (173)1–23.

Anderson, J.R. 1990. *Cognitive psychology and it implications* (3rd ed.). New York: W.H. Freeman and Company.

Brunye', T.T. & Taylor, H.A. 2009. When goals constrain: eye movements and memory for goal-oriented map study. *Applied Cognitive Psychology* 23: 772–787.

Department of Health 2007. *Report of registered cancer cases*. Taipei, Taiwan.

Eastman Kodak Company, 1983. *Ergonomic Design for People at Work(1)*, Nostrand Reinhold Company, New York.

Eysenbach, G. 2003. The impact of the Internet on cancer outcomes. *CA Cancer J Clin* 53: 356–71.

Henderson, J.M. & Hollingworth, A. 1999. High-level scene perception. *Annual Review of Psychology* (50)243–271.

Hochberg, J.E. & Books, V. 1962. Pictorial recognition as an unlearned ability: A study of one child's performance. *American Journal of Psychology* 75: 624–628.

Hyona, J., Lorch, R.F., Jr. & Kaakinen, J.K. 2002. Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *Journal of Educational Psychology* 94(1): 44–55.

Haramundanis, K. 2001. Learnability in information design, *SIGDOC 01'*, 7–11.

Kreps, G.L. 1988. The pervasive role of information in health care: Implications for health communication policy. In J. Anderson (Ed.), *Communication Yearbook* 11: 238–276.

Mandler, J.M. & Ritchey, G.H. 1977. Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory* 3(4): 386–396.

Messaris, P. 1994. *Visual literacy: Image, mind, and reality. Colorado.* West view Press, Inc.

National Health Insurance Bureau 2007. *Annual Report Bureau of Health Promotion*, Taiwan.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3): 372–422.

Rieber, L.P. 1995. A historical review of visualization in human cognition. *Educational Technology Research & Development* 43(1): 45–46.

Sanders, M.S. & McCormick, E.J. 1992. *Human Factors in Engineering and Design* (7th ed.), McGraw-Hill Books, New York.

Sweller, J. & Levine, M. 1982. Effects of goal specificity on means-ends analysis and learning, Journal of Experimental Psychology: Learning, Memory, and Cognition, (8)463–474.

William, Kritina & Jill. 2003, Universal Principles of Design, Rockport Publishers. Inc.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

The study to trimming and harvest performance by ergonomic grape scissors

HsuanChu Chen & Ching Yang
Department of Design Doctoral Program, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, ROC

ABSTRACT: The study is aimed at discussion over the tool and performance of grape industry in Changhua County, Taiwan. In the preliminary investigation to the tool used, agricultural performers are used to apply dedicated grape scissor. In case there is no suitable tool for carrying out the performance, they are inclined to use other tools such as general scissors or blade for substitution, or even direct hand-pick.

On the other hand, for performance, it can be divided into two categories as trimming and harvest work. Trimming is to remove the inferior grapes or prune them to clusters of shape and cut off the leaves; the other purpose of trimming is to enable sweetening and juicing of the fruits, nevertheless, it is necessary to notice not to harm the grape and sarcocarp. And the harvest work is to cut the stalk of grape and take the mature fruit while the other hand carefully holding it up in purpose of preventing it from dropping to loose and unstrung. For the sake, it is clear to see that the tool they are to use is almost from the beginning to the end as the two works said above are seasonal dense task that the post of arms and wrists acting for competition of these works likely incur Repetitive Motion Injury (RMI) and Carpal Tunnel Syndrome (CTS). After interview, investigation over the workers on site, most of them appear the foresaid two situations, and the pain areas are shoulder, arm and hand.

In the case, the way the study is to explore the tool and performance survey pre-improvement and makes five different modeling of grape scissors (G1~G5) by bettering angle of scissor opening and holding currently applied. There are two test missions to do, one is trimming task (T1) and the other is harvest task (T2) for testing grape cut performance and investigation of application while the test is undertaken by 30 people as subjects that are organized by new worker and experienced worker half and half. The outcome is found out that no matter new or experienced worker, they perform quite more excellent than the original one when they use the tool with shearing mouth bend G2 (30°); In respect to tool familiarity, experienced worker performs greater as for he/she is influenced by habit of usage, but in terms of investigation over application condition, he/she expresses no better performance. In the use of new tool, new worker seems showing the higher acceptance degree and performance than experienced worker. Therefore, the study result can provide significant references to exploration of improvement, habit of usage, mental relationship and performance of tool, as well as give important information to industry as for continuous research and development of ergonomic tool in future.

Keywords: Grape Scissors, Ergonomic tool, Trimming task, harvest task, neutral position, performance

1 INTRODUCTION

Taiwan's grape agriculture record began from Qing dynasty emperor Kangxi Year 12 (1684), it was advanced by test trial of grape wine in 1955 and gradually enlarged the measures for encouragement of production. Until around 1960, Taiwan initially implemented new species like Gu-Feng from Japan, the way it finally appeared small scope of economic planting. As it reaches the amazing annual output as high as NT $2.5 billions, thus it gradually became the vital, economic fruit in Taiwan. The major plant districts are Changhua, Taichung, Nangtou and Miaoli; the regular production reveals two harvest seasons a year (summer and winter). The often-seen cultivation species in Taiwan are Kyoho (Kyoho)' Italy (ltalia)' Jin-xiang (Golden Muscat) and Honey red, etc.

As being the instance, Taiwan owns a great deal of grape farmers but reserves very few of vocational injury data of the industry. In view of the fact, the study is executed to research grape trimming and harvest, applied tool as well as the influence on physical parts as wrist, elbow and arm, etc; the study is carried out by observation, record and interview of the grape garden to probe into the issues related to farm worker's body and mind, as well as the tool he/she uses in order to conclude the goal of improving grape scissors to the one fitting to ergonomic tool.

Trimming and harvest are the two major motions that grape worker regularly acts; the two tasks requires a heavy amount of motions by wrist, elbow, arm and shoulder to complete trimming and collection of grape. "Trimming work": A lot of motions like wrist bend and long-term arm and should raise high are involved. For the reason, for years, the conditions tend to come to accumulate pain, ache and injury of their bodies.

Ergonomics is also known for the role it plays in reducing inefficiencies by systematically searching for the solutions to reduce error. Ignoring continuous error and/or physical stress will cause trauma disorders such as Repetitive Motion Injuries (RMI). A common-seen RMI symptom reported today is Carpal Tunnel Syndrome (CTS). CTS also has been labeled as occupational overuse syndrome or repetitive strain (Konz & Johnson 2000). CTS is caused by compression of the median nerve, which passes through the carpal tunnel in the wrists (Jacksonet et al. 1995). Symptoms include pain, numbness, tingling, clumsiness and lack of sweating in the parts of hand innervated by the median nerve. These symptoms may be occasional rather than constant, with a majority of severe symptoms occurred during the night (Storti 1990).

Storti (1990) reported that bending the wrist forward 90° (acute volar flexion) resulted in a significant increase of pressure in the median nerve in the wrist. The evidence's strength supports the recommendations for workers to keep their wrists straight (neutral position) while rotating the forearm and hand. This led Konz & Mital (1990) to suggest guidelines as preventing and alleviating the symptoms associated with CTS in the design of hand tools. According to the literature mentioned in the post-injury, followed by tools that explore the design principles: four broad categories were listed as follows: frequency (reduce the number of cycles for a specific wrist), joint angle (keep the wrist in the neutral position), force (reduce the amount of force and its duration) and non-ergonomic (medical).

The intent of the prototype was to adhere to the basic guidelines for hand tool design. The following guidelines receive a special attention: #1: 'use of special-purposed tools', #2: 'design tools to be used by either hand', and #7: 'consider the angles of the forearm, grip and tool'. In other words put the "bend in the tool and not in the wrist" Konz & Johnson (2000). Principle of hand tool design: #1: maintain a straight wrist, #2: avoid stress on tissue, #3: avoid repetitive finger action, #4: design for safe operation, #5: remember women and left-Handed. Sanders & McCormick (1993).

2 METHOD

The study applies the methods of observation and record to tape the actual operation of trimming and collection work (Fig. 1), and further note down and analyze the relationship of Radial, Ulnar deviation and Palmer, dorsa flexion of hand and wrist during performance (Fig. 2), the test time remains 1 minute; statistic data of bend angle, direction and number of time is per Table 1.

2.1 *Workstation and operating environment*

Location of the experimental site in this study for observation is Xi-Hu town, Changhua County.

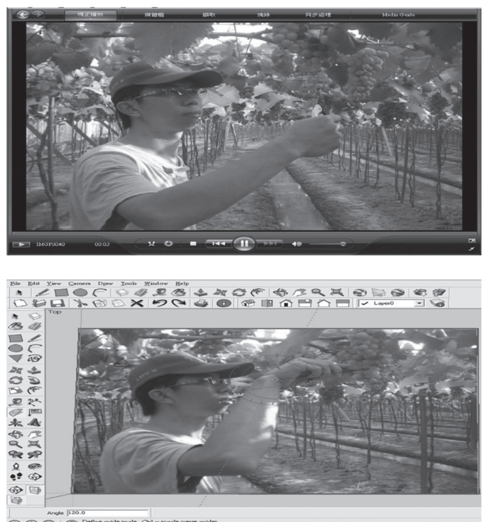


Figure 1. Using micro software media player to play the video and capture the pictures.

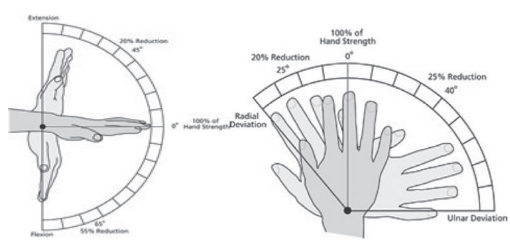


Figure 2. Using the 3D program, google sketch Up 7.0, to measure the angle of wrist and arm postures.

Table 1. Posture observation and recording.

Setting	Angles	Radial ulnar		Palmer dorsa		
Task	Degree	Deviation		Flexion		Total
Trimming	0~30	6	6	8	8	60
	31~60	6	6	12	8	
	61~90	0	0	0	0	
Harvest	0~30	0	0	27	0	39
	31~60	0	0	12	0	
	61~90	0	0	0	0	

Figure 3. Operational tool and work environment.

The height of the shack is 175~180 cm, the height of performance in the experimental site upon fruit maturity is about 165~170 cm while the distance between grape trees is about 1 meter. The grapes trimmed or collected are piled up in the plastic basket on the wagon (Fig. 3).

2.2 *Applied tools*

From the observation and interview, it is to see that the tools applied currently are roughly separated into two types as A and B; their common character is the pointed tip and long blade of scissors, and the hold handle can be divided into no spring flake (spring) and ring hold (handle). Through interview and investigation, if the users can select tool, a majority of them would choose type A; after exploration and inquiry, the reason for they choosing that is type A has a spring flake (spring) to prop up that tends to make user feel comfortable and easy in comparison with type B when they wear gloves in work. And type A is easy to slip as it does not have ring hold (handle).

2.3 *Grape scissor hand tool design*

The study implements the idea of Type A and Type B (Figs. 4–5) to make the minimum amendment and adopt to the most usual tool pattern, as well as refer to ergonomic hair cut literature (Jason L. Boylesa et al. 2003) and put a basis on hand tool design principle: Konz & Mital (1990), M. Sanders & E. McCormick (1993), Konz & Johnson (2000) to process tool alteration; utilize Morphological Analysis Zwicky (1940) to dismember each part to design and derive different grape scissors (Fig. 8). G1 is the original modeling style of Type A, it means Type A, and G2 is 30° bend off the original pattern while G3 is 45° bend off; G4 is single handle grape scissors and G5 is the original modeling style of Type B, it means Type B. The study make a further trial and test over the five grape scissors.

2.4 *Subjects*

Subjects under test of the study are totally 30 persons, half man half woman. All the participators must undergo a questionnaire survey, the contents are,

Figure 4. Scissors type A.

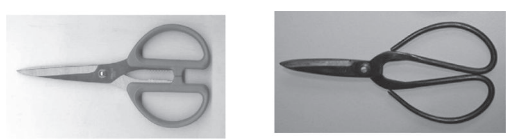


Figure 5. Scissors type B.

Type A Type B

Figure 6. Scissors type A and B.

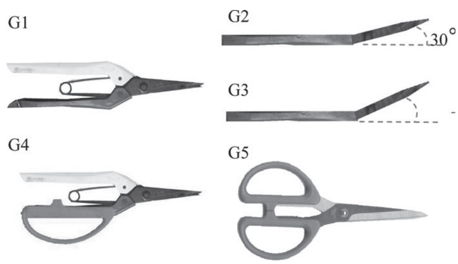


Figure 7. G1 is the original style, G2 is original style with 30° bend off and G3 with 45° bend off, G4 is the single handle Scissor and G5 is the original type B.

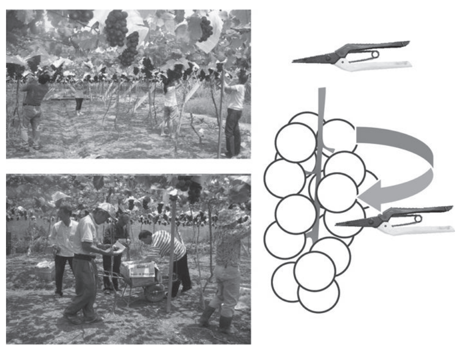


Figure 8. Trimming work task procedure and harvest work task procedure.

for example: gender, age, relative experience, hand injury, muscle/skeleton disease or surgery and handedness, etc., as well as the consent to join in the study. (see Table 2)

2.5 *Experimental design*

Design of the experiment is to have the testee use the five grape scissors to cut and complete the foresaid two missions that mean the tasks of (T1) "Trimming" and (T2) "harvest".

The first mission (T1) is trimming task. (Table 3)

Step 1. Select scissors to trim grape.
Step 2. Trim the full cluster of grapes; when finishing, trim the next string.
Step 3. Calculate the amount of trimming.
Step 4. Questionnaire and evaluation.

The second mission (T2) is harvest task (Table 3).

Step 1. Select scissors to trim grape.
Step 2. Cut the full cluster of grapes.
Step 3. Calculate the amount of harvest.
Step 4. Questionnaire and evaluation.

The above-mentioned two missions are carried out by random trimming and harvest. Test time of the above two missions are 1 minute respectively (Fig. 8).

Table 2. Characteristics of subjects (n = 30).

Characteristics									
Gender		Experience		Age (year)		Weight (year)		Height (cm)	
F	M	Yes	No	M	Range	Mean	Range	Mean	Range
15	15	13	17	40	29–57	57	46–78	164	156–174

Table 3. Processes of trimming and harvest task.

	Trimming task	Harvest task
Step 1	Select scissors to trim grape	Select scissors to trim grape.
Step 2	Trim the full cluster of grapes; when finishing, trim the next string	Cut the full cluster of grapes.
Step 3	Calculate the amount of trimming.	Calculate the amount of harvest.
Step 4	Questionnaire and evaluation	Questionnaire and evaluation

3 EXPERIMENT AND ANALYSIS

3.1 *Experiment and statistical analysis*

In regard of experimental analysis in the study, it is performed by SPSS data analysis. The ANOVA result shows that T1 reveals outstanding dominance (Tables 4–5).

Table 4 displays that the average work of tool G1-G5 is under the standard. In T1 mission, the work with G1 tool acquires 30 pellets while the work with G2 gains 37.8 pellets, with G3 obtains 26.2 pellets, and G4 gains 26.4 pellets, G5 acquires 25.4 pellets. In T2 mission, the work with G1 tool harvests 39.7 strings while the work with G2 collects 39.4 strings, with G3 obtains 34.9 strings, and G4 harvests 37.7 strings, G5 collects 40 strings.

Table 5 show ANOVA and F distributed results for (T1) and (T2) tasks. The indicate (T1) acts obviously excellent performance than (T2). The conclusion of contrast between different tools and tasks by Multiple Comparisons Analysis shows that in (T1) task, G2, G3, G4 and G5 have an evident difference of $\alpha < 0.05$, while in (T2) task, there is no any noticeable diversity. In Figure 9, the Average Line Chart and Error Bar clearly display that in (T1) task, G2 is better than G1, G5, G3 and G4. And in (T2) task, it is learnt from data that G1 is better than G2, G5, G4 and G3.

Apply ANOVA method to analyze performance: (T1) Trimming work appears with obvious differences $0.000 < \alpha$, the way it is not to set up null hypotheses; and in (T1) trimming task, G2 tool acts with much higher performance.

Table 4. G1-G5 tool performances and means standard deviations.

					95% Confidence Interval for mean			
Performance	Tool	Mean	Std. deviation	Std.error	Lower bound	Upper bound	Minimum	Maximum
(T1) Trimming task	G1	30.00	3.97	3.97	27.15	32.84	22.00	35.00
	G2	37.80	4.66	4.66	33.46	40.13	28.00	44.00
	G3	26.20	3.79	3.79	23.48	28.91	19.00	32.00
	G4	25.40	5.94	5.94	22.14	30.65	20.00	41.00
	G5	28.96	5.03	5.03	21.79	29.00	17.00	34.00
(T2) Harvest task	G1	39.70	39.70	10.60	32.11	47.28	24.00	55.00
	G2	39.40	39.40	7.82	33.80	44.99	32.00	57.00
	G3	34.90	34.90	12.73	25.79	44.00	19.00	55.00
	G4	37.70	37.70	10.19	30.40	47.53	24.00	53.00
	G5	38.34	38.34	10.53	32.46	41.24	23.00	56.00

Table 5. The ANOVA for T1 and T2 tasks.

		Sum of squares	df	M.Square	F	Sig.		Statistic[a]	df1	df2	Sig.
(T1) Trimming task	Between groups	893.92	4	223.48			Welch	9.428	4	22.36	0.000***
	Within groups	1014.00	45	22.53	9.918	0.000***	Brown-Forsythe	9.918	4	40.45	0.000***
	Total	1907.92	49								
(T2) Harvest task	Between groups	179.72	4	44.93			Welch	0.300	4	22.34	0.875
	Within groups	4955.50	45	110.12	0.408	0.802	Brown-Forsythe	0.408	4	41.48	0.802
	Total	5135.22	49								

Notes: (T1) Trimming task **p<0.01; ***p<0.001; Asymptotically F distributed, **p<0.01: *** p<0.001.

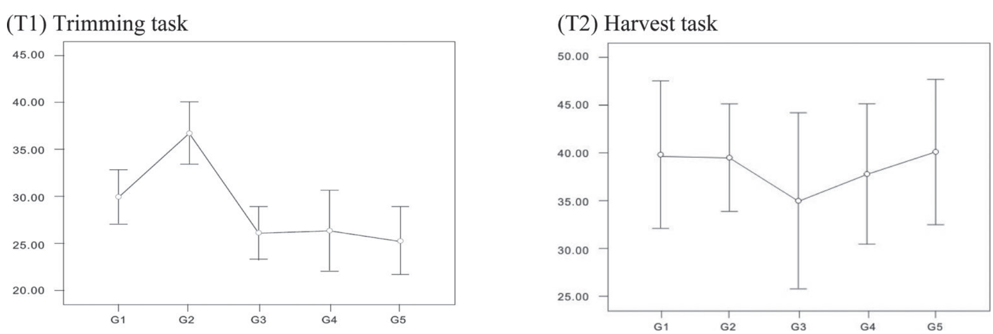


Figure 9. T1 and T2 tasks of means plots and error bar.

3.2 *Subjective ranking*

Post-experimental test, participant is asked to assess and list ranking of the tools based on the performance; in the subjective rank, we can see the ranking order in (T1) trimming task: G1 > G2> G4, G5 > G3, while the ranking order in (T2) harvest task: G1 > G2, G4 > G5 > G3. They indicate that perhaps these tools~G1, G2, G4 and G5 are friendly use, and only G3 is hard to use (Tables 6–7).

4 RESULT AND DISCUSSIONS

In tests of five tools for two missions, the study is aimed to probe into the relationship of performance

Table 6. Subjective rating.

Task	(T1) Trimming work					(T2) Harvest work				
Tools	G1	G2	G3	G4	G5	G1	G2	G3	G4	G5
frequency	30	21	3	15	15	27	21	6	21	15
Ranking	1	2	4	3	3	1	2	4	2	3
Group	A	B	D	C	C	A	B	D	B	C

Table 7. Subjective rating group.

Rank	(T1) Trimming work	(T2) Harvest work
Best	G1	G1
	G2	G2 G4
	G4 G5	G5
Worst	G3	G3

between new worker and experienced worker, and conclusively finds out the following results:

1. Tool performance apparently shows that in (T1) trimming task, performance of G2 tool is much higher than the other tools; and in (T2) harvest work, there is no evident difference for all the five tools.
2. In (T1) trimming task: There is an interesting phenomenon found in the (T1) operation and subjective evaluation, that is, for experienced worker, high performance tool does not mean that it certainly has a relatively high appraisal. In execution of actual performance, new worker's achievements in use of new tool are not inferior to experienced worker. In (T2) harvest task: The figure of Error Bar shows G3 tool results in a comparatively larger difference that indicates only some workers get with the tool but most of workers are unable to adapt to it. The tool with appropriate angle of bend will benefit the performance of a certain type of operation.
3. Mentally, new worker and experienced worker have different feeling in use of new tool; for experienced worker, he/she seems psychically exclude the new tool as he/she has been used to the original tool for a long time, in the case, even though the proof shows new tool reaches a much higher performance of experienced worker, he/she still subjectively evaluates and deems the original tool is the best. Nevertheless, for new worker, as any tool to him/her is a brand-new trial, the way he/she performs even a better result than experienced worker in use of some new tools.

The study explore the relationship between tool, performance and worker; in the following study, it will progressively broaden the scope to discuss the relationship between human—machinery—environment that tenders a profound investigation and research over each of variables such as every tool, operational link, etc; or it could be a research over mental measurement test of target tool, tool modeling design and detailed interpretation of usage and environment, or even the review, renew, R&D and design of the existing tool.

ACKNOWLEDGEMENTS

Thanks to Dr. Man-lai You for discussion and helpful criticism in the course of Special Topics on Human-Machine Systems Design.

REFERENCES

Boylesa, I.L., Yearouta, R.D. & Rysb, N.J. 2003. Ergonomic scissors for hairdressing. *International Journal of Industrial Ergonomics* 32: 199–207.

Greer, B., Jenkins, W. & Roberts, R. 1992. Carpal tunnel syndrome: a challenge for rehabilitation. *Journal of Rehabilitation* 58: 43–46.

Jackson, J., Kelly, B., Taylor, J. & Yearout, R. 1995. Carpal tunnel syndrome: case study of an intercollegiate athlete. *International Journal of Industrial Ergonomics* 15: 297–300.

Konz, S. & Johnson, S. 2000. *Work design: Industrial ergonomics*, Holcomb Hathaway Publishers, Scottsdale, AZ.

Konz, S. & Mital Anil. 1990. Guidelines carpal tunnel syndrome, *International Journal of Industrial Ergonomics* 5: 175–180.

Sanders, M. & McCormick, E. 1993. *Human factors in engineering and design*, 7th Ed, Mc Graw Hill Inc.

Storti, P.A. 1990. Getting a grasp on carpal tunnel syndrome. *Risk Management* 37: 40–48.

Zwicky, F. 1948. *Morphological method of analysis and construction*, Courant. Anniversary Volume. New York: Intersciences Publish: 461–470.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Relationship between performance-based and self-reported assessment of hand function in design of manageable packages

Karl W. Sandberg & Lena Lorentzen
Mid Sweden University, Sundsvall, Sweden

ABSTRACT: In the design of manageable packages for customers with hand function problems, we were faced with the problem of deciding which assessment tool to use. No studies have been done to show the relationship between performance-based and self-reported assessment of hand function to management of packaging. The purpose of the current study is to examine the relation between a self-reported hand function measure and performance-based measurements. The participants consisted of 40 adults who have impaired hand function. The Grip Ability Test (GAT) and Grippit test and self-reported assessment questioner were administrated. The findings of this study showed a slightly correlation between performance-based and self-reported assessment of hand function, therefore, further research is needed in this area.

Keywords: performance-based, self-report assessment, hand function, manageable packages

1 INTRODUCTION

The use of standardized tests has been well recognized in evaluating customers with upper-extremity or hand impairments because standardized tests provide accurate and objective information that can be used for treatment planning and determining treatment effectiveness (Rudman & Hannah 1998). The most common methodology used to evaluate self-assessment involves correlation analyses. A self-assessment score and a score based on some external measure or direct comparison of the absolute values of self-ratings and some external standard. Common standardized assessments do not consider the function of the hand in daily life (e. g. management packs), but instead the effect of artificial setting in which the testing occurs.

Questionnaires are a part of the evaluation of hand conditions and allow us to better understand what individual's experience (Sichuan et al. 2003, Michener & Legging 2001). Questionnaires are also easier to administer than physical measurements, for the reason that they require no special equipment, are less time consuming, and can be self-administered. To measure anatomic or physiologic impairments, many authors are of the view that physical measurements are the best (Amado 2001). There is a belief among practitioners that information gathered through self-report is subjective, less reliable, and less accurate than information obtained through standardized assessments. They concluded, however, that if a questionnaire is well designed and validated, the information collected from users through this method is also reliable (Michener & Leggin 2001).

2 OBJECTIVE AND SUBJECTIVE ASSESSMENTS

Several studies show the importance of either standardized assessments (Pap et al. 2003, Desrosiers et al. 1995) or self-report assessments (Amadio 2001, Barbier et al. 2001, Pap et al. 2003, Schuind et al. 2003).

However, only one study correlated the findings between the two objective measures; Sollerman Hand Function Test (Sollerman & Ejeskar 1995) and Sequential Occupational Dexterity Assessment (van Lankveld et al. 1996) with two subjective measures (Visual Analogy Scales for Pain and the Upper Limb Health Assessment Questionnaire (Bruce & Fries 2005) for individuals with rheumatoid arthritis.

Fitzgerald et al. (2000) make a distinction between the conception of self-assessment as an interindividual process or as an intraindividual process. The study of self-assessment according to the common methodological paradigm is an interindividual process. The subject asks, 'How easily can I open a package?' and generates a self-rating based on her assessment of her ability relative to others or to some ideal standard.

The purpose of the current study is to examine the relation between a self-reported hand function measure and performance-based measurements. We try to find new ways to assess and develop methods associated with design of manageable packages.

2.1 *Research design*

The research questions are: Does a relationship exist between objective assessment of hand function and subjective assessment of hand function to management packaging? We have developed an analysis model to compare performance-based and self-reported assessment of hand function of the participants' ability to manage packaging.

2.2 *Method*

Participants. The participants consisted of 40 adults who have impaired hand function. The participants were members of The Swedish Rheumatism Association, a non-profit organisation for people with rheumatic diseases.

Instruments. We used the Grip Ability Test (GAT) and Grippit test and self-reported assessment questioner. The Grip Ability Test (GAT) and Grippit are two tests used to determine hand functionality and strength in individuals with impaired hand skills. GAT consists of three tests designed to simulate daily activities: to pour a glass of water from a jug, putting on a glove and put a paperclip on a piece of paper. This is carried out on time, as a measure of the functionality of hands. Grippit test is used to measure the strength of the hands of one person. It is a machine with a power meter that sits in a handle. The result is given in N (Newton). For men tend to be on average approximately 400 N for women and about 200 N. A questionnaire was used to measure how participants perceive their disabilities in the hands, how they perceive their ability to use different hand and finger grip, and how they perceive to manage a reference pack. A 4-point and 9-point rating system are used for the participants to indicate how easy or difficulty for her to perform (Unicum 2008).

Procedure. First, the GAT and Grippit were administered to each participant, second, participant assess the degree of hand impairments in both hands, in a 4 point rating (from normal to severe impairments); third, participants asses their ability to use different grips, in a 9 point rating (from impossible to very easy); finally, how easy or difficult it was to manage packaging, in 9-point rating (from very difficult to very easy).

Table 1. Characteristics of performance-based and self-reported assessment instrument.

Instrument	Measures	Outcome type
Grippit	Strength	Objective in strength
Grip Ability Test (GAT)	Functionality	Objective time
Assessment of hand function (A-hand)	Experience	Subjective rating
Assessment using the grip (A-grip)	Experience	Subjective rating
Assessment in the management of packaging (A-manage)	Experience	Subjective rating

2.3 *Data analysis*

Descriptive statistics were calculated to determine the characteristics of the participants. Next, Spearman correlations were conducted to examine the relationships between the GAT, Grippit and subjective assessment ratings of the hand function to management packages.

2.4 *Results*

Thirty five women and five men (n = 40) participated in the study. Table 2 summaries the age of participants', age group, where the half of the participants falls in the 60–69 years old.

The participant's diagnosis is presented in Table 3, which half are diagnosed as Rheumatoid Arthritis.

No significant effects of left and right hand, age and diagnoses were obtained.

2.5 *Intercorrelation matrix*

The intercorrelation matrix for the research variables are shown in Table 4. The correlation analyses shown in Table 4, revealed that; A-hand is significant positively correlated with GAT ($p < 0.01$), A-grip are significant positively correlated with Grippit ($p < 0.05$) and A-manage ($p < 0.01$), A-grip are negatively correlated with GAT and A-hand

Table 2. Participants' age.

Age (years)	Number
25–39	4
40–49	2
50–59	6
60–69	20
>70	8

Table 3. Diagnoses of the participants.

Diagnoses	Number
Rheumatoid arthritis	20
Arthritis	6
Scleroderma	4
Guilliain barre syndrome	4
Parkinson's disease	1
Psoriasis arthritis	2
Rheumatoid arthritis and multiple sclerosis	2
Whiplash	1

Table 4. Correlations between performance-based (Grippit, GAT) and self-reported assessment (A-hand, A-grip, A-manage) measurement of hand function.

Constructs	1	2	3	4	5
Grippit	1.00				
GAT	–0.27	1.00			
A-hand	–0.24	0.44**	1.00		
A-grip	0.32*	–0.37*	–0.39*	1.00	
A-manage	0.25	–0.01	–0.01	0.42**	1.00

Note: *p < 0.05; **p < 0.01.

(p < 0.05). The results confirm earlier findings that there are few correlation studies that find a correlation between the performance-based and self-reported assessment.

3 CONCLUSION

The findings of this study showed a moderate correlation between a participants self-report assessment and performance-based measurement of hand function.

Results from present study give no promising assessment tool that designers can use as a valid measure to quickly and adequately evaluate a customer's hand function to design manageable packages. This tool may be helpful to many designers to better manage the time spent with their customers. Self-report questionnaires need to be short enough for easy completion and review and must contain items that will evaluate a client's functional limitations (Michener & Leggin 2001).

A limitation of the current study is that the severity of the participants' hand impairment was not categorised. Additional data from a larger sample size is needed to clarify whether a stronger correlation exists between the hand function assessments.

The information gained from the performance-based and self-reported assessment measurement is valuable. Both measures offer information on a participation's ability to function and complete tasks.

At the same time, the two assessments offer important as well as different additional information about the participation. The self-report provides insight into the participation's perception of her ability to function, whereas the performance-based measurement provides information on the speed and strength of performance.

Despite the theoretical value of self-assessment, the traditional measures could lead to the conclusion that self-assessment ability is poor. However, problems inherent in the traditional approaches for measuring self-assessment call into question this verdict on self-assessment. We have offered one alternative framework for research in this domain that conceptualizes self-assessment as an intraindividual as opposed to interindividual process. This new framework faces its own limitations, but it exemplifies the potential that exists to take the study of self-assessment in new directions.

ACKNOWLEDGMENTS

We wish to thank the Unicum (Nordic Design for All Center) who have funded the study and Mid Sweden Packing which will fund a journey to the 9th Pan-Pacific Conference on Ergonomics (PPCOE 2010), with theme and scope: Ergonomics for All, 7–10 November, 2010 in Taiwan. We also extend our thanks to the people who participated in the research.

REFERENCES

Amadio, P.C. 2001. Outcome assessment in hand surgery and hand therapy: an update. *Journal of Hand Therapy* 14: 63–67.

Barbier, O., Penta, M. & Thonnard, J. 2003. Outcome evaluation of the hand and wrist according to the international classification of functioning, disability, and health. *Hand Clinics* 19: 371–378.

Bruce, B. & Fries, J.F. 2005. The health assessment questionnaire. *Clinical and Experimental Rheumatology*, 23 (Suppl. 39), 14–18.

Desrosiers, J., Hebert, R., Bravo, G. & Dutil, E. 1995. Upper Extremity Performance Test for the Elderly (TEMPA): normative data and correlates with sensorimotor parameters. *Archives of Physical Medicine and Rehabilitation* 76: 1125–1129.

Fitzgerald, J.T., Gruppen, L.D. & White, C.B. 2000. The influence of task formats on the accuracy of medical students' self-assessments. *Academic Medicine* 75: 737–741.

Michener, L.A. & Leggin, B.G. 2001. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *Journal of Hand Therapy* 14: 68–76.

Pap, G., Angst, F., Herren, D., Schwyzer, H.K. & Simmen, B.R. 2003. Evaluation of wrist and hand handicap and postoperative outcome in rheumatoid arthritis. *Hand Clinics* 19: 471–481.
Richards, L., Stoker-Yates, J., Pohl, P., Wallace, D. & Duncan, P. 2001. Reliability and validity of two tests of upper extremity motor function post-stroke. *Occupational Therapy Journal of Research* 21: 201–219.
Schuind, F.A., Mouraux, D., Robert, C., Brassinne, E., Remy, P., Salvia, P., et al. 2003. Functional and outcome evaluation of the hand and wrist. *Hand Clinics* 19(3): 361–369.
Sollerman, C. & Ejeskar. A. 1995. Sollerman hand function test: A standardised method and its use in tetraplegic patients. *Scandinavian Journal of Plastic and Reconstructive Hand Surgery* 29: 167–176.
Unicum. 2008. *Self-assesment Questioner*. Unicum (Nordic Design for All Center), Bollnäs, Sweden.
van Lankveld, W., van't Pad Bosch, P., Bakker, J., Terwindt, S., Franssen, M. & van Riel, P. 1996. Sequential Occupational Dexterity Assessment (SODA): a new test to measure hand disability. *Journal of Hand Therapy* 9: 27–32.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Distraction affects the robot-assisted surgical performance depending on task difficulty

Shi-Hyun Park & Jung-Hung Chien
Nebraska Biomechanics Core Facility, University of Nebraska at Omaha, NE, USA

Dmitry Oleynikov
Department of Surgery, University of Nebraska Medical Center, NE, USA

Irene Suh & Ka-Chun Siu
College of Public Health, University of Nebraska Medical Center, NE, USA

ABSTRACT: This study investigated the impact of distractions on robotic surgical performance with different task difficulties. We hypothesized that distractions would negatively affect robot-assisted surgical performance in comparison to no distraction. In addition, the performance on distractions would be affected by the level of task difficulty. Sixteen subjects performed three inanimate surgical tasks: Bimanual Carrying (BC), Needle Passing (NP), and Suture Tying (ST) with three distraction groups (passive, active, combined distraction), and a control group with no distraction. Dependent variables were objective measurements (time, distance, speed, and smoothness). A two-way ANOVA was applied to examine the effect of distractions and tasks. Distractive environment significantly decreased performance (time to task completion ($p < 0.005$) and distance traveled ($p < 0.005$)), and the effect of distraction was significant on difficult tasks (NP and ST).

Keywords: active distraction, passive distraction, task difficulty, robot-assisted surgery

1 INSTRUCTIONS

1.1 *Distractions*

Distractions are known as a significant factor affecting human performance in the safety critical domain. Being distracted while performing a complex skill increases the complexity of a task and leads to performance impairment (Beyea 2007, Healey et al. 2006). However, it may not be just the presence or absence of a distraction that affects the performance. Instead, the characteristics of the distraction may play a critical role in affecting human performance (Szafranski et al. 2009).

1.2 *Distractive environment*

The operating room is a nosy and distractive environment. Many sources including patient monitors, suction machines, conversation during surgery, contribute to the distractive environment in the operating room. It is believed that such distractive environment could have a negative effect on surgical performance during delicate procedures such as laparoscopy (Goodell et al. 2006, Hsu et al. 2008). In addition, robot-assisted surgery is a complex task requiring a high level of precision and coordination of hand movement, and its complexity is further increased by distractions (Siu et al. 2010, Suh et al. 2009). Distraction in operating rooms not only can affect surgeon's performance, but also could be considered as one of the most important contributing for 40% to 50% factor of surgical errors (Sarker & Vincent 2005). Studies reported that many distractions were related to equipment, work environment, and procedural events (Healey et al. 2006). However, the effect of distractions on robot-assisted surgical performance with different levels of task difficulty has not been largely explored.

The purpose of this study was to investigate how distractions affect performance of fundamental surgical tasks. Also, we investigated the effect of distraction on different levels of task difficulty. We hypothesized that distraction would have a negative impact on robotic surgical performance, and the effect of distraction on surgical task performance could be driven by the nature of the task.

2 METHODS

We examined on surgical performance with the da Vinci™ Surgical System (dVSS; Intuitive

Surgical, Inc., Sunnyvale, CA, USA), which is commonly used for robot-assisted laparoscopic surgery.

2.1 *Subjects*

Sixteen medical students (22–24 years of age) were recruited to participate in this study. Novice users had no prior experience using the dVSS. All subjects had only basic surgical knowledge with no prior experience in robotic surgery and were right-handed. Informed consent was obtained from each subject before his or her participation in accordance with the Institution Review Board of the University of Nebraska Medical Center.

2.2 *Distractive conditions*

An active distraction (interactive distraction), a passive distraction (non-interactive distraction) and combined distraction (active and passive distraction) were used in this experiment. The active distraction group was required to respond to the distraction with decision making, and the distraction was composed of general science questions chosen from the Medical College Admission Test (MCAT) problem bank. A recorded series of general science questions was posed to subjects to solve and all answers were reported verbally. A set of pre-recorded noises (passive distraction) from an actual operating room (between 50 to 90 dB), no response required to the distraction, was used to mimic the noisy environment in the operating room during experiment. This recorded noise included patient monitors, suction machines, drill machines and conversations between surgeons and residents. Overall condition of the environment was the same for all subjects; no unusual amount of noise or other distractions.

2.3 *Tasks*

The following three inanimate robotic surgical tasks were performed in this study:

A. *Bimanual Carrying* (*BC*), a "pick and place" task: picking up five 15 × 2-mm rubber pieces from a 30-mm metal cap with the right and left instruments, respectively, and carrying them to the opposite caps simultaneously (Fig. 1a).
B. *Needle Passing* (*NP*), a "translational" task: passing a 26-mm surgical needle through six pairs of holes made on the surface of a latex tube (Fig. 1b).
C. *Suture Tying* (*ST*), a "precision navigation" task: passing a 150 × 0.5-mm surgical suture through a pair of holes made on the surface of a latex tube and making three knots using intracorporeal knot-ting (Fig. 1c).

All three tasks were designed to mimic real robotic surgical tasks, and to require consistent repetition of the same movements with bimanual coordination for quality performance. The participants were required to complete five BC, six NP, and one ST tasks for each trial.

2.4 *Procedures*

All participants were provided a 10 minute tour with a descriptive briefing of the dVSS, and performed ten practical trials for preventing learning within trials. Our previous studies have already established that ten trials are sufficient to learn a robotic surgical task (Judkins et al. 2009, Narazaki et al. 2006). A control group performing tasks without any distraction was included as a baseline. The order of tasks on inanimate material using the dVSS was fully randomized in each types of distraction.

2.5 *Objective measures*

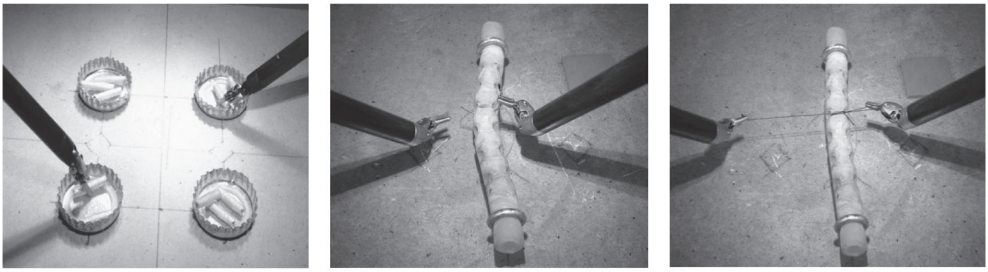
The objective performance measures were Time to Task Completion (TTC), total Distance traveled (D), Speed (S), and Curvature (κ). Time to task completion was the time required to complete a given surgical task. Start and end times were identified as the time when the instrument tips were within 1 cm of the starting positions. Total distance traveled was the sum of Euclidean distances between each time sample. Speed was calculated as the magnitude of the velocity. Curvature measured the straightness of the path and was calculated at each point on the path by the following equation (Gray 1997):

$$\kappa = \frac{|\dot{r} \times \ddot{r}|}{\dot{r}^3}$$

where $\dot{r}$_is the velocity of a point r on the three-dimensional path, and $\ddot{r}$ is the acceleration of point r. The median (κ_{med}) was computed for each trial. Values of κ_{med} close to zero indicate relatively smooth and straight movements, whereas larger values indicate curved and jerky movements. They were acquired using the dVSS Application Programmer's Interface provided by Intuitive Surgical, Inc (Sunnyvale, CA, USA). A custom program using LabView (National Instruments Corp., Austin, TX, USA) was written to interface to the dVSS via an Ethernet connection. Kinematic data was streamed at 100 Hz.

2.6 *Data analysis*

All data were post-processed by using MATLAB (MathWorks, Inc., Natick, MA, USA). A two-way ANOVA was performed to detect the effect



(A) (B) (C)

Figure 1. The tasks performed in the study using the da Vinci Surgical System (dVSS). (A) Bimanual carrying, (B) Needle passing, (C) Suture tying.

of the distractive conditions (passive, active, and no distraction) and the tasks (BC, NP, and ST). The post-hoc pair-wise comparison with Tukey's multiple comparison tests was performed when factors were significant. The significance level was set at $\alpha = 0.05$. The dependent variables were time to task completion, total distance traveled, speed, and curvature.

3 RESULTS

3.1 *Distraction effects*

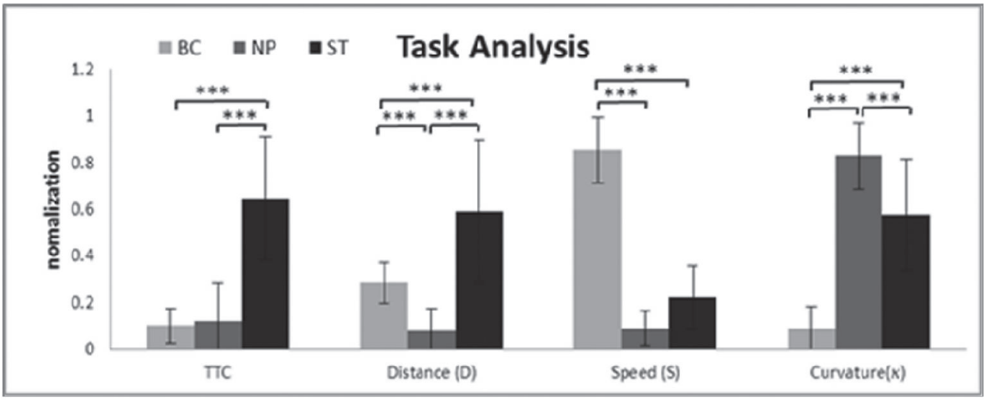
Significant distraction effects were found for temporal and kinematic dependent variables: Time to Task Completion (TTC) ($p < 0.005$), total Distance traveled (D) ($p < 0.005$) and average speed (S) ($p < 0.005$) (Fig. 2B). For the TTC, the follow-up pairwise comparisons indicated that the Bimanual Carrying (BC) task had no significant effect from all type of distractions. However, the Needle Passing (NP) task with the combined distraction showed significant increment of the completion time when compared with no distraction condition ($p < 0.005$). Also, the combined distraction on the NP significantly increased the time compared with the passive distraction. In terms of the Suture Tying (ST) task, the combined distraction condition showed significantly higher TTC compared to the no distraction ($p < 0.5$), the passive ($p < 0.005$), and the active distraction ($p < 0.005$). For the total distance traveled of the surgical instrument tips, only the ST task had significant effect with distractions. The active and combined distraction significantly increased the distance compared with the control group ($p < 0.005$). There were also significant differences between the passive and the active, between the passive and the combined, and between the active and the combined distractions ($p < 0.005$). However, there were no significant distractive effects on each task for the average speed and the curvature.

3.2 *Task analysis*

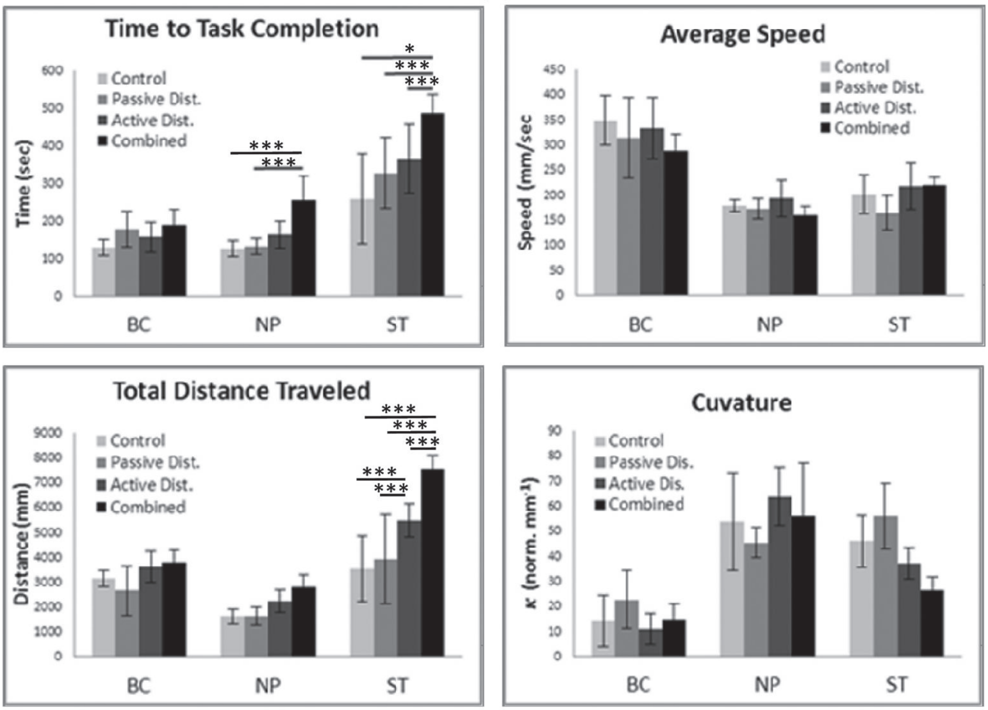
There were significant task effects to the different types of distraction for all kinematic dependent variables: TTC ($p < 0.005$), D ($p < 0.005$), S ($p < 0.005$), and curvature (κ) (Fig. 2A). The ST task showed significant higher time to task completion compared with the BC and the NP tasks ($p < 0.005$). The distance traveled was significantly different on all three tasks ($p < 0.005$) and the ST task required the longest distance. The speed of the BC task was significantly faster than other two tasks ($p < 0.005$), while there was no difference between the NP and the ST tasks. For the curvature, all three tasks had significant differences among each together. Especially, the BC task showed significant lower curvature. In other words, the motion of BC task had the most straightness and smoothness among the three tasks. Through these results, we confirmed that the suture tying task was the most difficult task with the greatest time to task completion, the longest distance traveled, and significantly lower speed and higher curvature (low motion smoothness) when compared with other tasks. The NP task was the second difficulty among three tasks with significantly lower speed and higher curvature compared to the control.

4 DISCUSSION

This study objectively demonstrated that the distractive environment has negative effects on performance. Our results supported our hypothesis, since we found that distractive environment affects performance during fundamental surgical tasks: a bimanual carrying, needle passing, and suture tying. In addition, the effect of distraction was different depending on the level of task difficulty. Active distraction (such as decision making) showed greater impact on performance as time to task completion and total traveled distance. These negative



(A)



(B)

Figure 2. (A) It shows the task effect on temporal and kinematic dependent variables: TTC: Total time to complete the task, Distance (D): total distance traveled of the surgical instrument tips, Speed (S): average movement velocity, Curvature (*k*): motion smoothness (lower value means more straight and smooth motion). (B) It shows the type of distraction effect and the result of pairwise comparison between the task and the distractions. (* $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$).

distraction effects are also supported by other studies, which showed that a distraction caused an attention shift and led to increased errors in performance of both cognitive (Pool 2003) and motor (Weerdesteyn 2003) tasks. Also, working memory comprises functional components of cognition that allow humans to comprehend to retain information and to solve problems (Baddeley 1996). It is noted that both active (decision making) and combined (active and passive) distractions showed more detrimental distraction effect than passive (auditory) distraction. These results suggested that the cognitive distraction was competing with the robotic surgical task for cognitive resources.

In terms of the nature of tasks, the bimanual carrying task, which was defined as the lowest difficulty task in the task analysis had no impact from all distractive condition, while the suture tying task determined as the most difficult task had negative effects on temporal performance and movement distance from all kinds of distraction. The needle passing task having the medium level of difficulty showed a degraded temporal performance only when distractions were present. Therefore, performing high level of difficulty task could be apt to have a negative impact from distractive environments. On the other hands, distractive effect could be ignored in performing low level of difficulty task. Therefore, more challenging tasks or complex tasks, such as suture tying, would have a larger negative effect of distraction on surgical performance. Future studies are required to investigate how surgical performance can be enhanced in the presence of distractions.

5 CONCLUSION

Results of this study represent one of the few efforts to investigate the distractive effect of surgical performance in a robot-assisted surgery. A positive attribute of this study was the ability to quantify both distraction effect and the level of task difficulty during performing fundamental surgical tasks. In addition, this study investigated and demonstrated the interaction between distractions and different task difficulties. One weakness was that the metrics used in the study was not sufficient to fully prove the effect of distractions and tasks, since the speed and the curvature were not able to show differences, although the result did show significance on the temporal performance and the movement distance. Thus, we need to develop more sensitive metrics, such as muscle activity from electromyography and error rate, for measuring human performance.

In conclusion, the study provided several insights that should be considered by an operator as part of a surgeon working in a distractive environment. It was found that the distractive effect degrading performance may vary among tasks having different levels of difficulty and complexity. Based on these findings, additional investigations are needed to find out diverse methods to reduce the negative impact from distractions, such as augment feedback information to assist an operator's decision making and cognition. Finally, the relationship between the distraction and the task difficulty should be considered when evaluating the work environment and safety for preventing human errors.

REFERENCES

Baddeley, A.D. 1996. Exploring the central executive. *Quarterly Journal of Experimental Psychology*. 49 A: 5–28.

Beyea, S.C. 2007. Noise: a distraction, interruption, and safety hazard. *AORN J*. 86(2): 281–5.

Campbell, M.J. & Gardner, M.J. 1988. Calculating confidence intervals for some non-parametric analyses. Br Med J (Clin Res Ed). 296(6634): 1454–1456.

Goodell, K.H., Cao, C.G. & Schwaitzberg, S.D. 2006. Effects of cognitive distraction on performance of laparoscopic surgical tasks. *J Laparoendosc Adv Surg Tech. 1*6: 94–98.

Gray, A. 1997. *Modern differential geometry of curves and surfaces with Mathematica*. Boca Raton, FL: CRC Press LLC.

Healey, A.N., Sevdalis, N. & Vincent, C.A. 2006. Measuring intraoperative interference from distraction and interruption observed in the operating theatre. *Ergonomics*, 49: 589–604.

Hsu, K.E., Man, F.Y., Gizick, R.A., Feldman, L.S. & Fried, G.M. 2008. Experienced surgeons can do more than one thing at a time: Effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. *Surg Endo*. 22(1): 196–201.

Judkins, T.N., Oleynikov, D. & Stergiou, N. 2009. Objective evaluation of expert and novice performance during robotic surgical training tasks. *Surg Endosc*. 23(3): 590–7.

Narazaki, K., Oleynikov, D. & Stergiou, N. 2006. Robotic surgery training and performance: identifying objective variables for quantifying the extent of proficiency. *Surg Endosc.* 20: 96–103.

Pool, M., Koolstra, C. & Voort, T. 2003. Distraction effects of background soap operas on homework performance: An experimental study enriched with observational data. *Educ Psychol*. 23: 361–380.

Sarker, S.K., & Vincent, C. 2005. Errors in surgery. *Int J Surg*. 3(1): 75–81.

Siu, K.C., Suh, I.H., Mukherjee, M., Oleynikov, D. & Stergiou, N. 2010, The impact of environmental noise on robot-assisted laparoscopic surgical performance. *Surgery.* 147(1): 107–113.

Suh, I.H., Siu, K.C., Mukherjee, M., Oleynikov, D. & Stergiou, N. 2009. The effect of distraction on robot-assisted surgical performance. *J Am Coll Surg.* 209(3): 107–112.

Szafranski, C., Kahol, K., Ghaemmaghami, V., Smith, M. & Ferrara, J.J. 2009. Distractions and surgical proficiency: an educational perspective. *The American Journal of Surgery.* 198: 804–810.

Weerdesteyn, V., Schillings, A., Galen, G., & Duysens, J. 2003. Distraction affects the performance of obstacle avoidance during walking. *J Motor Behavior*. 35: 53–63.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Development of a knowledge-based system based on Collaborative-Filtering recommendation for training knowledge sharing

Shu-Ling Wang & Wei-Ham Hung
Department of Information Management, National Taichung Institute of Technology, Taichung, Taiwan

ABSTRACT: Knowledge sharing has been an important issue to the outcomes of knowledge management. However, traditional knowledge management systems only focus on seeking how to store the knowledge, as documents, which has been retained by enterprises while ignoring the core advantages of knowledge sharing such knowledge as the experiences, technical capabilities and organization culture among employees. Therefore, how to support employees retrieve what knowledge they require for improving their work efficiency is an important issue. This study aims to developing a knowledge-based system for training knowledge sharing and management. Firstly, we apply Case-Based Reasoning (CBR) approach to support enterprises in obtaining and managing case knowledge, and promote knowledge sharing and accumulation by developing personal services and interactive functions based on Web 2.0. Especially, the system allows the knowledge to be more effectively communicated, shared and exchanged among users by integrating the Collaborative-Filtering recommender, which automatically produces and recommends knowledge-based on users' preferences and requirements.

Keywords: case-based reasoning (CBR), knowledge management (KM), knowledge sharing, Collaborative-Filtering recommendation, Web 2.0

1 INTRODUCTION

One of the essential aspects of virtual enterprising is determining how to integrate the processes, resources and knowledge from different local branches and enterprises through enterprise alliances, in order to quickly respond to customer requirements (Zhen, L. et al. 2010). However, traditional knowledge management systems only focus on seeking how to store the knowledge, as documents, which has been retained by enterprises while ignoring the core advantages of knowledge sharing such knowledge as the experiences, technical capabilities and organization culture among employees. Therefore, how to support employees retrieve what knowledge they require for improving their work efficiency is a important issue.

Although such tacit knowledge is usually accumulated by employees themselves cannot be easily reproduced and transferred, it is a major challenge for enterprises to enhance their core competitiveness while trying to circumvent the loss of this knowledge base through personnel resignation and transfer. Therefore, if knowledge can be obtained and organized by effective means, it can become the driving force for improving enterprise competitiveness and increasing value (Nonaka, I. 1994).

This study aims to develop a "collaborative knowledge-based system" for knowledge sharing and management; this requires a knowledge management system that integrates the Particle Swarm Optimization, K-mans algorithm and Case-Based Reasoning to support enterprises in obtaining and managing case knowledge, and promotes knowledge sharing and accumulation by developing personal services and interactive functions based on Web 2.0. The knowledge scattered across departments or enterprises is effectively integrated by the platform, making the users of an enterprise able to obtain knowledge related to their tasks. The platform allows the knowledge to be more effectively communicated, shared and exchanged in the enterprise by integrating the Collaborative-Filtering recommender, which automatically produces and recommends knowledge-based on users' preferences and requirements.

2 RELATED WORK

2.1 *Collaborative-Filtering recommendation*

The concept of the Collaborative-Filtering Recommender is mainly based on the hypothesis that a user group with identical interests may be found by means, and items that might interest such users can be further recommended by analyzing the preferences of such a group (Breese, J.S. et al. 1998). Items are recommended to target users for use from the group's perspective, and the similarity of user preferences is calculated by historical records of past users, to find the nearest neighbors whose preferences are most similar; information is also given to target users for reference through the opinions of the nearest neighbors. The Collaborative-Filtering recommender has been successful applied in various fields; a few examples are provided below. For instance: TAPESTRY is an experimental mail system developed at the Xerox PARC for solving the increased use of electronic mail. Users' preferences determine the type of email filtering of the system (Goldberg, D. et al. 1992). Usenet News is a recommender system for new for users interested in content that provides rating, and then calculates the rating data to generate recommendation lists (Konstan, J.A. et al. 1997). Other similar recommender systems include movies, music (Li, Q. et al. 2007) and e-learning (Bobadilla, J. et al. 2009) (Hsu, M.H. 2008). Applying the recommender system to enterprises allows members to interact for knowledge sharing and management (Zhen, L. et al. 2010).

While studies have reported that Collaborative-Filtering may be the most successful recommendation technology up to now, it still harbors two limitations, as follows (Burke, R. et al. 2002) (Cho, Y.H. et al. 2004).

2.1.1 *Sparsity problem*

The number of items that need to be assessed is so great that the matrix of the assessment materials is too sparse to produce a proper recommendation item to the user through similarity calculation.

2.1.2 *Scalability problem*

Because the recommendation system calculates the similarity between users by the Nearest Neighbors algorithm, which will, in time, become ever more complicated as a square of equal ratio to the historical use records of users and the number of items to be assessed, and when users and the recommended items gradually increase in number, system scalability will become problem affecting the efficiency of the recommendation system.

In order to improve the problem of sparsity and scalability, scholars have proposed different technological solutions, such as Hybrid Recommender System (Burke, R. 2002) and Clustering technology (Rashid, A.M. et al. 2006) (Sarwar, B.M. et al. 2002). Other researchers have pointed out that using clustering techniques can indeed lead to good recommendations. The applications of clustering techniques reduce the sparsity and improve the scalability (Rashid, A.M. et al. 2006). Therefore, this study adopts the method that K-means algorithm. First of all, large amounts of data will be clustered in advance and user clusters of identical preference will be defined once clusters are created that reduce the sparsity and improve scalability of recommender system.

2.2 *Knowledge-based system*

The primary purpose of knowledge-based systems is to collect knowledge from experts and then systematically present or modularize the knowledge. Having systematic representations and modules allows computers to perform inferences and problem solving on various fields. If the steps in a problem solving process or the logics and reasoning under an inference can be formalized, then knowledge-based systems can be brought to resolve complicated problems that otherwise must be solved by human experts.

According to the Computer user High-tech Dictionary, knowledge-based system is a computer system that is programmed to imitate human problem-solving by means of artificial intelligence and reference to a database of knowledge on a particular subject (Huang, H.C. 2009). Knowledge-based systems typically contain four core components: 1. knowledge base, 2. inference mechanism, 3. knowledge engineering tools, 4. users interface (Liao, S.H. 2005).

Knowledge base is constructed through explicit knowledge representation method supplied and built by domain experts. As knowledge-based systems extensively refer to knowledge and experiences of domain experts when they are searching for solutions, knowledge-based systems are very suitable to applications of unstructured and uncertain knowledge. Example applications of knowledge-based systems are as follows: KBDSS is used to assist doctors in analyzing factors presented in the working environment that cause musculoskeletal disorder, especially common causes that lead to shoulder and neck pain (Padma, T. et al. 2009). BSCKBS analyzes business data and assist users to make decisions through knowledge-based system. It also helps construct specific business strategy and marketing plan by deploying BSC (Gasparetti, F. et al. 2009). Business letter generator deploys both web-based training and self-directed learning. It assists new employees to write business and provides trainings on business writing (Huang, H.C. et al. 2009).

As discussed previously, a complete knowledge-based system will contain four core components. Every component can affect the system's performance. Hence, this study utilizes CBR and collaborative recommendation algorithm to develop a smart knowledge-based system and smart user interface.

Use enterprises that provide knowledge base services as an example. The purpose is to build a knowledge-based system that assists users to effectively search and use their company's training knowledge. The methods and algorithms for developing the system are explained in the following chapter.

3 THE METHOD OF SYSTEM DEVELOPMENT

3.1 *System framework*

This study integrates Case-Based Reasoning, Knowledge Management method and Particle Swarm Optimization to obtain and integrate the current knowledge resources of virtual enterprises, based on the concept of spiral of knowledge (Nonaka, I. 1994), to analyze the functional structures that might be needed in the four processes: socializing, externalizing, combining and internalizing tacit and explicit knowledge. In order to improve the knowledge sharing and processes within an enterprise, information technology is applied to develop intelligent knowledge sharing and management platform architecture and further formulate a personalization system for knowledge recommendation by employing a Collaborative-Filtering recommender. The architecture of an intelligent knowledge sharing and management platform proposed in the study includes: (1) the case knowledge of representation and digitalizing. (2) A case knowledge sharing platform based on Web 2.0. (3) Personal knowledge management portal. (4) Knowledge Q & A community. The relevant descriptions are as follows:

3.1.1 *Case-based knowledge of representation and digitalizing*

In this study, enterprise knowledge is represented as cases, such as engineering knowledge, R&D documents, training textbooks, etc. A case knowledge base is intended to be built by interviewing experts and acquiring virtual enterprise knowledge.

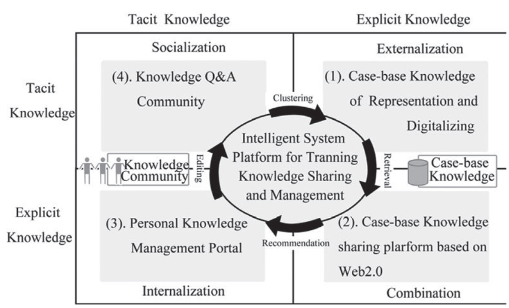


Figure 1. The architecture of intelligent system platform for training knowledge sharing and management.

3.1.2 *Case-based knowledge-sharing platform based on Web 2.0*

Users can share their experiences, review case knowledge and retrieve case knowledge on the platform, which further recommends knowledge to workers through its recommendation function, to enable the knowledge to flow in the organization and achieve interactive, sharing and exchange effects.

3.1.3 *Personal knowledge management portal*

The information needed by individual workers is managed by a single portal. Users can choose the knowledge they need according to their preference or task goals, facilitating users to quickly acquire the internal knowledge resources and improve the work efficiency via the personal knowledge management portal.

3.1.4 *Knowledge Q & A community*

With many experts in various fields participating in the community, users can interact by finding experts or colleague in relevant fields, asking questions and discussing, thus promoting communication of tacit knowledge and further sharing knowledge.

3.2 *System function*

This study aims to build a framework for users to first retrieve case knowledge through combined Case-Based Reasoning, Particle Swarm Optimization and K-means algorithm, thereby obtaining the relevant knowledge they need; the structure is also capable (based on users' preferences on the case knowledge and further integrated with the Collaborative-Filtering recommender) of producing a list of relevant recommended knowledge elements that users might prefer if they had never been previously retrieved. The modules of the platform are described in detail as follows:

3.2.1 *Case-based knowledge module*

This module contains the users' database and case knowledge database, which respectively record the personal data of users and the digitalized documents as stated in the above "Case knowledge of representation and digitalizing" section, which serve as the data recommended to be input into the module.

3.2.2 *Case-based knowledge recommendation module*

This section integrates case knowledge sharing, management and the Collaborative-Filtering recommender, to construct the knowledge recommendation module. A personal recommendation system is

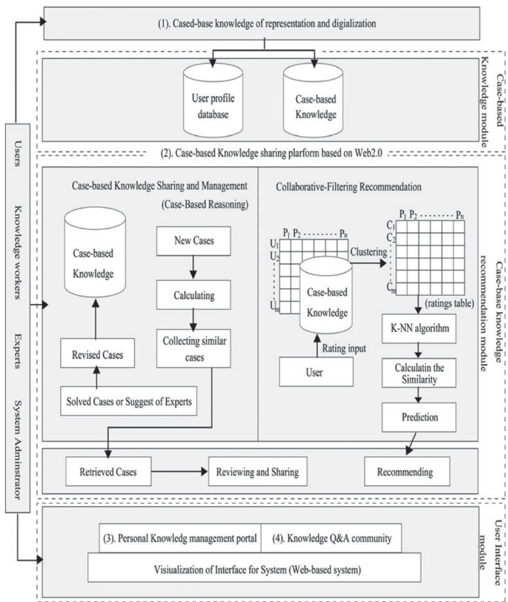


Figure 2. The architecture of system function.

provided by applying the Collaborative-Filtering recommender and the opinions of uses that retrieve similar cases found by CBR to retrieve information which was not previously read by such readers but which might possibly match their preferences.

3.2.3 *User interface module*

Users can quickly obtain and master the knowledge they need by their preferences or task goals via the single interface. A knowledge Q & A community is the system by which users can interact by asking questions, discussing and finding experts in relevant fields for advice, and strengthening the value and practices of knowledge exchanges by topic discussion.

4 SYSTEM METHOD

The aim is to solve the sparse and scalability problems (Cho, Y.H. et al. 2004) to enhance the efficiency and prediction accuracy of the Collaborative-Filtering recommender. This study proposes combining Particle Swarm Optimization, K-means algorithm and the Collaborative-Filtering recommender; the former is used for clustering, according to user preference, with the specific cluster center as index. When sparse problems occur, the prediction accuracy of the traditional Collaborative-Filtering recommender is improved by the clustering method. The results of clustering in advance can also solve the system scalability problem (Roh, T.H. et al. 2003) (Sarwar, B. et al. 2001) (Sarwar, B.M. 2001) (Ungar, L.H. et al. 1998) caused by too large a data load. The procedure of this method can be summarized as follows:

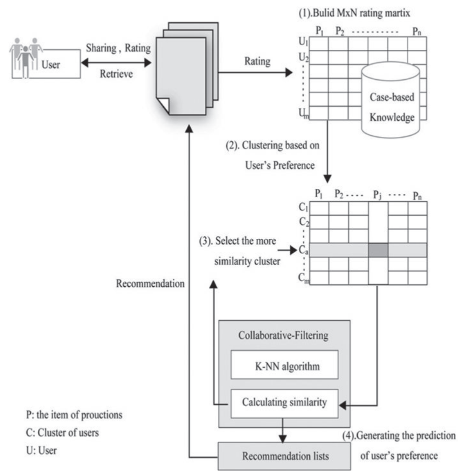


Figure 3. A simple architecture of the method that combine cluster and Collaborative-Filtering recommender.

4.1.1 *Build $M \times N$ rating matrix table*

This is used for recording the past transaction records of users and locating other users whose behaviors are similar, who will be taken as the nearest cluster. The preference of the nearest cluster is also analyzed and used as the basis for recommendation.

4.1.2 *Clustering based on user preference*

The users' preference is clustered using PSO combined with K-means algorithm (Shelokar, P.S. et al. 2007); and large amounts of data will be clustered in advance and user clusters of identical preference will be defined.

4.1.3 *Calculate the similar clustering*

Use the K-NN algorithm (Roh, T.H. et al. 2003) to find user preferences and interests in the most similar clusters.

4.1.4 *Generating recommendation lists*

The similarity between the target user and nearest user is calculated using Pearson's correlation coefficient (Rashid, A.M. et al. 2006) as the weight for predicting recommendations. In the clusters obtained by assessing target users and using the K-NN algorithms, similarity is calculated by taking preference value between two cases into the similarity equation (Rashid, A.M. et al. 2006).

5 INTELLIGENT USER INTERFACE

5.1 *CBR retrieval interface*

The smart user interface of knowledge base constructed provides case search and retrieval interface. The system performs inference based on case's related attributes. Users can search and retrieve knowledge of related cases by providing different case attributes. With the help of knowledge retrieval, the process of obtaining knowledge teaching materials becomes more efficient. The functionalities of this system include: 1. Generate the best recommendation by calculating similarities among attributes of cases that were initially input by users. 2. Present results in the user interface. 3. Users can redefine case attributes in order to obtain different retrieval result.

5.2 *Personal recommendation interface*

Traditional CBR is done by finding solutions to similar cases that happened previously. However, when users are facing new problems, this method cannot completely satisfy the required information. Since there is a lot more knowledge hidden in various filed knowledge cases, the system must actively recommend it to users and provide users with more opportunities to discover hidden knowledge. Therefore, this study further integrates CBS with Collaborative-Filtering recommendation mechanism. The system finds knowledge of similar cases through case-based reasoning and assesses

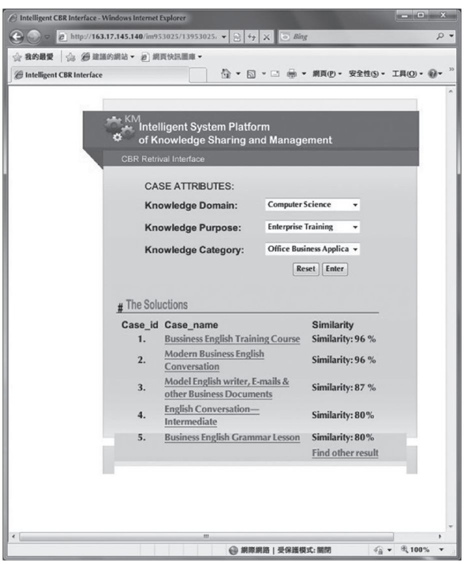


Figure 4–1. According to the case attributes to reasoning the similarity case.

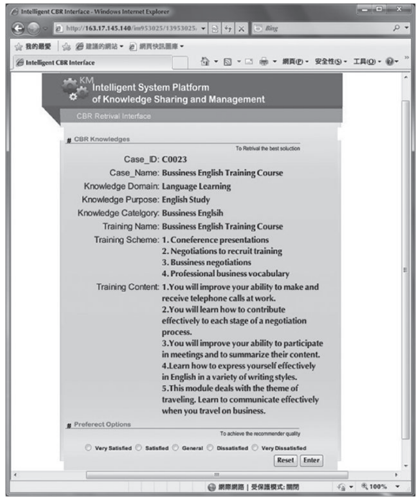


Figure 4–2. The similarity case content.

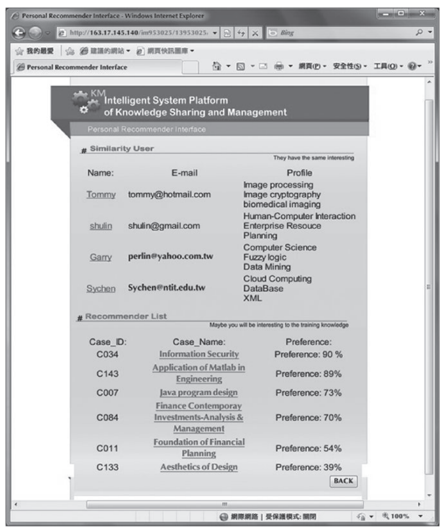


Figure 5. Personal recommender interface.

the result by receiving feedbacks given by users after they have finished reading the result. The system keeps cumulative such data and analyzes user's preferences for active recommendation. Overall, the system achieves personal knowledge recommendation and facilitates the flow, sharing and transmission of knowledge.

6 CONCLUSION

This study uses CBS as the foundation and integrates Collaborative-Filtering recommendation with Particle Swarm Optimization (PSO) to develop knowledge sharing and smart personal knowledge recommendation interface. First, this study applies knowledge spiral framework to effectively perform systematic data organizing and classification. This study then constructs case knowledge base. By going through case-based reasoning retrieval interface and by using case attributes provided by users, the retrieved result will be more accurate. The personal recommendation interface enables the system to possess smart learning mechanism. The system actively recommends according to users' preferences. The system facilitates the flow, sharing and transmission of knowledge and increase the performance of knowledge management.

REFERENCES

Breese, J.S., Heckerman, D. & Kadie, C. 1998. Empirical analysis of predictive algorithm for Collaborative-Filtering. *In Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence.* Madison, WI.

Bobadilla, J., Serradilla, F. & Hernando, A. 2009. Collaborative-Filtering adapted to recommender system of e-learning. *Knowledge-based Systems* 22: 261–265.

Burke, R. 2002. Hybrid recommender systems: Survey and experiments. *User Modeling User-Adapted Interaction* 12: 331–370.

Cho, Y.H. & Kim, J.K. 2004. Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert Systems with Applications* 26(2): 233–246.

Goldberg, D., Nichols, D., Oki, B.M. & Terry, D. 1992. Using Collaborative-Filtering weave an information tapestry. *Communications of the ACM* 35: 61–70.

Gasparetti, F., Micarelli, A. & Sciarrone, F. 2009. A Web-based training system for business letter writing. *Knowledge-Based Systems* 22: 287–291.

Hsu, M.H. 2008. A personalized English learning recommender system for ESL students. *Expert Systems with Applications* 34: 683–688.

Huang, H.C. 2009. Design a knowledge-based system for strategic planning: A balanced scorecard perspective. *Expert Systems with Application* 36: 209–218.

Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Goron, L.R. & Riedl, J. 1997. *Communication of the ACM* 40: 77–87.

Liao, S.H. 2005. Expert system methodologies and applications a decade review from 1995 to 2004. *Expert System with Applications* 28: 93–103.

Li, Q., Myaeng, S.H. & Kim, B.M. 2007. A probabilistic music recommender considering user opinions and audio features. *Information Processing and Management* 43: 473–487.

Nonaka, I. 1994. A Dynamic Theory of Organizational Knowledge Creation. *Organization science* 5(1): 14–37.

Padma, T. & Balasubramanie, P. 2009. Knowledge-based decision support system to assist work-related risk analysis in musculoskeletal disorder. *Knowledge-Based System* 22: 72–78.

Rashid, A.M., Lam, S.K., Karypis, G. & Riedl, J. 2006. ClustKNN: A highly scalable hybrid model & memory-based CF algorithm. *In: WebKDD.*

Roh, T.H., Oh, K.J. & Han, I. 2003. The Collaborative-Filtering recommendation based on SOM cluster-indexing CBR. *Expert Systems with Applications* 25: 413–423.

Sarwar, B.M., Karypis, G., Konstan, J. & Reidl, J. 2001. Item-based Collaborative-Filtering recommendation algorithms. *In: Proceedings of the 10th international conference on World Wide Web.*

Sarwar, B.M., Karypic, G., Konstan, J. & Riedl, J. 2002. Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering. *In: Proceedings of the Fifth International Conference on Computer and Information Technology.*

Shelokar, P.S., Siarry, P., Jayaraman, V.K. & Kulkarni, B.D. 2007. Particle swarm and ant colony algorithms hybridized for improved continuous optimization. *Applied Mathematics and Computation* 188: 129–142.

Ungar, L.H. & Foster, D.P. 1998. Clustering methods for Collaborative-Filtering. *In: workshop on Recommender Systems 15th National Conference on Artificial Intelligence.*

Zhen, L., Huang, G.Q. & Jiang, Z. 2010. An inner-enterprise knowledge recommender system. *Expert System with Applications* 37: 1703–1712.

Ergonomics for All: Celebrating PPCOE's 20 years of excellence – Lin & Chen (eds)
© 2011 Taylor & Francis Group, London, ISBN 978-0-415-58608-5

Managing the telematics use during drive: What does driver wants? A cross countries study

Fang Chen
Interaction Design, Department of Applied IT, Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT: It is a well-known knowledge that telematics use during drive can cause driver distraction and hence cause the accidents. Researchers and automotive manufactures developed a concept called "workload manager". It is intended to alert or control the use of telematics equipment during drive. But there is no general agreement and acceptance of how to do the control. The present study was focused on "what the driver wants" when regarding the tread-off balance between telematic use and drive safety, and specially focused on culture differences. It based on interview studies of the drivers from China and Sweden. The result showed very big culture differences. Further research work is needed in this area.

Keywords: drive distraction, workload management, drive safety, culture difference

1 INTRODUCTION

Driver distraction has been identified as one of the major contribution factors to the car accident (Dingus and Jermeland 2005). The distraction can also be considered as the driver being overloaded by different tasks and information on the road. Driver workloads are contributed by:

- Road the traffic conditions
- Secondary task performance
- Roadside presentations
- Different in-car presentations.

The telematics equipments use, such as mobile phones, GPD devices, mp3 players, etc., is the typical secondary task performance during drive. It is now a well-known knowledge that telematics use during drive can cause driver distraction and hence cause the accidents. Many research works have been published in this area (Lee, L. Ries et al. 2000; Hancock 2003; Green 2004; Anttila and Luoma 2005; Liu and Lee 2005).

Green summarizes the three major key points made in crash literature regarding telematics (Green 2004): 1) that there are crashes in which the use of telematics has been a contributing factor, 2) in the crashes where telematics use has contributed, the driver has become so involved in the in-vehicle task that they fail to give the necessary level of attention to the driving task, and 3) the crashes due to telematics use are more likely to occur during benign driving conditions.

But the impact from these studies to the real drive situation is very little. As the life cycle of the telemtics devices are short (about 6 months), people often bring in new devices and use them during drive. In the opposite, the vehicle's life cycle is much longer (at least 4 years). It is not easy for the automotive industrial to keep updating and take effective control of the telematic use during drive. The telematics devices are still not specially design for using during drive, and people brought in new devices and used during drive more and more often.

To solve this conflict, researchers and automotive manufactures developed a concept called "workload manager". According to Green (Green 2004), "a workload manager is a device that attempts to determine if a driver is overloaded or distracted, and if they are, alters the availability of telematics and the operation of warning systems".

Green (Green 2004) further on, categorizes workload managers into four broad categories depending on what they measure: 1) driver state, 2) driver input, 3) vehicle performance and response, and 4) the driving situation. The collection of data for triggering the workload management system is done by various sensors that measure a variety of human, automotive and driving parameters such as steering wheel angle, speed throttle angle, brake pressure, acceleration, lane position, headway, etc. The control is taking place by limiting access to functionality or warning against device use accordingly. For instance, incoming telephone calls can be diverted if the system deems the driving situation critical (i.e. if it assesses that it would be dangerous for the driver to answer the call at the time).

Chen (2008) proposed the zonal adaptive workload management system. In this system, the drive workload was evaluated according to the external road and traffic condition, rather than the driver's performance.

A key issue associated with workload management is how, at any given moment, to evaluate the level of access that drivers should have to the range of functionality available.

There are three possible ways to "alter the availability of telematics". One is totally blocking the availability of the devices (total control system). Another one is to give advices to the driver of the potential danger; it is up to the driver to decide if they shall use the devices or not (advice system). The third solution is something in between. What is the best way to do? There is not a common agreement among users and automotive industry.

There are some studies about how culture differences can affect the human interface design and human interaction. Marcus (Marcus 2003) has argued that many different aspects needs to be considered, such as culture dimension and cognitive difference, metaphor, mental model, navigation, interaction, appearance and presentation. There are some studies regarding the culture issue in drive and interaction design (Xie and Parker 2002; Zhang 2006; Chen 2009; Chen 2009; Jordan 2009). All of these studies have implied that the interaction design for in-vehicle systems that fits one country may not be suitable to another country, especially, the design results from Western studies are not suitable for Asian. There are very limited reports that can be found about drive distraction and telematics use in China and there is no report about how the workload management system shall be designed for Chinese drivers.

The purpose of present study was focused on "what the driver wants" when regarding the tread-off balance between telematic use and drive safety. More specific, the comparison between Swedish driver and Chinese driver was investigated.

2 METHOD

The study is carried out in both China and Sweden.

There are 18 Swedish and 42 Chinese drivers participated into the study. All of them have over 5 years drive experiences and used different telematics devices during drive.

The study was based on structured interview. In the beginning of the interview, the basic knowledge of telematics devices use and drive distraction was explained to the subjects, and made sure that they understood it perfectly. Then we explain to the subject that there are three possible control solutions to reduce the possible risk of using these devices during drive.

Alternative I: *A total control system (on-off)*: By this system, it is totally forbidden of using telematics devices during drive.

Alternative II: *Drive behaviour system*: According to the drive performance. This system is attempting to determine if a driver is overloaded or distracted through different drive parameters. If so, alter the availability of telematics and the operation of warning systems.

Alternative III: *Advices system*: This system is essentially location-based system which try to take into account the road and driving conditions. This system adjusts and limits the performance of secondary task based on the needs and complexity of secondary task performance. The system will only provide some advices to the driver of the potential danger of using telematics devices during driving. It is up to the driver, if he would follow the advice or not.

After the introduction, the subject was asked if he or she understood it perfectly? Then we asked them to select one system that they prefer the best and can consider implementing it in his (her) car. We even asked then to describe the advantages and disadvantages of each system to make sure that they understand the differences of the three systems.

After a general question, we went on with a lot of specific traffic and drive situations and asked the drive during that moment, which system they prefer. There are three independent variables in this special interview: The complexity of the tasks that the driver may perform during driving, different types of roads and different traffic situations.

We selected 17 tasks that most drivers may perform during drive.

The combination of road type and traffic conditions can be found in Table 1. We categorized the road condition into three levels. In each level, the difficulties in drive are similar. The numbers in Table 1 indicated the levels in this category.

Level 1: It is easy to drive.

Level 2: It is relatively easy to drive, but some extra attention is needed compare to level 1.

Level 3: It is not easy to drive, therefore, most of the driver's attention shall be on driving.

Under each road and traffic situation, the subjects is asked to select what kind of system they

Table 1. The road type and traffic conditions.

Road type	Traffics
Highway	Busy traffic (2); Difficult areas (2); Easy traffic (1)
City traffic	Busy traffic on big road (no pedestrians) (3); Not busy traffic in big road (2); Busy traffic and smaller roads (3); No busy traffic with smaller roads (3)
Country road	Busy traffic on big road (no pedestrians) (2); Not busy traffic in big road (1); Busy traffic and smaller roads (many pedestrians) (3); No busy traffic with smaller roads (3); Passing small villages (3)

Table 2. Shows the detail tasks that are included in each level.

Task level	Tasks included
1	Looking for traffic information; Weather forecast; Special scheduler; Special information; Listen to radio only; Listen to music but do not do anything;
2	Change TV channel; Recording voice communication; Change channels of radio; Modulate the radio; Remind the impartment date and time.
3	Dial number (searching numbers the installed in phone book); Receive a call; SMS—send and receive; Check personal routine schedule; Looking for a special song in a long list; Watch car TV

prefer to have (Alternative I, II, or III) when they perform each of the 17 tasks.

Among the 17 tasks, we also divided it into three levels (See Table 2) according to how difficult it can be to perform:

Level 1: Easy to perform, which probably only need one or two button to push. It does not demand long time visual attention to the task.

Level 2: Relatively easy to perform, but slightly more complicated than level 1.

Level 3: difficult to perform, which has to press a few more buttons and requires long time visual attention on the task.

In total, there were over 3900 combinations of questions that the subjects were asked to select the answers.

3 RESULTS

The results from selecting one preferable system before discussing the detail of drive show in Table 3.

There are significant differences among Swedish drivers and Chinese drivers when it comes to the preferences of control system for the telematics use.

Then, when it comes to the detail drive situation, it shows very different figure. The results from Swedish driver showed in Table 4.

If we disregard the traffic condition and task complexity, in average, 40% select Alt. I, 32,8% select Alt. II and 27,2% with Alt. III. It showed very different figures from Table 3.

Among the 42 Chinese drivers, there were 4 of them did not complete the interviews, therefore, during data analysis, we took off their data. So the results in Table 5 came from 38 subjects.

Table 3. The preference of the control for telematics use during drive.

Control alternatives	Swedish driver (total 18)	Chinese driver (total 42)
I	0	2%
II	39%	9,5%
III	61%	83%

Table 4. The preferences of Swedish driver.

Traffic L	1	2	3	Average
Task L 1				
Alt. I	44.5%	50.2%	50.6%	48.5%
Alt. II	31.6%	25.4%	22.3%	26.5%
Alt. III	23.8%	24.4%	27.2%	25.1%
Total	240	480	714	
Task L 2				
Alt. I	25.0%	28.5%	29.4%	27.6%
Alt. II	40.5%	42.3%	39.0%	40.6%
Alt. III	34.5%	29.3%	31.6%	31.8%
Total	200	400	595	
Task L 3				
Alt. I	32.9%	53.3%	45.5%	43.9%
Alt. II	36.7%	25.7%	31.1%	31.2%
Alt. III	30.4%	20.9%	23.3%	24.9%
Total	240	540	714	

Table 5. The preferences of Chinese driver.

Traffic L	1	2	3	Average
Task L 1				
Alt. I	20,5%	29,7%	30,3%	26,8%
Alt. II	3,8%	5,2%	4,2%	4,4%
Alt. III	72,0%	65,1%	65,5%	67,5%
Total	468	936	1404	
Task L 2				
Alt. I	31,5%	39,5%	41,6%	37,5%
Alt. II	4,1%	6,4%	4,5%	5,0%
Alt. III	64,4%	54,1%	53,8%	57,4%
Total	390	780	1170	
Task L 3				
Alt. I	33,3%	47,9%	49,1%	43,4%
Alt. II	3,6%	5,4%	3,4%	4,2%
Alt. III	63,1%	46,7%	47,4%	52,4%
Total	468	936	1404	

If we disregard the traffic condition and task complexity, in average, 35,9% select Alt. I, 4,5% select Alt. II and 59,1% with Alt. III. It showed very different figures from Table 3.

4 DISCUSSION

The comparison studies between Swedish driver and Chinese driver have been carried out earlier (Lindgren 2006; Lindgren 2007; Lindgren 2008; Lindgren 2008; Chen 2009; Jordan 2009). All of these studies pointed out that there is a culture differences regarding interactive design for driver safety. In present study, again, indicated that there is a difference when considering the attitude of using telematic devices while drive.

Among the three different alternative way of control the use of telematic devices, Alt. II is the drive behavior system, which is the principle of how the most workload management system works. Our results showed that there are 32% of Swedish drivers appreciated such a system. But only 4,5% of Chinese driver would select such a system to their car. The reason behind the results can be very complicated, and it is not the intention of present study to find out why.

Through the personal experience, I would assume that the reason behind this was that Chinese driver were not sure how the workload management system worked in practice while Swedish driver probably already seen or heard about the system.

The highest preference of the control system among Swedish driver is actually the on-off system (40%). This means the Swedish driver prefer the vehicle can broke certain functions on the telematic devices during drive. For Chinese driver, most of them prefer the advice system (59,1%). The reason is probably that in Sweden, drivers were educated, through many different channels and media, that there is a potential danger of using telematics devices during drive. Some Countries have even set up the laws about forbidden using mobile phone while drive. In China, the public education, about how the why it is dangerous to use telematic devices while drive, was very limited. The drivers consider of accessible and freedom of using telematic devices are more important. They may over-estimate their capacity of handling secondary task performance during drive.

Off course, the number of interviewees in this study is very limited. It is not possible to extend the result with general meaning, rather consider it as a pilot study to indicate that some further investigation is needed in this area.

Another interesting issue in this study is that when we ask the general question about which control system they prefer, without close connection to tasks and traffic, very few drivers would consider using the total control system. The results showed the same for Chinese driver and Swedish driver. All of them said that they would like to take the control of the car, not the car take over the control of them, especially regarding the control of telematic devices (see the results in table 3). But when we combine the situation of traffic situation and detail task performance, the outcome became very different. This implies that research methodology is an important issue.

A result from subjective study can be very different from the study of real drive. Still, the study indicated that the culture factor cannot be ignored. More research work is needed.

5 CONCLUSION

The fact that telematics use during drive can cause driver distraction and potential accidents. Therefore, automotive industry is interested in workload management system. One unsolved problem appeared: how shall the vehicle “control” driver’s access of telematic devices.

Present study is based on structured interview with drivers from China and Sweden. Three independent factors are considered: the task complexity when regarding using the devices; the traffic condition and road conditions. Three alternative controls are evaluated.

The results indicate a strong culture impact when regarding what the driver’s attitude towards the control of accessibility of telematic devices. Further investigation is strongly needed.

REFERENCES

Anttila, V. & Luoma, J. 2005. Surrogate in-vehicle information systems and driver behaviour in an urban environment: A field study on the effects of visual and cognitive load. *Elsevier 8 (Transportation Research Part F)*: 121–133.

Chen, F. & Jordan P. 2008. Zonal adaptive workload management system: Limiting secondary task while driving. *IEEE Intelligent Transportation System* IVs' 08, Eindhoven, The Netherlander, June 2–5.

Chen, F. & Jordan, P. 2009. Reducing secondary-task workload while driving through interactive interfaces. *17th World Congress on Ergonomics, IEA 2009*, August 9–14, 2009, Beijing, China.

Chen, J., Liu, Z., Lindgren, A., Mendoza, P., Chen, F. (2009). Acceptance of integrated active system in China. *17th World Congress on Ergonomics (IEA 2009)*, Beijing, China.

Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J. & Knipling, R.R. 2005. *The 100-car naturalistic driving study, phase ii—results of the 100-car field experiment.* Technical report, Virginia Tech Transportation Institute, 3500 Transportation Research Plaza (0536), Blacksburg, Virginia 24061.

Green, P. 2004. Driver distraction, telematics design, and workload managers: Safety issues and solutions. In *Proceedings of the 2004 International Congress on Transportation Electronics* (SAE publication P-387).

Hancock, P.A., Lesch, M. & Simmons, L. 2003. The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention* 35: 501–514.

Jordan, P. & Chen, F. 2009. Attitude towards advanced driver assistance systems—An international study. *17th World Congress on Ergonomics (IEA 2009)*, August 9–14, 2009, Beijing, China.

Lee, J.D., M.L. Ries, et al. 2000. *Can collision warning systems mitigate distraction due to in-vehicle devices?* Iowa, Cognitive Systems Laboratory, Department of Industrial Engineering, University of Iowa: 9.

Lindgren, A., Broström, R., Chen, F. & Bengtsson, P. 2006. Cultural differences in driver attitude towards advanced driver assistance systems (ADAS). *IEA World Conference of Ergonomics*, Maastricht, Holland, July 10–14.

Lindgren, A., Chen, F., Jordan, P.W. & Zhang, H. 2008. Requirements for the design of advanced driver assistance systems—The differences between Swedish and Chinese drivers. *International Journal of Design* 2(2): 41–54.

Lindgren, A., Chen, F., Jordan, P.W. & Ljungstrand, P. 2008. Cross-cultural issues and driver requirements for advanced driver assistance systems. *2nd international conference on Applied Human Factors and Ergonomics*, Las Vegas, NV, July 14–18.

Lindgren, A.B., Chen, F. & Bengtsson, P. 2007. Driver attitudes towards advanced driver assistance systems—Cultural differences and similarities. *Human Factors Issues in Complex System Performance*. G.R.J.H.D. de Waard, P. Nickel. & K.A. Brookhuis (eds.), Maastricht, The Netherlands: Shaker Publishing: 205–215.

Liu, B.-S. & Lee, Y.-H. 2005. Effects of car-phone use and aggressive disposition during critical driving maneuvers. *Elsevier 8 (Transportation Research Part)*: 369–382.

Marcus, A. 2003. User-interface design and China: A great leap forward. *Interaction* 10(1): 21–25.

Xie, C.-Q. & Parker, D. 2002. A social psychological approach to driving violations in two Chinese cities. *Pergamon* 5: 293–308.

Zhang, W.H., Huang, Y-H., Roetting, M., Wang, Y. & Wei, H. 2006. Driver's views and behaviors about safety in China—What do they NOT know about driving? *Accident Analysis and Prevention* 38: 22–27.