Textbook of Cognitive Psychology

M. Asch

IVY PUBLISHING HOUSE
DELHI-110095

Sole Distributors

SARUP & SONS
NEW DELHI-110002

Published by

IVY PUBLISHING HOUSE
Delhi-110095

Sole Distributors

SARUP & SONS
4740/23, Ansari Road, Darya Ganj
New Delhi-110002
Ph. : 3281029, 3244664

Textbook of Cognitive Psychology

Ist Edition : 2002

ISBN—81-7890-044-0

Printed in India

Published by IVY Publishing House, Delhi, Laser Typesetting at Mayank Printers and Printed at Mehra Offset Press, Delhi.

Preface

This text is a handy and reliable source for the current concepts, related to Cognitive Psychology. The language of the concepts has been simple. The aim of this text is to present the principles and basics of Cognitive Psychology as a conceptual framework. In this manner it can realistically serve the needs of graduate and post-graduate students of Psychology. The present publication would hopefully be a welcome addition to the rather meagre literature in the vital area.

—Editor

Contents

1

Cognitive Psychology

Cognitive psychology is concerned with information processing, and includes a variety of processes such as attention, perception, learning, and memory; it is also concerned with the structures and representations involved in cognition. The greatest difference between the approach adopted by cognitive psychologists and that followed by the Behaviorists is that cognitive psychologists are interested in identifying in detail what happens between stimulus and response. Some of the major ingredients of the information-processing approach to cognition were spelled out clearly by Lachman, Lachman, and Butterfield (1979). In essence, it is assume that the mind can be regarded as a general-purpose, symbol-processing system, and that these symbols are transformed into other symbols as a result of being acted on by different processes. These processes require time to be carried out, so that reaction-time data can provide a useful source of information. The mind has structural and resource limitations, and so should be thought of as a limited-capacity processor.

It is generally accepted that the human mind is an information-processing system, and the same is true of computers. A key issue is the extent to which these two kinds of information-processing systems resemble each other. Some psychologists have had no doubts about the answer. For example, according to Simon (1980, p. 45), "It might have been necessary a decade ago to argue for the commonality of the information processes that are employed by such disparate systems as computers and human nervous systems. The evidence for that

commonality is now over-whelming." The consensual view is probably that there are, indeed, striking similarities between computers and human minds, but there are also substantial differences.

One of the most significant developments within cognitive psychology in recent years has been a proliferation of the number of different approaches to human cognition. As Eysenck and Keane (1990) pointed out, it is possible to identify at least three major categories of psychologists whose main interest in in cognition: experimental cognitive psychologists; cognitive scientists; and cognitive neuropsychologists. Experimental cognitive psychologists concentrate on empirical research into cognition on normal subjects. Cognitive scientists develop computational models, and attach great significance to the computer as a metaphor for human cognition. Cognitive neuropsychologists investigate cognition in brain-damaged patients, and claim that the study of such patients can provide valuable insights into normal cognitive functioning.

Another major development within cognitive psychology has involved the application of the a paradigms and theories of cognitive psychology to several different areas of study. For example, explicitly cognitive approaches have been adopted in social and developmental psychology.

Cognitive Psychology, History of

Cognitive Psychology, history of The evolution of cognitive psychology during the second half of the twentieth century can usefully be discussed in the context of the dominant approach which preceded it, i.e. Behaviorism. In the early years of this century, John Watson argued that the way to make psychology a legitimate experimental and scientific discipline was by focusing only on observable entities. This meant that the emphasis was very firmly on the relationship between observable stimuli and observable responses. As part of the Behaviorist approach, there was a reluctance to introduce hypothetical mental constructs into theoretical psychology.

The emergence of Behaviorism occurred to some extent because Watson and other Behaviorists wanted psychology to emulate the more established sciences such as physics and chemistry. It was argued by logical positivists such as Carnap, for example, that theoretical constructs in any science are meaningful only to the extent

that they can observed. In addition, scientific theories can only be justified by appealing to observed facts. The views of the logical positivists led some of the leading Behaviorists (e.g. B. F. Skinner) to argue that physics and chemistry had been more successful than psychology because physicists and chemists had adhered more firmly than psychologists to those characteristics of good science advocated by the logical positivists.

Behaviorism was extremely influential for a long time, especially in the United States of America. However, even there it gradually lost its appeal, particularly during the 1950s and thereafter. There were basically two reasons why this happened. First, Behaviorism never produced detailed or adequate accounts of complex cognitive functioning. It was possible (although probably erroneous) to explain many phenomena in conditioning in terms of associations between stimuli or between stimuli and reponses, but it made little or no sense to attempt to account for our knowledge of a complex system such as language in stimulus-response terms. Precisely the same was true of attempts by Behaviorists to account for cognitive activities such as Creativity and Problem Solving.

Second, philosophers of science during the twentieth century increasingly challenged traditional views of the scientific enterprise. Popper (e.g. 1972), for example, attacked the notion that scientific observation possesses objectivity. He claimed instead that scientific observations are based very much on preconceived ideas and theories. He used to make this point when lecturing by instructing the members of the audience that they should observe, to which their response was typically, "Observe what?" In other words, observations does not proceed in a vacuum, but is influenced strongly by what we are looking for.

Traditional views of science were challenged most strongly by Feyera bend (1975). He argued that there are remarkably few rules which actually constrain the activities of scientists. In practice, the only rule which is followed extensively is "anything goes". Science differs more from non-science than Feyerabend admitted, but there is no doubt that his views and those of many other philosophers of science had a liberating effect on psychology. If the hard sciences such as physics and chemistry did not follow very strict rules, then there was no need for psychology to do so either. This meant that the

way was open for the rigidities and limitations of Behaviorism to be discarded in favor of more flexible approaches, of which congnitive psychology soon established itself as the main one.

It is extremely difficult to identify the starting point for a major academic discipline such as cognitive psychology. One of the issues is that one needs to distinguish between early work which has clear affinities with contemporary cognitive psychology but which nevertheless had minimal impact on the development of cognitive psychology, and work which actually played a part in the emergence of cognitive psychology. A very clear example of work falling into the former category is the research carried out by neuropsycholgists toward the end of the nineteenth century. They endeavoured to account for language impairments in brain-damaged patients by postulating damage to specific language-processing components of the brain, and they also tried to locate the parts of the brain in which these components were to be found. In spite of the fact that their work has strong relevance to that branch of contemporary cognitive psychology knownas Cognitive Neuropsychology, the research and theory of the nineteenth-century neuropsychologists had practically no impact on the emergence of cognitive psychology in the 1950s.

There is general agreement that the ideas of William James (1890) were extremely influential in the development of cognitive psychology. He was primarily as theorist, but many of his ideas on Attention and on Memory still seem acceptable today. For example, he distinguished between "primary memory," which he regarded as forming the psychological present, and "secondary memory," which he defined as the psychological past. Cognitive psychologists such as Atkinson and Shiffrin (1968) proposed a very similar distinction between Short-Term Memory and Long-Term Memory.

Another important influence on cognitive psychology was the work of Bartlett (Sir Frederic) (1932). As early as the First World War, he began investigating memory in a relatively naturalistic way by considering how well stories could be remembered at different retention intervals. Of particular importance was his theoretical approach to memory. He argued that what is remembered is determined in part by the Schemata (= organized knowledge) which are possessed by those reading a story. Bartlett's (1932) schema theory had little impact on memory research during the 1940s and 1950s, but became the focus

of considerable interest to cognitive psychologists during the 1960s and thereafter.

Some of the other important antecedents of contemporary cognitive psychology are to be found within Behaviorism itself. Tolman (1932) was one of the leading Bevhaviorists, but his research led him to modify Behaviorism in various ways that made it resemble cognitive psychology more closely. Hull and other researchers had argued in strict Behaviorist terms that rats learned to run through mazes by associating the maze stimuli with specific responses, including particular muscle movements. Tolman (1932) became convinced that maze running in the rat involved much more than these simple stimulus-response connections. He discovered that rats who had initially been trained to run through a maze were remarkably good at swimming through the maze when it was flooded with water, in spite of the fact that the muscle movements involved were quite different to those that had allegedly been acquired. This led Tolman to argue that a rat running through a maze several times gradually builds up a 'cognitive map'; this is an internal representation of the maze which permits the rat to run or to swim through the maze depending on the prevailing circumstances. What is significant here is that Tolman discovered that learning in the rat could be understood only by considering internal processes and structures.

The advent of the digital computer as a metaphor for the human cognitive system played an important role in the development of cognitive psychology. Psychologists have historically shown a strong tendency to make use of recent technological developments as metaphors for major aspects of human functioning. This can be seen very clearly in the case of theoretical attempts to describe memory (Roediger, 1980). The ancient Greeks likened the functioning of the memory system to wax tablets and avaries. Over the centuries, these metaphors were replaced by others involving switchboards, gramophones, tape-recorders, libraries, conveyor belts, and underground maps. In the case of the digital computer, it was argued that there are important similarities between its functioning and that of the human mind. According to Simon (1980, p. 45), "It might have been necessary a decade ago to argue for the community of the information processes that are employed by such disparate systems as computers and human nervous systems. The evidence for that

community is now overwhelming."

Gardner (1985) has charted the development of cognitive psychology. He argued justifiably that 1956 was a crucial year. There was a meeting held at the Massachusetts Institute of Technology in that year at which Gorge Miller talked about the magic number seven in short-term memory, Newell and Simon discussed their computational model known as the General Problem Solver, and Noam Chomsky presented a paper on his theory of language. During the same year, there was the well-known Dartmouth Conference, which was attended by Chomsky, McCarthy, Miller, Minsky, Newell, and Simon. It has often been argued that this conference saw the birth of Artificial Intellegence. The same year also saw the publication of the first book in which concept formation was investigated from a cognitive-psychological viewpoint (Bruner, Goodnow, & Austin, 1956).

Cognitive psychology during the 1960s and 1970s was strongly influenced by the seminal theorizing of Broadbent (1958). In essence, it was accepted that there are important relationships among the phenomena of attention, perception, short-term memory, and long-term memory. All of them could be understood by assuming that information flows through a complex cognitive system consisting of many interdependent processes. Within this theoretical framework, stimulus processing proceeds through a relatively invariant sequence of stages from modality-specific stores to its ultimate destination in long-term memory.

One of the best attempts to characterize the dominant information-processing framework in cognitive psychology was that of Lachman, Lachman, and Butterfield (1979). This framework incorporates a number of assumptions. One assumption was that the mind can be regarded as a general-purpose, symbol-processing system. Another assumption was that the goal of cognitive pyschology is to identify the symbolic processes and representations which are involved in performance on all cognitive tasks. A further assumption was that the mind is a limited-capacity processor which has limitations of both structural and resource kinds.

This general framework still seems to be reasonable. However, one of the major weaknesses with this approach as it was usually implemented in the 1960s and 1970s was that the emphasis tended to be on data-driven rather than on conceptually driven processes. In

other words, the ways in which stimulus processing is modified as a function of the individual's past experience and expectations were ignored. It was often assumed that processing occurs in a serial fashion (i.e. one process is completed before the next process starts). While strictly serial processing probably occurs on certain tasks, it is now recognized that the assumption that processing is always serial is invalid. Alternative views which acknowledge that processes frequently overlap and interact with each other have become increasingly popular in recent years.

Another major limitation of the research in cognitive psychology of the 1960s and 1970s was that it was mostly carried out in laboratory conditions and addressed academic rather than applied questions. In other words, cognitive psychology lacked what is generally termed Ecological Validity, meaning relevance to real-life problems and issues. The position has changed considerably in recent years. For example, there has been an enormous increase in research on language, and language is, of course, of central importance in real life. Key applied issues such as Eye-Witness Testimory have been investigated in great detail. In addition, and perhaps most noticeably, there has been a dramatic rise in the amount of research within cognitive psychology which is concerned with examining the cognitive performance of numerous special groups within society (e.g. brain-damaged patients; mood-disordered patients).

If one considers the current position of cognitive psychology, then it is apparent that cognitive psychologists differ considerably among themselves in terms of their aims and their approaches. Indeed, it is probably correct to claim that the most obvious difference between contemporary cognitive psychology and the cognitive psychology of ten or twenty years ago lies in its diversity. Cognitive psychologists are to be found within social psychology, developmental psychology, and personality psychology. Perhaps most surprisingly of all, cognitive psychologists have even begun to attack the citadel of Behaviorism, i.e. conditioning phenomena. For example, it is increasingly recognized that conditioning depends on information processing, and that it involves the selection of relevant information and its integration with stored information about relevant past experiences and events (Alloy & Tabachnik, 1984).

Eysenck and Keane (1990) argued that it is possible to divide

cognitive psychologists into at least three major groups. First, there are the experimental cognitive psychologists, who follow the traditional cognitive-psychological approach of focusing on data collection and theory construction. Second, there are cognitive scientists, who develop computational models, and who argue that the computer provides a good metaphor for human cognition. Congnitive scientists differ among themselves in terms of their assessment of the value of traditional experimentation. Third, there are cognitive neuropysychologits. They are interested in the patterns of cognitive impairment shown by brain-damaged patients in part because they believe that the study of brain-damaged patients can be informative about the processes involved in normal human cognition. In essence, their claim is that several processing modules or components are involved in cognition. Since different patients have different modules impaired, it should in principle be possible to identify most (or even all) of the modules involved in cognition by detailed investigation of brain-damaged patients.

There are arguments for identifying a fourth group of cognitive psychologists who might be called applied cognitive psychologists. It is certainly true that applied cognitive psychologists differ from other cognitive psychologists in terms of what they investigate and the methods they employ. However, it is much less clear that applied cognitive psychologists and other cognitive psychologists differ systematically in terms of their theoretical preconceptions and orientations, and it is for this reason that it seems unnecessary to extend the categorization scheme beyond the three groupings discussed above.

Of course, there are many cognitive psychologists who do not fit neatly into any of the above categories. For example, there are many cognitive psychologists in the United Kingdom who sometimes behave like experimental cognitive psychologists but at other times like cognitive neuropsychologists. As a consequence, the distinctions between the three categories of congnitive psychologists cannot be regarded as absolute. However, Eysenck and Keane (1990) argued that there are very many cognitive psychologists who fall squarely into one or other of the categories, and so the categorization scheme is of value.

The various categories of cognitive psychologists differ in terms of their adherence to empiricist and rationalist perspectives.

Experimental cognitive psychologists and cognitive neuropsychologists tend to be empiricists, in that they assume that the way to understand human behavior is via observation and experimentation. In contrast, cognitive scientists tend to be rationalists – that is to say, they believe that the construction of formal systems resembling those to be found in mathematics is the appropriate way to proceed, Of course, there are many cognitive psychologists who adopt a compromise position which is neither entirely empiricist nor completely rationalist.

The recent history of cognitive psychology has not indicated that one of the approaches (i.e. experimental cognitive psychology, cognitive science, cognitive neuropsychology) is intrinsically superior to the others. Each of the approaches is of value in its own right, and what is of particular importance is the attempt to demonstrate that all three approaches produce *converging evidence.* In other words, it is possible to have more confidence that a theory is on the right lines when it is supported by all three approaches than when it is supported by only one or two. That means that in future there is likely to be an expansion and development of the three approaches rather than an abandonment of any of them.

In sum, the history of cognitive psychology indicates that congnitive psychology has become increasingly influential and diverse over the years. At one time, cognitive psychology was rather narrowly focused on laboratory phenomena. However, the methods and theoretical perspectives of cognitive psychology have now permeated nearly every area of psychology. Thomas Kuhn has argued famously that a scientific discipline is generally dominated by a particular theoretical orientation which he called a "paradigm." There are strong arguments for supposing that the information-processing approach of cognitive psychology constitutes such a paradigm.

Bartlett

Bartlett, Sir Frederic FRS (1886-1969). First professor of Psychology at the University of Cambridge, Sir Frederic Bartlett was a leading figure in the development of British psychology and the author of a number of books, including *Psychology and the soldier* (1927) and *The problem of noise* (1934). However, within cognitive psychology, Bartlett's lasting influence comes from his work *Remembering: An experimental and social study,* published in 1932.

Bartlett took the view the experimental psychology should relate to the real world – a fact reflected in the assistance he gave to the Ministry of Defence during both world wars. In *Remembering,* Bartlett advanced the view that much of the experimental work concerned with human memory lacked validity because the experiments were unrealistic. At the time memory research was dominated by the empirical methods originating from Ebbinghaus in the nineteenth century. Ebbinghaus stressed the importance of using meaningless material such as nonsense syllables in memory experiments – the argument being that greater experimental rigor could be obtained because subjects could not bring pre-existing knowledge to bear on the learning task.

Bartlett argued that these methods were unlikely to discover anything significant about memory. He proposed that the basis of human learning was an "effort after meaning" and that to exclude it from investigation was to miss the essence of human memory.

In *Remembering,* Bartlett describes a number of experiments investigating how people remember meaningful material. Most well known is "The War of the Gohsts" experiment in which he examined subjects' successive recall of a Red Indian folk story. The most striking aspect of the data was that subjects introduced aspects of their own knowledge about the world so as to make the story more coherent from their own point of view.

To account for these results Bartlett introduced the term schema. Originated by the neurologist Head, the term schema was used to describe the internal body image that enables us, for example, to know the relative position of our limbs and their relation to other features of the environment. Bartlett extended the idea by suggesting that we also have internal Schemata dealing with our knowledge of how the world is. When remembering he believed that these schema were brought into play as a means of aiding retrieval. Schema could take many forms. At one level schema could take many forms. At one level schema reflecting the subject's attitude to the whole story could influence the manner in which it was recalled. At a more specific level a schema might result in a distortion or transformation of detail more consistent with the subject's view of the world.

Because of methodological shortcomings associated with tasks such as successive recall, and the difficulty of devising specific

theories, Bartlestt's arguments had little impact on memory research. However, in 1978. Neisser revived the objections first raised by Bartlett and there has since been a growing interest in studying memory outside the laboratory. Bartlett's views on schema are also reflected in recent developments in cognitive psychology, most notably Schank and Abelson's (1977) concept of scripts.

Broadbent

Broadbent, Donald Eric, FRS, CBE (b. 6 May 1926). British psychologist. Donald Broadbent is one of the key figures in the development of cognitive psychology. His book *Perception and communication,* which was published in 1958, served in many ways to set the agenda for the subsequent direction taken by cognitive psychology. In that seminal book, he argued persuasively that one can understand such apparently diverse phenomena as Perception, Attention, and Short-Term Memory by putting forward a theory in which information flows through a cognitive system. In other words, instead of considering all of these phenomena in isolation from each other, it is preferable to treat them as interdependent ingredients in a single cognitive system.

Some of the ideas to be found in Broadbent (1958) represent extensions of earlier theories of communications systems (e.g. Shannon & Weaver, 1949). However, Broadbent (1958) was the first psychologist to propose a systematic and detailed account of the cognitive system viewed as consisting of a set of separate but interacting components. This general approach is still influential and continues to be endorsed by cognitive psychologists, cognitive scientists, and cognitive neuropsychologists.

At the more specific level, Broadbent (1958) made particularly major contributions to our theoretical understanding of attention and short-term memory. There had, of course, been previous theories of both attention and short-term memory. However, Broadbent's (1958) theoretical contributions provided the impetus for a substantial reawakening of interest in these, and other, topics. For example, he addressed the issue of the fate of unattended information both theoretically and empirically, and his pioneering work set the stage for subsequent theories of attention (e.g. Treisman, 1964).

Broadbent subsequently developed and extended his theoretical

views in various ways (e.g. Broadbent, 1971). For example, he greatly increased our theoretical understanding of the relationship between stress and cognitive functioning. In addition, he modified his theoretical views on attention in light of the new research which had been carried out in response to his earlier book. During the 1980s, he and his colleagues carried out much innovative research concerned with some of the factors associated with cognitive failure. Broadbent's Cognitive Failures Questionnaire (Broadbent, Cooper, Fitzgerald, & Parkes 1982) was the main measure of cognitive failure that was used. Among the important findings that emerged was the suggestion that there is a general factor of cognitive failure.

In sum, Broadbent is the most influential of all British cognitive psychologists. Apart from the outstanding contribution he has made at the theoretical level within academic psychology, he also has the rare ability to carry out practically important applied research having a genuine theoretical focus.

Emotion

The term "emotion" denotes various enjoyable or distressing mental states. There has been a tendency to think of emotions as biological and of little interest for cognitive analyses. Indeed, they have even been somewhat neglected in psychology generally. The study of emotions has, however, come to be of considerable importance for cognitive pyschology and there is a growing body of evidence and theory about the attributes and functions of the these states.

History of Concepts of Emotion in Psychology

The reason emotions have previously been neglected in human psychology is no doubt due to the tradition in which they have been seen as enemies of reason. This view was held by Plato, who thought that they were like drugs that distort rationality. This distrust was continued by the Stoics who argued that emotions were diseases of the soul.

Modern investigations of emotions began with Darwin, whose work continued this line of thinking. After writing the *Origin of species,* in which humans were scarcly mentioned, he devoted much of his time to collecting evidence for human evolution. He found similarities

of emotional expression between humans and other animals, and between adults and children. He observed that these expressions occurred whether or not they were of the least use. He was therefore able to argue that when they occur in adult humans, they are echoes of evolutionary history, and of the history of our development from childhood. They are behavioral vestiges, no longer necessarily functional, rather like the small bones at the base of our spine which are anatomical remnants of tails that our ancestors once had.

Following Darwin, James (William) argued that emotions are perceptions of bodily happenings, and Walter B. Cannon argued that they are upwellings of impulses from lower, non-rational parts of the brain. Such proposals constituted the mainstream of psychological theorizing about emotions.

Influential though this tradition has been, there has also been another tradition of thought about emotions in the history of Western psychological ideas. It started with Aristotle, who argued that emotions are based on cognitive judgements; and he gave componential analyses of emotion terms. So, for instance, he analyzed anger as a state caused by a belief that one has been insulted. It is directed at the perpetrator, and it is diminished by an apology or remedy. It does not occur if one is treated justly. This program of analyzing the meaning of emotion terms was continued by Spinoza.

In this cognitive tradition, the work which marks the modern starting point is that of Frederic Paulhan, who in 1887 argued that emotions occur when there is an interruption of an ongoing activity or tendency. An emotion then produces mental disturbance which entirely occupies conscious attention. This has become known as the conflict theory because emotions are seen as arising either with conflicts between a goal and an unexpected outcome, or when one goal conflicts with another.

Emotions in the Early Days of Cognitive Psychology

From around 1960 cognitive psychologists began to realize that, in computation, interruptions of one process by another are needed where resources are limited and where there are multiple goals. Systems must allow interruptions and establish goal priorities. These issues are coextensive with those of emotions in humans. Such arguments implied that we can take steps toward understanding

emotions by considering the design of cognitive systems, and by analyzing emotions in terms of theories of action.

During the 1960s, also, Schachter and Singer (1962) published a paper which is the most influential recent research on emotions. They gave injections of adrenalin which produced arousal. If their subjects had no other explanation for the arousal produced by the adrenalin, they attributed it to the social context in which it occurred. So arousal of light-hearted social interaction was experienced as happiness, and where frustrating and intrusive things were going on it was experienced asanger. An emotion, according to this idea, has two components, a noncognitive, undifferentiated arousal – not very different from the kind of input postulated by James and Cannon – together with a cognitive attribution of the cause of the arousal. At one stroke Schachter and Singer united the biological, cognitive, and social-psychological traditions. They also gave a theoretical basis for understanding psychosomatic interactions which has still not been much improved upon.

The implications of Schachter and Singer's theory of misattributing the causes of arousal have been only partly supported empirically, but their idea has provided foundations for cognitive theory. When an event interrupts an ongoing plan, arousal occurs. The evaluation of the event gives rise to the emotion and prompts the next phase of action.

Current cognitive-psychological interest in emotions is in three main areas: cognitive theories of emotions as evaluations of circumstances, interactions of emotions with memory and thinking, and linguistic analyses of inferences that can be made about emotions.

Cognitive Theories and Evidence of Emotions as Evaluations

Since the 1960s there has been growing interest in theories of emotions as evaluations of events. The most comprehensive review of empirical evidence in this area, together with a scholarly discussion of the kind of cognitive theory that can account for it, is Frijda's (1986) book.

Frijda's type of theory now has rather general agreement. It is that we can think of emotions as occurring when events are evaluated in relation to a person's important goals – concerns, as he calls them. An emotion is a process that typically starts with an eliciting event.

This is perceived and coded, and it is then appraised in relation to concerns. Its significance is assessed in terms of priorities and what can be done about it. Readiness for action occurs, which may be accompanied by physiological changes such as increasing the heart rate. According to Frijda, then, readiness is the core of an emotion. It may issue in actual actions, including emotional expressions.

Emotions differ in the kinds of appraisals and evaluations that are made. For example, if an event occurs that frustrates an important goal, and you judge that it was deliberately caused by someone who had no right to do it, and that you might be able to obtain some redress – then you are likely to feel anger. But if you judge that the even was an accident, or if it is impossible to do anything about it, then anger is less likely.

Of course, this is close to the kind of analysis that Aristotle gave. What has happened recently is that there is now substantial empirical evidence drawn from a range of psychological methods, in different cultures, on adults and children, that emotions do indeed occur like this.

It is now accepted that, despite Schachter and Singer's use of adrenalin of cause arousal, emotions are psychological. Whereas physical exposure of a person to a cold atmosphere has reliable bodily effects of making that person feel cold and shiver, emotions do not occur purely as a result of physical events. They typically occur because of evaluations of events in relation to a person's individual concerns and expectations. For instance, the same physical pattern of pressure of a hand on the skin could elicit a feeling of happiness, of anger, or of disgust, depending on who was doing the touching and how its recipient evaluated it.

A variation on this type of cognitive theory, which relates more closely to issues raised by Darwin, is due to Oatlay and Johnson-Laird (1987) who argue that in mammalian life there is a small number of basic types of evaluation of event in relation to goals. These evaluations produce a set of basic emotions, corresponding to (a) happiness, (b) sadness, (c) anger, (d) fear, and (e) disgust. These five types of emotion occur when there is: (a) achievement of subgoals, (b) loss of a goal, (c) frustration of a plan or goal by another person, (d) conflict of goals including conflict with a self-preservation goal, and (e) perception that something is noxious or toxic.

These emotions depend on each of our goals being monitored whether they are contributing to an ongoing plan or not. From the monitoring processes simple signals are sent out whenever progress toward any significant goal changes substantially, for better or worse. These signals have no propositional structure, but each tends to set the cognitive system into a mode which, during evolution, has been selected as appropriate to whatever juncture (or types a to e) has been detected. There may be more than five of these general modes. Determining the number will depend on physiological and cross-cultural evidence.

The different modes are not equivalent to the general arousal postulated by Schachter and Singer. Rather, happiness, sadness, anger, fear, and disgust are physiologically distinguishable, as shown by Ekman, Levenson, and Friesen (1983). However, as with Schachter's theory, Oatlay and Johnson-Laird postulate that the non-propositional part of the emotion is typically accompanied by propositional information about what caused it, about whom it concerns, and so on. This theory, moreover, resolves the difficulty of most conflict theories in explaining happiness. Happiness occurs when subgoals in a sequence of actions are being achieved, with problems that arise being solved from available resources, and when other goals do not conflict or distract.

Emotions are not just vestiges. They maintain their importance in human cognition because action is typically influenced by many simultaneous goals. Human action takes place without perfect mental models of the world, and often involves coordination with other people. Since, therefore, the world is not entirely predictable, and since conflicting goals cannot always be reconciled either within ourselves or between people, our action can hardly ever be perfectly rational. So to argue that emotions are irrational, as Plato did, misses an important point: rational solutions to problems of human action are only occasionally possible. Emotions function to help manage the vicissitudes of action in a world that can only be very imperfectly known. We detect junctures at which the unexpected happens or when progress toward any of our goals changes significantly. At such points emotions make the system ready in a certain kind of way, prompting us towards the next phase of action, and communicating nonverbally to others nearby. Thus, if a subgoal is achieve, we move

happily towards the next phase of action. If some danger is detected, the current plan is interrupted, checks are made anxiously on what had been done, the environment is scanned, escape or other actions to gain safety are prepared. As Oately and Johnson-Laird put it, the function of emotions is to communicate these changes in readiness to ourselves and to others.

Emotions have evolved, according to this argument, because achieving subgoals, suffering losses, detecting conflicts of goals, and so on recur frequently in the life of mammals. Each of these general junctures is recognizable, and it has been cognitively efficient for a general readiness appropriate to each of these junctures to be available, so that it can be invoked along with the default plans and interpersonal communications that are stored in association with it. These default plans are either those that have proved successful in evolution and have been genetically transmitted, or they are habits acquired in individual development.

If we were to analyze each individual circumstance of subgoal achievement, loss, goal conflict, and so on only by thinking about it, we would take too long, and would be too prone to errors. We would often not have the necessary evidence for good solutions. Emotions as automatic evaluations of recurring classes of basic evaluations, each prompting an appropriate readiness to act, has proved advantageous in evolution, not because emotions offer perfect solutions - there is often no perfect solution in such circumstances. They have been selected, presumably, because on average the genetically stored programs and individually acquired habits have been more successful than other kinds of solution.

We humans are not ants, biological machines emitting fixed action patterns in response to particular stimuli. Nor are we gods who have complete knowledge of the consequences of all actions and power to control all outcomes. We are somewhere in between. For us there is a substantial problem: we have imperfect mental models and limited resources, although our own plans do have some of the effects we aim for. We need, therefore, a means for managing those junctures in which all has not gone quite as expected. For such beings, readiness and default plans must be invoked quickly and without thinking, to prompt the next phase of action when a change of evaluation has occurred. Then, where something has gone wrong, the readiness to

act must be followed by longer-lasting mental preoccupation in which we can concentrate on some cognitive reprograming. These are indeed the properties of emotions, which are part of the biological-cognitive solution to these problems of goal and plan management.

Interactions of Emotions with Memory and Thinking

A second area of active cognitive interest is in the effects that emotions have on memory and other pyschological functions. Bower's (e.g. 1981) work has been influential: he and his colleagues did a series of experiments in which happy or sad moods were induced by hyponsis. When recalling a list in the same mood as that in which they had learned it, subjects recalled more words than when in the opposite mood.

Effects of mood have now been found using methods other than hypnosis, and there is now a wide range of evidence on this issue, collected in many different types of experiment. Mood affects attention, problem solving, imagery, free association, social perceptions, and response to stories. And as well as memory for laboratory tasks, autobiographical events are affected by mood. Remembered events that are congruent with mood during recall come selectively to mind.

These results have aroused clinical interest, since it has been shown that mood-congruent ideas come to the minds of patients with clinical depression or anxiety. So when patients are sad or anxious, sad or anxious thoughts come to their minds, and these then reinforce their distressed moods. Cognitive therapists, therefore, have become interested in cutting into this self-reinforcing cycle of dysphoria. Cognitive therapists also share the Platonic distrust of emotions: their favored way of cutting into the cycles of depression and anxiety is a version of the ancient Stoic idea that thinking about things in certain ways can minimize emotions. Again, there have been modern additions to this classical idea: there is now convincing evidence that cognitive therapy is more effective than other forms of therapy, including drugs, for alleviating depression, and that it has effects of a magnitude comparable to drugs for anxiety states.

Analysis of the Language of Emotions

The third area in which there has been considerable recent

research has been in the language of emotions. Two kinds of work have recently become important. First, there has been continuation of the work begun by Aristotle on componential analyses of emotion terms.

Emotions are mental states to be distinguished from states that are not primarily mental, lite pain or sleepiness. Many terms imply knowledge of emotions, embedded in folk theories that support the semantics of ordinary language. Thus, in English, many emotion terms, perhaps the majority, indicate both a kind of emotion and indicate that something elicited it or that it has effects. So to use the term "happy" means merely that I have a particular kind of emotion. But to use the term "glad" indicates that I also know that something caused my happiness. To say correctly that someone is jealous refers to a basic feeling of anger or disgust, and also to a complex social situation involving the possibility of losing a loved person to another, and to the kinds of actions that might occur.

Other research has implicated the idea of scripts, sequences of events occurring in a predictable order. Since the emotion process runs through a sequence of cognitive events, some researchers have proposed that it has properties of a script. The argument is that terms such as happiness, anger, anxiety, and so on may refer to this general script, in a prototypical way. Any given emotion term may thus be thought of as indicating a more or less good exemplar of a prototypical emotion. This type of proposal is at variance with componential analyses, and the question of how best to understand emotion terms is controversial.

Terms for emotions occur in all languages so far studied, though some languages have only a few such terms, and some have no generic term for "emotion." Not all cultures perceive emotions primarily as subjective experiences. Some focus on their interpersonal effects, and languages reflect these differences. Interpersonal issues are present also, although perhaps less obviously, in European and American culture. If the meaning of emotion terms is concerned with semantics and the meaning of individual experiences, these interpersonal effects can be thought of as concerned with pragmatics and on effects on others communicated by facial and other expressions. So, for instance, sadness tends to prompt others to help. Anger is directed toward someone with an intention of redress and is often the opening phase

of renegotiation in a relationship.

Because meanings and interpersonal implications of emotions are important to humans, they are crucial for any understanding of narrative. Narratives of many kinds tend to turn around incidents in which emotions occur often an emotion brings an episode to a close, and is followed by a new episode. Emotion words are informationally rich, and it is possible to make inferences about story characters, particularly about their goals and plans, from the emotions that the storyteller describes. Thus, for computational understanding of narratives, these kinds of inferences must be available. They must be made from goals and plans to the likely emotions of characters in the narrative, and also from emotion terms to implications for goals and plans. For these reasons, understanding the semantics and pragmatics of emotions will be necessary for computational understanding of narrative.

Conclusion

From a state in which understanding emotions seemed to have little part to play in cognitive psychology, research on emotions has become important, as is appropriate to the role of these states in managing problems of acting in the ordinary world with multiple goals and imperfect knowledge.

James, William (b. New York, 11 January 1842; d. Chocura, New Hampshire, 26 August 1910). American psychologist and philosopher, the brother of the novelist Hendry James. He entered Harvard Medical School in 1863 and graduated in 1869. He subsequently returned to Harvard, becoming in turn assistant professor of philosophy in 1880, professor of philosophy in 1885, professor of psychology in 1889, and then again professor of philosophy in 1897.

The contribution of William James to cognitive psychology rests very largely on his *Principles of psychology,* which was published in 1890. It is a difficult book to describe briefly. As Peters (1962) pointed out in his analysis,

Already James had begun is opposition to systems and was determined to be systematically erratic. He announces that "the reader will in vain seek for any closed system in the book; it is mainly a mass of descriptive details, running out into queries which only a metaphysics alive to the weight of her task can hope successfully to deal with."

The unsympathetic nature of the *Principles of psychology* occurred in part because of James's various interests: his medical training provided him with a good knowledge of physiology and biology, and his expertise in philosophy led him to a fascination with issues such as the stream of consciousness and free will.

It is impossible here to do more than identify a few of the contributions which James made directly to cognitive psychology. In the field of memory, he distinguished between primary memory or the psychological present, and secondary memory or the psychological past. This distinction has been very important over the past 30 years or so, although in modern terminology it is customary to distinguish between Short-Term Memory and Long-Term Memory. James emphasized the importance of ATTENTION, and particularly the necessity of selective attention. Indeed, according to James (1890), "Selection is the very keel on which our mental ship is built." Many of James's ideas about attention have had a strong influence on contemporary theories. While interest in attentional phenomena waned considerably during the Behaviorist era, it is now generally accepted that attention is of major significance. According to Keele and Neill (1978), the concept of attention has come to be regarded as lying "at the very core of cognitive psychology."

Another of the major contributions of William James was to the area of Emotion. The so-called James-Lange theory of emotion argued in essence that the cognitive expericence of emotion is heavily influenced by prior physiological changes. More specifically, there is first of all the perception of some external object or creature. This is followed by physiological changes and/or by movement. Finally, the observation of these changes produces the emotional experience. As James himself expressed it, "We are sad because we weep."

Language

Language, pragmatics of In Charles Morris's (1938) scheme, pragmatics is the study of "the relation of signs to interpreters." It is distinguished from semantics, the study of " the relation of signs to the objects to which the signs are applicable," and syntax, the study of "the formal relation of signs to one another". So whereas semantics is the study of what words and sentences mean, pragmatics has come to be the study of what speakers mean in using them.

Most approaches to pragmatic have come out of the work of three Oxford philosophers, H. Pual Grice, Peter Strawson, and John Austin. It was Grice (1957) who noted that one must distinguish what words and sentences mean (world and sentence meaning) from what speakers mean in using them (speaker's meaning). The meaning of the sentence "I am now over here," for example, is built around the meanings of the words "I," "am," "now," "over," and "here." But that does not specify what a speaker is doing in uttering the sentence on a particular occasion. When Deirdre, say, utters it in London in 1945, she may use "I" to refer to herself, "now" to refer to that moment in 1945, and "over here" to refer to the side of Piccadilly opposite Dan, her addressee. She may be asserting to Dan that she is standing across the street from him. And, indirectly, she may also be requesting him to cross the street to meet her. Pragmatics, indeed, has focused on three issues illustrated in this example: (1) reference; (2) speech acts such as assertions and requests; and (3) implicatures, or what people imply in what they mean (Levinson, 1983).

A phrase like "the woman in the blue dress," Strawson (1950) argued, does not by itself refer to a particular woman. That comes only when a speaker uses it on a particular occasion. The problem is how addressees identify who or what the speaker is referring to. In many accounts, speakers and addressees presuppose a certain common ground—certain mutual knowledge, beliefs, and assumptions. Two people's common ground is base on their shared perceptual experiences, previous conversations (including the words just uttered), and shared culture. The way speakers succeed in referring is by designing their utterances to be interpreted against the common ground with their addressees. When Deirdre chooses the phrase "the woman in the blue dress", she assumes that Dan will be able to identify the woman being referred to as the most salient woman in a blue dress in their current common ground.

Reference comes in many varieties. With demonstrative reference, the speaker performs an accompanying demonstration. Deirdre might say "That is Julius Marx" while pointing at a man, an empty chair, or a Rolls Royce; to understand the reference, Dan must find the most salient connection from the man, the chair, or the car to something that could be named Julius Marx. With anaphoric reference, the speaker refers to things mentioned earlier. Deirdre might say "Sean let the dog

off the leash and then put it in a drawer," using "it" anaphorically; Dan must recognize that she is referring to the leash and not the dog. With the proper noun Sean, Dan must determine which Sean she is referring to. And, finally, there are definite descriptions such as "the woman in the blue dress."

Language use, according to Austin (1962), is a form of action. When a speaker utters a sentence, he or she is indeed doing several things at once. Suppose Deirdre says to Dan, "Where is your car?" Among other things, she is performing the "utterance act" of issuing the utterance. She is also performing the "illocutionary act" of asking Dan where his car is. And she is performing the "perlocutionary act" of trying to get him to tell her where his car is. All of these acts are called speech acts. Most research has focused on illocutionary acts.

Illocutionary acts are of five main types (Searle, 1975). With assertives, speakers express a belief, as in saying "It's raining out." With directives, they try to get addressees to do things; commands, requests, and questions are types of directives. With commissives, speakers commit themselves to a future action; promises and offers are two types of commissives. With expressives, such as thanking, apologizing, congratulating, or greeting, speakers express certain feelings toward addressees. And finally, with declarations, speakers change some institutional state of affairs, as when judges sentence prisoners, referees call fouls, and ministers christen babies. Not all illocutionary acts fit neatly into these categories.

The issued is how speakers get addressees to recognize what acts they are performing. For illocutionary acts, the form of the sentence uttered is not enough. A request for the time, for example, can be made with a declarative sentence ("I'd like to know the time"), an imperative ("Please tell me the time'), or an interrogative ("What time is it?" or "Do you know the time?"). Which form is chosen depends on the speaker's purpose, status, and politeness.

What complicates the picture, Grice (1975) argued, is that speakers generally mean more than they say. When Deirdre asks Dan on the telephone, "Is Margaret there?" She may be asking him whether Margaret is there. She may also be *implicating* that he should call Margaret to the telephone if she is. Addressees derive such implicatures, according to Grice, by interpreting each utterance against the current purpose or direction of the talk exchange. They assume that the

speaker is cooperative and, therefore, is trying to be informative, accurate, relevant, and clear. They compute what the speaker must mean based on this assumption. Implicatures too take many forms.

In the final analysis, what people do with language is accomplished within a discourse, and the primary form of discourse is face-to-face conversation. For a conversation to be coherent, the participants must make sure their common ground accumulates in a systematic way, and that takes collaboration. References, illocutionary acts, and implicatures in conversation a really collaborative efforts. It is not enough for Deirdre to utter to Dan the phrase "the woman in the blue dress" or the sentence "I am now over here." The two of them must do what it takes to reach the mutual belief that he has understood who she is referring to, or what she means. That may require several turns as Dan gives Deirdre evidence of understanding and they correct any misunderstandings that arise (Clark & Schaefer, 1989). The pragmatics of language cannot be divorced from the study of discourse, especially conversation.

Self-focus and Self-attention

The concept of self is one of the oldest and most basic in all of psychology. The word pertains simultaneously to the sense of personal continuity that characterizes every individual's personality; to the organized body of knowledge that everyone has about who one is, what one feels and believes, and who one wishes to be; and to the subjective sense of being at the center of the experiences in which one is involved. The self also exhibits a curious and unique property that has been termed reflexivity. That is, the self has the capability of somehow turning backward on itself, of taking aspects of its own content and its own functioning as the object of its awareness. When a person is making use of this reflexive capability, that person is in a state of self-focused attention.

Background

The self's reflexive capability has captured interest at least since the time of William James. Though many eminent theorists made reference over the years to the potential importance of self-directed attention, the idea led to virtually no empirical work until, in 1971, Wicklund and Duval first published experimental research on the

possibility that self-directed attention might exert an influence on people's behavior and judgements. The success of these studies led to an expanded program of research and theoretical work on their part (Duval & Wicklund, 1972). These efforts sparked the interest of others, and a good deal of further work ensured.

Relationship to congitive psychology. The work described here was done by social and personality psychologists rather than cognitive psychologists. Unlike most researchers in "social cognition," those studying self-focus have only rarely adopted the research techniques of cognitive psychology and have examined aspects of behavior that are relatively molar rather than molecular. With few exceptions, the work has tended to focus on regulation of people's overt actions and on subjective judgements such as emotional states, attributions, and the like.

The theorists working in this area have also been less precise in their treatment of the concepts that underlie the construct of self-focus than most cognitive psychologists might prefer. That is, they have not fully articulated either theories of the self or theories of attentional processes. What is of interest is the consequence of directing more versus less attention to the self. No stand is usually taken on the question of whether attentional resources are divided or are instead timeshared. Indeed, even the usual assumption of a relatively fixed attentional capacity often seems pro forma.

Operationalizations. The typical procedure for increasing self-focus is simply to place subjects before a stimulus that serves to remind them of themselves: e.g. mirrors, TV cameras, or audiences. The introduction of the manipulation usually is timed so as to coincide with the point in the experiment at which a brief state of self-focus should in theory be most relevant to the behavior of interest. Although unorthodox, these manipulations do produce the psychological state they are intended to produce.

Another technique was also developed to vary self-focus in research. In 1975, Fenigstein, Scheier, and Buss published the Self-Consciousness Scale, a measure that captures individual differences in the tendency to spend time thinking about the self. The individual differences measured by this scale are indicative of differences in the ease or the frequency with which people slip naturally into states of self-focus.

Researchesrs who use both experimental and personality techniques have adopted a terminological shorthand to distinguish between the two operationalizations. In general, the term "self-awareness" is used to reflect experimental manipulations; the term "self-consciousness" is used to reflect individual differences. Terms such as self-focus and self-attention are used more indiscriminately. We have adopted this convention throughout our discussion here.

Cognitive Consequences of Self-Focus

What are the consequences of attending inward to the self ? Because the informational content that defines the self has many aspects, the phenomenology of self-directed attention can take several forms, and the consequences of self-focus can vary considerably. The effects of self-focus range, in fact, from very simple to very elaborate. We begin our discussion by addressing relatively simple effects, which we have characterized here as "congnitive" consequences of self-focus.

Salience and weighing effects. A very simple consequence of self-attention is to cause a self-aspect that is salient to seem phenomenologically more prominent. This tendency toward greater phenomenological prominence is reflected in several different ways, depending on what aspect of self is situationally salient.

In some circumstances, the most salient aspect of self is some transitory internal state. For example, when people have been exposed to an affect-inducing stimulus, affect tends to be talent as an aspect of the self's experience. When attention is self-focused after a feeling state has been created, the feeling is subjectively more intense. On the other hand, there are also occasions when people are led to expect internal states that fail to occur. In these situations, self-focus appears to lead to greater awareness of the *absence* of the expected condition. Thus, self-focused people are less likely to be misled about what is happening inside them.

Sometimes what is salient is not so much an internal state as a more global sense of the self as an entity that plays a role in events in the day to day world. In some instances, this awareness means that the self becomes more salient as a potential causal force. Thus, when people make causal attributions for hypothetically experienced events, greater self-focus seems to produce a greater weighting of the self as

a causal agent, and causes greater attribution of causality to the self (e.g. Duval & Wicklund, 1972).

If self-focus enhances the salience of the self's involvement in events, self-focus should also increase the extent to which people view themselves as being the target of others' actions. There is substantial evidence that this is so (Fenighstein, 1984).

Self-focus, self-knowledge, and articulation of self-knowledge. Does the scrutiny of self-focus provide a clearer or more accurate reading of stored information about the self? The implicit assumption underlying this possibility is that enhanced self-attention prompts a more thorough search of relevant memories. The result would be a more accurate self-report of one's prior actions, or of one's more general action tendencies. This hypothesis has been supported in a number of studies, using self-reports of such qualities as hostility, dominance, and sociability to predict actual behavior in subsequent sessions (reviewed in Carver & Scheier, 1981).

An additional derivation from this line of thought is that attending to the self over a long period should promote greater ariculation of self-knowledge (Nasby, 1985). Indeed, it is arguable that this is why self-reports of very self-conscious people are more valid than those of less self-conscious people. That is, having spent so much time thinking themselves, people high in self-consciousness have well developed and highly articulated schematic representations of themselves.

The idea that self-consciousness leads to better developed representations of the self leads to several predictions, two of which have been tested in research paradigms adopted from cognitive psychology. In one paradigm, subjects in an initial phase rate the self-descriptiveness of a large number of adjectives. In phase two, they view a subset of the adjectives, and indicate (yes or no) whether each is self-descriptive. After a distracter task, another set of adjectives is presented in phase three, some of which had been presented in phase two (old) and some of which had not (new). The task in phase three is to identify which items are old and which are new.

The tendency to misjudge new words as being old is termed the "false alarm" effect. It is most common among words rated initially as highly self-descriptive. This tendency is strongest among subjects high in self-consciousness. Among words previously rated as *not*

self-descriptive, in contrast, highly self-conscious subjects were more likely to correctly judge new words as new. Taken together, these findings are consistent with the idea that highly self-conscious persons have a more articulated representation of themselves than do less self-conscious persons (Nasby, 1985).

A similar conclusion emerges from an incidental learning (depth of processing) paradigm, in which subjects make a decision about each of a series of words, decisions ranging from superficial (involving only shallow processing of the word) to more complex (involving deeper processing). After a distractor task, subjects are given a surprise, recall test for the words in the first task. Incidental recall (reflecting deeper processing) is particularly high for words applied to the self in the first task (via the question "Does this word describe you?"), a finding that is commonly termed the "self-reference" effect. Hull and Levy (1979) found that the self-reference effect is stronger among subjects high in self-consciousness than among those lower in self-consciousness. Again, this finding is consistent with the idea that highly self-conscious people have highly articulate self-schemas.

Self-Regulatory Consequences of Self-Focus

A more complex set of consequences follows from self-focus when self-attention is invoked under other circumstances. Because these effects appear to have an impact on the guidance of behavior, we will characterize them here as self-regulatory in nature.

Attention to standards and adjustments in self-regulation. A person who is engaged in goal directed activity has adopted some value, standard, or point of reference as a guideline for his or her action. In these circumstances, self-focus appears to have two interrelated effects: it enhances the tendency to compare one's present state or action with the salient point of reference, and it promotes an adjustment in one's behavior so that the former comes into closer conformity with the latter.

Evidence that self-focus causes an increased tendency to compare one's current behavior with salient reference values is difficult to obtain. The difficulty stems from the fact that the postulated comparison is an internal mental check, a process of monitoring that takes place at an abstract level and is not directly observable. Indirect access to this even can be obtained, however, in situations where the

abstract comparison requires concrete information that is available to the subject but must be sought out. A series of studies based on this reasoning has yielded evidence that seeking of comparison-relevant information is increased by self-focus-before starting a task, while engaged in a task, and even after the task is finished (Scheier & Carver, 1983).

The comparison between behavior and standard is presumed to lead to beahvioral adjustments (when possible) that reduce discrepancies between the two values. This discrepancy reduction effect was one of the first consequences of self-directed attention to be shown experimentally (Wicklund & Duval, 1971) and it has since been obtained in numerous conceptual replications. Subjects follow instructions more closely, adhere more closely to salient norms, and adjust their actions more in response to systematic situational pressures, when they are more self-focused at the time than when they are less so.

Why does self-focus lead to discrepancy reduction? Two explanations have been offered, which differ substantially in their metatheoretical underpinnings. According to Duval and Wicklund's (1972) view, any discrepancy between one's current state and a salient standard creates an aversive drive state. The larger the discrepancy, the more aversive is self-focus and the stronger is the drive state. One option for reducing that aversive state is to reduce the discrepancy.

Carver and Scheier (1981) have proposed a different view (a view more compatible with the thrust of contemporary cognitive psychology): that the discrepancy reduction effect is a straightforward illustration of the process of feedback control. In a feedback loop, a sensor perceives a current quality. The perception is compared against a reference value. If a discrepancy between the two is discerned, an adjustment is made in a behavioral output. If the adjustment is appropriate, the subsequent perception of present state either will be closer to the reference value or will not deviate from it at all.

This, according to Carver and Scheier, is what happens when attention is self-focused in a setting where a reference value for behavior has previously become salient. People monitor the fit between what they see themselves doing and what they intend to be doing, and make adjustments as needed to diminish discrepancies between the

two. In this view, the discrepancy reduction effect is a natural aspect of the feedback-based self-regulation of action.

Self-focus and disengagement. Though self-directed attention promotes behavioral conformity to salient standards in many circumstances, sometimes the opposite effect occurs. This is most likely to occur when movement toward the goal seems beyond one's capabilities or is precluded by some aspect of the situation one is in. When a giving-up response is evoked, self-focus enhances the extent to which it is carried out. As a result, self-focus sometimes results in diminished persistence or even in withdrawal from the behavioral situation entirely (Carver & Scheier, 1981, 1986).

Carver and Scheier's interpretation of this sort of effect is that it involves a second critical element besides self-attention. In this view people continue to exert effort (and thus are facilitated by self-focus) until they become convinced that they are unable to attain the goal they are pursuing. Only when they pass this pyschological watershed of loss of confidence does the impetus to disengage begin to emerge and self-focus yield a decrement in behavior. Thus, the important co-determinant of this effect of self-focus is an unfavourable outcome expectancy (Carver & Scheier, 1986).

Sometimes the disengagement impulse that follows from these unfavorable expectancies is expressed indirectly, rather than overtly. In some behavioral settings, overt withdrawal is prevented by implicit social sanctions. In other settings physical constraints prevent behavioral disengagement. The Carver and Scheier position is that in these settings the disengagement impulse is often expressed covertly, via activities such as daydreaming or off-task thinking (Fig. I). If the person is supposed to be engaged in a task during this period, the off-task mentation can result in performance impairment.

The model displayed in Fig. I appears to be consistent with the results of research that examined the consequences of self-focus under circumstances of difficulty or adversity (Carver & Scheier, 1981, 1986, review issues in this literature; see Duval & Wicklund, 1972, for an alternative view). That is, even under conditions of adversity, if subjects have confidence of being able to overcome the adversity, self-focus enhances efforts and outcomes. If subjects are doubtful enough about attaining their goals, however, self-focus is associated with reduced efforts and impaired outcomes.

We should make one further point here, which deviates slightly from the theme of self-focus per se, but which also has implications for that theme. Disengagement (behavioral or mental) often is a viable response to impediments. Life would be impossible if people did not give up some of the goals they took up. On the other hand, it is difficult to disengage from a goal which is proving difficult or impossible to attain but to which one is very committed. In this situation, when attention is self-directed a cycle occurs. The person focuses on the comparison with the desired goal, assesses expectancies, finds them unfavorable, experiences the impulse to disengage, cannot execute it fully because of the importance of the goal, and is precluded from moving toward the goal by the disengagement impulse. Because the goal is important, it is repeatedly reconfronted and focused on. This cycling produces a phenomenology of self-deprecatory rumination and enhanced distress.

As a concrete example consider the situation of a test-anxious student taking an exam, whose most fervent with is to escape from the test room. Though he or she does not do that, the student lapses into off-task thinking as a way of disengaging mentally. This withdrawal of mental resources makes failure more likely. Moreover, even this mental disengagement cannot be sustained for long, as attention repeatedly returns to the test and thence to the unfavorable expectancies that induced the disengagement impulse in the first place.

It has long been recognized that test-anxious persons experience the resulting phenomenology of self-deprecatory rumination while in test settings (Wine, 1971). Though this phenomenology is often termed self-focus, we believe that such a label for that peenomenon is too simple and is potentially misleading (Carver & Scheier, 1986), because it implies that self-focus is antithetical to task focus. The fact is that self-focus impairs performance under some circumstances, but *enhances* performance under other circumstances – and does so even among the test anxious.

One final point, which stems from considering disengagement and discrepancy reduction effects together. We have characterized the difference these two classes of phenomena here in terms of differences in people's expectancies for the outcomes of their goal directed efforts. It is important to recognize that this characterization places these effects of self-attention squarely in the tradition of

expectancy-based analyses of motivation. From this viewpoint, self-focus effects are not a curious and somewhat esoteric side topic. Rather, they are embedded in a major perspective on behavior.

Public and private aspects of self. There is one more important issue to address before closing our discussion of self-directed attention. This issue stems from the multifaceted nature of the self. That the self is multifaceted should already be apparent, given the range of topics we have touched on thus far. However, the issue has also been treated in a more systematic manner than is apparent from our discussion to this point. More specifically, a distinction has been made between public self-aspects and private self-aspects (Fenigstein et al., 1975).

Private aspects of the self are self qualities that are covert, hidden, and bear on personal desires and preferences. Public aspects of self are self qualities that are overt, more explicitly open to others' scrutiny, and bear on self-presentational desires or communal values. This simple dichotomy appears to capture a fundamental and important division within the self, although it will be obvious that each of these self-facets can also be further subdivided.

Given this dichotomy among self-aspects, a dichotomy can also be imposed among instances of self-focus. At any given moment, it is possible to direct attention selectively to one or the other of these facets of the self. The consequences of self-attention can be expected to depend on which aspect of self is taken as the object of one's inward focus. This statement is simple, and the processes to which it refers are also simple, but the idea has far-reaching implications.

To go further, we must first consider how the distinction has been operationalized in research. The individual difference measure of self-consciousness incorporates the distinction in its design. Separate scales of the Self-Consciousness Scale measure tendencies to attend to public and private aspects of the self, respectively. The existence of separate scales implies that the two tendencies are not assumed to oppose each other. Indeed, the scales correlate positively, implying that people who tend to think about one aspect of self also tend to think about the other aspect of self. On the other hand, the correlation is usually not so high as to make it difficult to study the effects of each tendency separately.

Many researchers have now concluded (based partly on

convergence between effects of manipulations and effects of dispositions) that experimental manipulations of self-focus also embody the public-private distinction. Small mirrors, placed in such a manner as to make it clear that there is no hidden observer behind them, appear to make people selectively attentive to private aspects of the self. TV cameras, the sound of the subject's own voice, and the presence of observers (particularly evaluative observers) make people selectively attentive to public aspects of the self.

Behavioral consequences of the public-private distinction have been examined in a substantial number of studies. Of greatest interest is a set of studies that focused on situations in which the private self and the public self can be expected to exert *opposite* influences on behavior. Due to the juxtaposition of competing values in these situations, it is very important which self-aspect the person takes as focal. The results of these studies are strikingly uniform. When placed in situations that engender this sort of conflict, people act in accord with their private beliefs and perceptions to the extent that they are focused on their private selves. They act in accord with consensual preferences and self-presentational considerations to the extent that they are focused on their public selves.

Applications to the Conceptualization of Pathology

An outline of principles relating to self-attention should include at least some mention of two additional literatures. One of them deals with the effects of alcohol, the other with depression. Neither case is a simple application of the ideas already outlined – in each case the theorists have developed statements of their own – but both cases fit reasonably well with the ideas already outlined.

Alcohol and the reduction of self-awareness. The first of these literatures grows from the hypothesis that an important consequences of alcohol ingestion is a reduction in self-awareness. This hypothesis, developed by Hull (1987), accounts for a variety of observed effects of drinking. For example, if self-focus promotes self-regulation (as we argued earlier), a reduction in self-focus via alcohol should cause behavior to become more poorly regulated. Consistent with this, it is well known that people's behavior when intoxicated is more responsive to cues of the moment, and less responsive to personal values and principles, than is the behavior of the same people when sober.

More important, Hull and his colleagues have gone on to investigate directly the effects of alcohol consumption on various manifestations of self-awareness in a series of studies. In our view, the most interesting of the findings concerns the idea that people use alcohol as a vehicle for what Hull calls "the strategic avoidance of self-awareness." We said earlier that if people feel unable to attain desired goals, they feel an impulse to disengage from those goals. Diminishing self-focus via alcohol would produce an effect similar to successful disengagement. Moreover, this tactic can be used even when actual disengagement is precluded. This strategic use of alcohol tends to put certain people (those who have a strong motive to disengage) at an elevated risk for sustained drinking problems. The considerable support that Hull (1987) has obtained for his predictions make this theory an important candidate for further exploration.

Depression and self-focus. The second literature to be addressed here bears on depression. It has frequently been noted that depression is correlated with self-consciousness. Pyszczynski and Greenberg (1987) have proposed a mechanism by which that association comes to exist. In their view, depression develops when a person loses some important source of self-worth (thus having a large discrepancy between perceived self and desired self) and is unable to move toward regaining it. The person becomes immersed in self-deprecatory rumination (and concomitant distress) in the manner we outlined earlier.

As this pattern of negative self-focused thought begins to solidify and stabilize, the person's self-image becomes more negative, and the person tends to focus selectively on negative self-attributes. The result is a "depressive self-focusing style," in which the person is more likely to seek out self-awareness after failure than after success (Pyszczynski & Greenberg, 1987), a pattern that is opposite to that found among nondepressed persons. This selective focusing on the negative helps to maintain and even exacerbate the depression. Though the precise reason for this selective focus on the negative remains to be determined, this approach to understanding aspects of the phenomenology of depression seems a promising one.

2

Cognitive Science

Cognitive science refers to the interdisciplinary study of the acquisition and use of knowledge. It includes as contributing disciplines: artificial intelligence, psychology, linguistics, philosophy, anthropology, neuroscience, and education. As one might surmise from this list, the cognitive-science movement is far-reaching and diverse, containing within it several viewpoints.

General Aspects of Cognitive Science

Cognitive science grew out of three developments: (1) the invention of computers and the attempts to design programs that could do the kinds of tasks that human do; (2) the development of information-processing psychology, where the goal was to specify the internal processing involved in perception, language, memory, and thought; and (3) the development of the theory of generative grammar and related offshoots in linguistics. Cognitive science was a synthesis concerned with the kinds of knowledge that underlie human cognition, the details of human cognitive processing, and the computational modeling of those processes.

With regard to the overall architecture of an intelligent system, the traditional view has been inherited from on Neumann. Roughly, sets of symbols are moved about from one memory store to another, and are processed by explicit rules applied in sequence. In recent times, this "symbol-system view" has been best articulated by Newell and Simon (e.g. 1972); their emphasis has been on the components

of the human information-processing system, particularly working memory and long-term memory, and how information is stored and used in these memories to solve problems.

The major challenge to the rule-based symbolic architecture is a "connectionist architecture," sometimes called a "parallel distributed processing system" (e.g. Rumelhart, McClelland, & the PDP Research Group, 1986). Connectionists disavow that intelligent processes consist of the sequential application of explicit rules. Instead, they posit a large number of simple processing units, operating in parallel, where each unit sends excitatory and/or inhibitory signals to other units. In some cases, the units stand for possible hypotheses about the state of the world; in other cases, the units stand for goals and actions. In all cases, information processing takes place through the interactions of a large number of units.

A variety of methodologies has evolved in cognitive science. One methodology that is linked to AI is "protocol analysis" (e.g. Newell & Simon, 1972). In it, subjects are asked to think aloud while they solve a problem or do a task. Subsequent analysis attempts to specify the knowledge and processes needed to generate such protocols. A "computational model" is then usually constructed that replicates the problem-solving process manifest in the protocol. Other common methodologies are borrowed from experimental psychology, including memory studies and reaction-time analyses. Still other methodologies come from other disciplines, including, for example, discourse analysis (linguistics), dissociations caused by brain damage, and the kind of formal analyses characteristic of philosophy.

Topic Areas in Cognitive Science

Cognitive science can also be defined in terms of its topic areas. Reserach has centered on five major areas that are addressed below: knowledge representation, language, learning, thinking, and perception.

Knowledge representation. One of the earliest formalisms of representing knowledge about objects or concepts was the "discrimination net". A discrimination net is a tree structure; each branch point of the tree discriminates one set of objects or concepts from another on the basis of some feature, and the final leaves of the tree represent the objects or concepts themselves. To illustrate: knowledge about animals might be represented by a tree that contains

separate branches for birds, mammals, and so on, where the branch from birds contains subsequent breaches for songbirds, birds to prey, and so on, and the leaves emanating from songbird included species like bluejays, orioles, and so on. Such a discrimination net may be used in categorizing a novel animal by determining the values of the animal that correspond to the branch points in the tree, e.g. "It's feathered and flying, so it's a bird, it's small and sings so it's a songbird, etc."

Quillian (1968) introduced a "semantic network" as a formalism for representing concepts. In a semantic network, a concept is represented by a set of properties, which in turn consists of pointers to other concepts. The properties are made up on attribute-value pairs, so that an oriole might be represented by the pairs: superset = bird, color = orange, location = Eastern United States. Any amount of information can be represented in this format. Along with this representation. Quillian proposed "spreading activation" (or "marker passing") as a means for searching a semantic network. When asked, for example, "Does an oriole have skin?," the concepts corresponding to oriole and skin are activated, and the activation from both sources spreads through the network; when the activation from both sources intersects, a possible answer to the question has been found. A semantic network goes beyond a discrimination net in that it permits different kinds of relations, (or links), which are differentially labeled and which are critical in using the network or answer questions. These ideas inspired a host of models in both artificial intelligence and cognitive psychology.

Semantic networks evolved into even more structured representations. Minsky (1975) extended the idea of a concept-in-semantic network into the notion of a large organized data structure, or "frame," embedded in a "frame system." Each frame contains "slots," which accept only certain input values, and frequently have "default values" to use if there are no input values. The frame for a bird, for example, has slots for shape, size, and color which can be instantiated by any particular bird that one happens across. Similarly, the frame for a bedroom has a slot for a bed, which can instantiated by any object playing the role of bed, and also has as a default value the prototypical bed. Frames are essentially models of parts of the world; to understand input information is to determine which frame,

or configuration of frames, best fits it. An important variation on the idea of frame is the notion of a "script," which is a representation for a stereotypical action sequence like "going to a restaurant". Scripts contain defaults not only for the specific actions likely to be encountered in such a sequence, but also for the actors and props that are likely to figure in such a situation (Schank & Abelson, 1977).

Another important formalism for representing knowledge is a "production system" (Newell & Simon, 1972). Knowledge is represented as condition-action pairs called "production rules" (e.g. "If an animal has a long neck and brown blotches, then infer it is a giraffe"). A production system operates roughly as follows: if the current state satisfies one of the rules, the rule fires, which changes the current state, so that either a new rule fires or the system stops.

One final representational approach in the symbolic tradition, a mental model, specifies the domain-specific knowledge needed to understand a dynamic system or natural physical phenomenon, specifically the knowledge needed to qualitatively simulate the system or phenomenon. The key idea is that one can decompose a complex system (a door buzzer, for example) into a set of component models (switches, coils, and the like), whose characterization is independent of the particular system in which they are embedded (e.g. de Kleer & Brown, 1983). Then, to construct a qualitative simulation of the system, one must know two things: (I) the topology of connections between the various components (that is, what is connected to what), and (2) the incremental input-output functions of the various components (that is, if a particular input to a component goes up, what happens to the output).

The connectionist tradition has not focused on knowledge representation, because all knowledge is assumed to be represented in the same way – a neural-like network of nodes.

Language. Though psycholinguistic research is covered elsewhere in this dictionary, it has been a central research focus in cognitive science. We choose one topic from each of the three major divisions in language, semantics, and pragmatics.

Cognitive-science research on syntax often focuses on how a language understander parses a sentence into syntactic units like noun phrases and verb phrases. Such a parse is frequently a prerequisite for understanding the sentence's meaning. Parsing is complicated by

the fact that at numerous points during the process, more than one analysis is possible. Consider the sentence, "The French bottle smells." When the parser reaches the word "bottle," it has to decide whether this word should be grouped with the proceeding two to form the noun phrase "The French bottle," or whether "bottle" is the start of the verb phrase "bottle smell". A critical question is whether the human language parser considers both parses simultaneously, or instead commits itself to one analysis and then backtracks to the other only if needed. Other foci of interest are the extent to which linguistic constraints can be used to expedite parsing (e.g. Berwick & Weinberg, 1984), and whether syntactic analysis can be done solely by a connectionist architecture (e.g. Rumelhart, McClelland, & the PDP Research Group, 1986).

Research on semantics has emphasized the decomposition of words into more primitive meaning units. Thus, the verb "kill" can be composed into "cause to die," while "give" can be decomposed into "transfer of possession." One virtue of "semantic decomposition" is that it makes it clear why some verbs are related to one another (they share some components). What remains controversial, however, is whether the semantic decomposition of a word is a necessary step in understanding the word. Some experiments have found evidence for this link between decomposition and understanding, but others have not.

Work in pragmatics has been dominated by the "ordinary language" philosophers. Particularly important in Grice's (1975) analysis of implicit of conversation. According to Grice, participants in a conversation adhere to a "cooperation principle", which means, among other things, that the speaker tries to make his or her contribution relevant to the aims of the ongoing conversation. When a speaker violates this "relevancy criterion," a listener assumes that the communication is still cooperative, and draws inferences that will make the speaker's utterance relevant. To illustrate, suppose a host says to a guest, "Do you want some coffee?," and the guest replies, "Coffee would keep me awake." The direct meaning of the guest's reply is irrelevant to the host's question; but the host makes it relevant by using knowledge that generally people do not want to be kept awake, which leads to the inference that the indirect meaning of the guest's response is "No".

Learning. Research on learning in cognitive science involves both machine learning and human learning. There are three traditions in machine learning: similarity-based learning, explanation-based learning, and neural-net learning. Similarity-based learning is concerned with the question of how a system that is presented multiple exemplars of a concept and multiple non-exemplars can induce a rule that distinguishes exemplars from non-exemplars. Explanation-based learning induces a concept-membership rule from only a single exemplar by using detailed knowledge about the domain. Perhaps the most important strand of research in neural-net learning grew out of earlier work on learning in perception systems, in particular the "delta rule." The delta rule assumes there is training or feedback with respect to what the output should be from each node in a neural (or connectionist) net. If the expected output for a node is greater than the actual output, then the rule increases the weight of any inputs to the node that were acting to increase its output (and decreases the weight of inputs acting to decrease the output). Similarly, if the expected output for a node is less than the actual output, the rule decreases the weight of any input to the node that acts to increase its output.

Most cognitive-science research on human learning has focused on learning procedures. There are three competing "architectures" for such learning. Brown and VanLehn (1980) postulate that learning occurs only at "impasses." They looked extensively at human errors in subtraction and found that when students are doing a problem that their subtraction procedure cannot handle, they reach an impasse. At that point they invoke one of the repair (or problem solving) procedures to deal with the impasse. This repair may lead to constructing a correct procedure, but more often it leads to constructing an incorrect procedure.

Anderson (e.g. 1983) developed a three-stage model of procedural learning within his ACT ("Adaptive Control of Thought") theory. The theory uses semantic networks to represent "declarative knowledge" ("knowing that"), and Production Systems to represent "preocedural knowledge" ("knowing how"). In the first, or interpretative stage of procedural learning, people use declarative knowledge to solve problems, much like a computer program is used as data by a computer-language interpreter. The second stage is the knowledge compilation stage; declarative knowledge is converted into

procedural knowledge, and productions that occur repeatedly are compose by joining together pairs of productions and proceduralized by instantiating variables. The third stage, tuning, involves refining this procedural knowledge by means such as generalization and discrimination.

Laird, Newell, and Rosenbloom (e.g. 1987) have recently proposed a new architecture for human learning that is implemented as a computer model, called Soar. It is a problem-solving system built on production rules. When it encounters an impasse in problem solving, it sets up a subgoal to deal with that impasse. If it succeeds, it stores the solution as an operator that it uses when the same impasse occurs again . Like Brown and VanLehn's theory, the system learns at impasses, but in the context of a general problem-solving system.

Thinking. Cognitive-science research on thinking has centered on problem solving. Almost any cognitive activity can be construed as an instance of problem solving, but the kinds of situations labeled "problem solving" typically require putting together novel sequence of processes to achieve some difficult goal. Such situations include mathematical puzzles, games like chess, and scientific problems.

Newell and Simon (1972) view problem solving as a search through a "problem space." The space consists of "states" with paths between them. The initial gives in a problem (e.g. the opening board position in chess) constitute an "initial state"; the solution of the problem (e.g. the taking of the opponent's king in chess) constitutes the "goal state"; all possible states of the problem constitute the other states in the problem space; and transitions between states constitute permissible "operators." Problem solving consists of searching through the space for a path or sequence of operators that takes one from the initial state to the goal state.

Almost inevitably the problem space is sufficiently large that blind search is useless (thousands if not millions of paths would have to be checked). The problem solver therefore needs to employ some kind of "heuristic search." Newell and Simon (e.g. 1972) have proposed a number of such "heuristic strategies," the best known of which is "means-ends analysis".

Perception. Most cognitive-science research on perception has centered on visual processing and imagery. Some of the leading work has been inspired by Marr (e.g. Marr & Poggio, 1979). This research

concerns the early stages of perception, and constructs computational models of how the visual system uses retinal information (points of varying brightness) to construct two-dimensional and then "two-and-a half-dimensional" descriptions of input objects. (Their two-and-a-half-dimensional sketch represents what is in front of what, but omits three-dimensional aspects of the objects).

Other researchers have concentrated on the later stages of perception, in particular on matching input descriptions (like those mentioned above) to stored representations of objects. Some of the most seminal work has dealt with simple perceptual objects, namely letters and words. In the McClelland and Rumelhart (1981) model, for example, matching is implemented via a connectionist network using top-down and bottom-up activation and inhibition processes, operating over separate levels for features (i.e. line segments at different orientations), letters, and words. The model explains a host of empirical findings, and has fostered the development of connectionist models in general.

The basic idea guiding imagery research is that imaging is like perceiving. For example, studies of mental rotation have found that the time people need to decide that two forms are identical increases regularly with the rotational differences between the forms. This result suggests that people are "mentally rotating" one of the forms to "see" if it coincides with the other. But the idea that imaging is like seeing has proved controversial. Pylyshyn (1981), for example, argues that imagery consists of nothing more than the use of general thought processes to simulate physical or perceptual events, based on tacit knowledge of how physical events unfold. For example, "mental rotation" occurs because people are *thinking* (as opposed to *imagining*) in real time about the course of a physical rotation; they know implicitly the time needed to physically rotate an object, and they wait the corresponding amount of time before indicating their response. The debate about imagery continues, and has led proponents of imagery to turn to neuropsychology and the study of brain-damaged patients for evidence that imagery and perception are mediated by the same brain structures.

Cognitive Therapy

An increasingly important approach to therapy, especially with

mood-disordered patients, is based upon the notion that treatment can usefully focus on producing changes in patients' cognitive systems. Major cognitive therapists such as Beck (1976), Ellis (1962), and Meichenbaum (1977) all agree that patients suffering from the main mood disorders of anxiety and depression have unduly negative and self-defeating thoughts about themselves and their circumstances, and that producing major positive changes in these thought patterns can lead to recovery.

Historically, the views of Ellis (1962) are important, because he was one of the first therapists to propose a systematic. According to his approach, unhappiness occurs as the result of a three-point sequence. Point A is the occurrence of some unpleasant event (e.g. rejection by a valued other), point B is the cognitive reaction to that unpleasant occurrence (e.g. "I must be worthless to be rejected in this way"), and point C is a state of great anxiety or depression. In other words, anxiety or depression does not occur as a direct result of unpleasant events; instead, it occurs because of the irrational thoughts that are triggered off by the occurrence of unpleasant events.

The rational-emotional therapy developed by Ellis (1962) involves eliminating the irrational and self-defeating thoughts which all of us experience some of the time, and replacing them with more rational and positive thoughts. As Ellis (1962) expressed it, "If he [i.e. the individual] wants to be minimally disturbable and maximally sane, he'd better substitute for all his absolutists *It's terrible's* two other words which he does not parrot or give lip-service to but which he incisively thinks through and accepts–namely, 'Too bad!' or 'Tough shit'!"

The approach to cognitive therapy adopted by Meichenbaum (1977) focuses on the patient's internal dialog, and in particular on the negative or unhelpful things which he or she says to him - or herself when problems are encountered. These negative thoughts need to be replaced by more adaptive self-talk, in which the patient actively attempts to cope with the situation by talking to him- or herself about the appropriate actions required to achieve the immediate goals. This form of self-talk makes it easier to succeed, and can have the further beneficial effect of leading to positive changes in the information stored in memory.

Beck (e.g. Beck & Clark, 1988) has proposed a somewhat more

complex theoretical framework as the basis for cognitive therapy. He argues that there are important differences between the thought processes involved in anxiety and in depression. Anxious individuals have a heightened sense of vulnerability, and are concerned about physical and/or psychological threats. In addition, their negative appraisals of people and situations tend to be rather specific, and their thoughts focus on possible negative events in the future. In contrast, depressed individuals are concerned primarily about loss, and they have negative attitudes toward themselves, the world, and the future. Their negative thoughts tend to be more pervasive and global than those of anxious patients, and are often focused on past events.

In Beck's congnitive therapy, the first step is to make the patient aware of his or her negative and irrational thoughts. When this has been achieve, the therapist endeavors to challenge these negative thoughts, often by presenting the patient with new and incompatible information. This can involve presenting the patient's problem in a different way, by suggesting different strategies of handling the patient's current problems, or by re-analyzing the patient's previous difficulties.

There have been several attempts to assess the efficacy of cognitive therapy as a form of treatment. While there are many comlexities in assessing recovery, equating groups receiving different forms of therapy in terms of severity, and so on, the evidence generally suggests that cognitive therapy is at least as effective as most other forms of therapy, and possibly more effective. However, even if this evidence is taken at face value, there are two major interpretative problems. First, most forms of cognitive therapy involve a mixture of several different ingredients. For example, cognitive therapy typically makes use of encouragement and persuasion, various different techniques for altering irrational thought patterns, and assignments designed to alter behavior in key situations. As a consequence, it is usually extremely difficult to know which particular ingredient or ingredients are responsibility for any recovery which is observed.

Second, behavioral treatments (sometimes known as 'behavior therapy') have sometimes been contrasted with cognitive therapy. However, while it is true that the emphasis in such treatments is more on changing behavior than on changing thinking and other aspects of the cognitive system, the distinction is relative rather than absolute.

For example, as Brewin (1988) points out, behavioral treatments usually begin by offering the patient encouragement and a detailed rationale for the proposed treatment. Such ingredients in behavioral treatments closely resemble some of the ingredients incorporated into cognitive therapy. It is a matter of current controversy as to the precise relationship between behavior and cognitive therapy. Some cognitive therapists (e.g. Meichenbaum, 1977) would go so far as to claim that most behavior therapy is effective because it produces beneficial effects on cognitive functioning, in spite of the assertion by many behavior therapists that they are interested only in changing behavior rather than cognition. What is certainly becoming increasingly clear is that the similarities between behavior therapy and cognitive therapy are greater than used to be believed, and that cognitive interpretations of the effectiveness of behavior therapy are often at least as plausible as the traditional interpretations based on conditioning principles.

Imagery and Perception

Recent scientific studies on mental imagery have greatly increased our understanding of the relation between imagery and Perception. These studies have helped to identify the spatial, perceptual, and transformational properties of images, how imagery interacts with perception, and how visual discoveries can be made using imagery.

Scanning Mental Images

The spatial structure of mental images has been explored using the technique of mental image scanning. Kosslyn, Ball, and Reiser (1978) reported that the time it takes to imagine scanning between locations on imagined maps is proportional to the actual distances between the locations. A similar relation between scanning distance and response time has also been shown for distances among imagined objects in three-dimensional space. Such findings suggest that mental image scanning corresponds to the actual scanning of perceived objects and their features.

Mental image scanning enables people to judge, from memory, whether an object lies in a particular direction with respect to a new, unexpected location. In such cases, image scanning is used spontaneously; people report imagining that they are scanning from one location to the other along the indicated direction. This implies

that image scanning is not merely an artefact of instructions to deliberately simulate an actual physical scan. Rather, mental images appear to represent the true spatial properties of objects and their configurations.

Distribution of Spatial Attention

Another way in which imagery resembles perception concerns the manner in which attention is distributed within an imagined or presented pattern. Using a probe-detection task with real and imagined block letters, Podgorny and Shepard (1978) found that reaction times to verify that the probe is presented on some part of the letters varies with probe location in the same way in imagery and perception conditions. In particular, reaction times are most rapid, in each case, when the probe falls on the intersection of barshaped segments of the patterns, and as the overall compactness of the patterns increases.

Mental Image Acuity

Imagery also resembles perception with regard to constraints on resolution for fine visual details. Kosslyn (1975) reported that the time it takes to locate features of imagined objects increases as the features decrease in size. This finding could not be explained in terms of the association strength between the features and the object; large, poorly associated features can be detected more quickly than small, highly associated features.

As in perception, the visual field in imagery appears to be restricted, because images will "overflow" when the image is formed at too large a size. Direct comparisons between the visual fields in imagery and perception reveal a close correspondence between the two; the fields are the same shape, and increase in size in the same way as the features of the object or pattern become easier to distinguish. Criticisms of studies on image acuity, however, have suggested that experimenter bias may be responsible for at least some of these findings.

Imagery After- Effects

There is some evidence that imagery can lead to the adaptation of certain types of feature analyzing mechanisms in the visual system. Finke and Schmidt (1977) reported that orientation-specific color after-effects could be induced when bar gratings were visualized in association with actual colors presented during the adaptation procedures. These after-effects were not obtained, however, when colors were visualized in assocation with actual bar gratings, suggesting that there are limits on the extent to which imagery might involve the visual system.

Subsequent findings have raised questions about whether imagery can actually lead to the adaptation of visual mechanisms sensitive to bar gratings. There are important differences between orientation-specific color after-effects obtained using real and imagined gratings, and there have been failures to obtain imagery-induced after-effects using testing procedures that reliably reveal the perceptual after-effects. Moreover, prior visualization of a bar grating fails to produce orientation-specific changes in sensitivity in a subsequent grating detection task. Taken together, these findings suggest that imagery probably does not involve the activation of visual mechanisms responsible for the early processing of color and feature information.

Visual-Motor Adaptation

Visualizing errors of movement can lead to adaptive changes in visual-motor coordination that correspond to changes resulting when errors are actually observed. Finke (1979) reported that prism adaptation could be achieved by simply visualizing the same kinds of errors that prisms normally create. Imagery adaptation procedures also result in visual-motor after-effects resembling those following actual prism adaptation .

Image Facilitation of Perceptual Processes

Imagining a pattern can facilitate the detection or identification of the pattern. Farah (1985) reported that letters can be detected more easily when the letter is concurrently visualized in the correct location. Similarly, a tone can be recognized more accurately when exactly the same tone is imagined. This type of facilitation may result from the Priming of perceptual pathways during imagination.

Imagery can also facilitate perception by providing helpful visual contexts. Objects can be detected more quickly when people imagine visual scenes that are semantically compatible with the objects. Line length discriminations can be made more rapidly and accurately when

a helpful surrounding pattern is mentally superimposed over the lines. This facilitation resembles that obtained using actual context patterns.

Interference Between Imagery and Perception

Under certain conditions imagery can also interfere with perception. Segal and Fusella (1970) found that the detection of faint visual and auditory stimuli is impaired whenever images are formed in the same sensory modality. Images can also function as pattern "masks," interfering with detection of the pattern. Whether imagery facilitates or interferes with perception probably depends on a number of factors, such as the intensity of the stimulate, the type of perceptual judgement, and the extent to which the imagined and presented object correspond.

Perception can also interfere with visualization. Brooks (1968) showed that performance on difficult imagery tasks, such as identifying the shape of successive corners of block letters, worsens when the responses are given by pointing to visually presented characters, compared to giving the responses verbally. These interference effects, however, are due mainly to the spatial properties of images, rather than to their visual characteristics.

Reality Monitoring

Imagery and perception can sometimes be confused when one attempts to real whether or not an event actually occurred. These errors in Reality Montioring increase the more frequently the event is imagine, and the more easily the images are formed. In general, reality monitoring depends on the perceptual qualities of the memory, the context of the memory, and the amount of cognitive effort used in establishing the memory (Johnson & Raye, 1981).

Mental Rotation

A considerable number of studies have shown that imagined rotations of objects correspond in many respects to actual rotations. Shepard and Metzler (1971) found that imagined rotations are used to verify the equivalence of shape between three-dimensional objects that, in perspective drawings, are depicted as being rotated with respect to each other. The imagined rotations are performed at a constant rate, regardless of whether the rotations occur in the picture plane or

in depth. In each case, the response times are proportional to the angular separation distances.

Subsequent studies on mental rotation have shown that the imagined rotations are holistic – that is, performed on images of the entire object or pattern; exhibit analog properties – that is, are approximately continuous; and reflect biomechanical constraints on the degrees of freedom of natural motions. In addition, mental transformations other than rotation have been reported; these include imagined transformations of size and color. Eye movements can be a possible artefact in some of these studies, but this is a problem only in those studies where the stimuli are presented simultaneously.

Some practical applications of mental rotation include identifying rotated characters, such as letters of the alphabet, detecting the symmetry of rotated patterns, distinguishing right and left turns, and judging relative directions from rotated perspectives.

Representational Momentum

Imagined transformations, such as mental rotation, exhibit a property analogous to the momentum of physically moving objects. Freyd and Finke (1985) reported that memories for the final position of a pattern that had been depicted as rotating are forwadlly shifted, in the same direction as the implicit rotation. Like physical momentum, these memory shifts increase as the implied velocity of the pattern increases. The memory shifts appear to have a spontaneous onset, and represent continuous changes along the path of motion. This effect, called "representational momentum," suggests that mental transformations have a ballistic, inertial quality.

Image Reinterpretation

Like actual visual patterns, mental images can sometimes be reinterpreted, yielding novel, unexpected discoveries. Pinker and Finke (1980) found that people could imagine rotating a configuration of objects and then "detect" familiar patterns "formed" by how the objects would look from different viewing perspectives. The recognized patterns could not be anticipated from the way the objects appeared from the initial viewing perspective.

However, there are limits on the extent to which mental images can be reinterpreted. Structurally "hidden" figures in patterns can

rarely be detected using imagery, suggesting that images are constrained by structural descriptions initially given to the patterns. In addition, there is evidence that classic ambiguous figures cannot be perceptually "reversed" in imagery. On the other hand, unexpected patterns can often be recognized following the imagined construction and transformations of a familiar, starting pattern, or when people are given a randomly chosen set of parts to use during creative mental synthesis. Whether image reinterpretation is possible may therefore depend on the type or level of perceptual ambiguity depicted in the image.

Conclusions

Imagery appears to resembly perception in a number of important respects. Like actual objects and patterns, images can be scanned, rotated, inspected, and assembled. In addition, images exhibit other characteristics in common with perceived objects, such as spatial extent, limited resolution, and inertia. There are also important differences between imagery and perception. Imagery appears not to involve levels of the visual system where information about simple visual features is initially extracted. Nor can images the structurally reinterpreted to the same extent as perceived objects.

With regard to their practical implications, images can be used to improve perceptual performance, modify visual-motor coordination, verify spatial relations among objects, compare rotated objects, and discover patterns made up of novel combinations of features. Future studies in imagery research will most likely focus on these, and other practical uses of the perceptual properties of images.

Learning Styles

The term "learning style" has been used in the literature in two distinct ways: (1) it has been used to indicate a broad description of relatively consistent behaviors related to ways of going about learning; it is treated as an individual difference of generality comparable to intelligence or personality, but describing consistency in the ways people tackle learning tasks; (2) the definition has been narrowed considerably to parallel the idea of cognitive style, with the use of bipolar traits, but described in relation to the learning tasks commonly found in educational contexts, as opposed to scores on psychological

tests. Both uses of the term imply that learning style is related both to cognitive processes and to personality, but in the broader definition the emphasis is more cognitive, while the narrower definition is closer to personality. In both senses, the term has been used to cover a range of concepts which have emerged from attempts to describe aspects of student learning.

Processes, Skills, Tactics, Strategies, and Styles

The next step in clarifying the meaning of the term learning style is to see it in relation to other terms describing cognitive activities arranged in ascending order of generality. "Cognitive processes" are basic cognitive activities taking place within the memory, involving coding or thinking processes, which are usually investigated within laboratory experiments. When these processes are described in terms of the ability of individuals consistently to carry out certain types of task, such as certain psychological tests, they are referred to as "cognitive skills." In everyday life, skills are brought into play in order to solve particular problems or to deal effectively with certain situations. Here a decision is required to select particular skills, or to apply processes in succession, within some overall plan. This can be seen as adopting "tactics." And if an organized series of tactics is required, perhaps depending also on a person's attitudes and motives, then these may be called "strategies." Finally, if a person tends to adopt a similar set of strategies consistently across different tasks and settings, this can be taken to indicate the existence of "learning styles". Schmeck (1988, pp. 3-20) provides valuable additional clarification of these and other related terms. Unfortunately, in the literature there is still considerable overlap in the use of this set of terms, not always consistent with the level of generality indicated here. For the purposes of clarity, however, these distinctions are useful in looking at the research into learning style.

Cognitive Learning Styles

Research into cognitive learning styles has developed from attempts to predict different levels and types of academic performance in students. Some of the earliest attempts at describing different learning styles came from the attention given to divergent

thinking by J. P. Guilford. The tests of divergent thinking or "creativity," such as the Uses of Objects Test, which were subsequently developed encouraged comparisons between scores on conventional intelligence tests and those on the new open-ended tests. As a result, Wallach and Kogan (1965) and Hudson (1966) drew attention to the existence of two distinct groups of children. Those with much higher scores on intelligence than creativity were labeled "convergers," while those with the reverse pattern of scores were called "divergers." Hudson reported that these two groups showed markedly different pallerns of thinking, contrasting subject choices in school, and even varying personality configurations. The divergers were seen as being more impulsive, emotional, expressive, and humorous, while the convergers were more logical, rational, consistent, restrained, and conventional. Wallach and Kogan argued that the existing educational system appeared to favor the convergers and to penalize the divergers. Such arguments had a considerable impact on education at the time, encouraging the movement toward informal or open education, and also introduced researchers to the idea of individual differences in which cognitive differences were paralleled with personality differences. This pattern later became the hallmark of learning style.

Kolb (1983) also used the terms convergers and divergers, but with a rather different meaning. He has distinguished four learning styles based on two sets of distinctions – whether people prefer to deal with concrete or abstract information, and whether they process if reflectively or actively. Within this terminology, the converges prefer abstract material and process it atively, while the divergers look for concrete information and process it reflectively. The remaining two styles are described as assimilators (abstractions processed reflectively) and accomodators (active, concrete). There are parallels here with C. G Jung's "psychological types" and with the distinctive thinking processes attributed to extraverts and introverts.

For as in the former case the purely empirical heaping together of facts paralyses thought and smothers their meaning, so in the latter case introverted thinking shows a dangerous tendency to coerce facts into the shape of its image, or by ignoring them altogether, to unfold its phantasy image in freedom.

(quoted by Entwistle, 1981, p. 186)

Again we see the connection between cognitive processes and personality.

A more recent set of distinctions can be found in the work of Biggs (1987), who has used ideas derived from information-processing theories to investigate the ways in which students learn and study. His three categories also describe characteristically different strategies for dealing with information, but these are related to the intentions and underlying motives of the learner. The styles here are actually referred to as "approaches to learning" due to the similarities which were demonstrated with the concept introduced by Marton (Marton, Hounsell, & Entwistle 1984;) on the basis of interview with students. Biggs's categories, however, were identified through factor analyses of inventories of study behavior and learning processes which showed the consistent and close link between strategy and motive. A "deep" approach brings together intrinsic and competence motivation with a learning strategy which involves the attempt to understand the meaning of what is being learned. An "achieving" approach is rooted in competition and ego enhancement and leads to a strategy which depends on well-organized study methods. Finally, a "surface' approach is driven by fear of failure and leads to a dependency on reproduction through rote learning.

Although all three of these descriptions of learning styles do contain both cognitive and non-cognitive components, or correlates, the descriptions put more emphasis on their cognitive origins. In contrast, the next set of concepts uses the narrower definition which emphasizes the bipolar nature of learning style.

Bipolar Learning Styles

The research discussed in this section shows a more direct relationship to ideas in cognitive styles. In distinguishing between abilities and cognitive styles Messick (1976, pp. 9-10) wrote that:

Abilities are value directional: having more of an ability is better than having less. Cognitive styles are value differentiated: each pole has an adaptive value....[which] depends upon the nature of the situation and upon the cognitive requirements of the task in hand....Abilities are specific to a particular domain of content or function. Cognitive styles, in contrast, cut across domains. They appear to serve as high-level heuristics that organize lower-level

strategies, operations, and propensities – often including abilities – in such complex sequential processes as problem-solving and learning.... In factor-analytic terms, if abilities and temperament traits are first or second-order factors, then cognitive styles are higher-order factors linking those domains.

Cognitive styles have most commonly been measured through perceptual tests such as the Embedded Figures Test or the Matching Familiar Figures Test. In contrast, learning styles have been identified either by observing students working on learning tasks, or from inventory scores. In Biggs's categories, only students adopting a deep approach could be expected to reach a thorough understanding of what they were learning. Pask (in Schmeck, 1988) found that even when students were required to demonstrate understanding, they still showed differing strategies in the way they tackled the learning materials. Based on extended work in this area, he has concluded that some students have relatively strong and consistent *preferences* for adopting a particular type of strategy, which he takes to indicate their "learning style."

Some students adopt a "holist" style in which, right from the start , they try to see the task in the widest possible perspective, establishing an overview which goes well beyond the task itself. Their learning process involves the use of illustrations, examples, analogies, and anecdotes in building up an idiosyncratic form of understanding deeply rooted in personal experience and beliefs. Other students prefer a "serialist" style in which they begin with a narrow focus, concentrate on details and logical connections in a cautious manner, and look at the broader context only toward the end of the topic, if at all. Extreme holists are often impulsive, even cavalier, in their use of evidence, tending to generalize too readily and to jump to unjustified conclusions. Extreme serialists are often too cautious, failing to see important relationships or useful analogies, thus leaving their understanding impoverished.

These two main categories of description do, however, need qualification, as there are different ways of exhibiting stylistic preferences. Pask himself (in Schmeck, 1988, pp. 92-3) has described these differences as follows:

Among all holists it is possible to recognize (a) those who depend upon the use of valid analogies or generalizations given within the material ...., and (b) those holists who depend upon the creation of (their own) valid analogies and generalizations between topics. This latter method incidentally seems to be one of the most productive methods of learning - by invention or discovery.... Among serialistic learners, it is possible to distinguish at least the following two types. One is the operation, or local rule, learner who progresses logically, step-by-step, moving to a different context only when he or she has assimilated one portion thoroughly and often asking for tutorial guidance on which topics to tackle next. The other is the rote learner, who follows the prescribed and narrow path. Unless the subject matter is trivial or the learner virtually has Eidetic Imagery, rote learners have little chance of success in making use of the knowledge they have acquired, although they may perform quite competently in multiple choice or other types of examination where special memorization skill places them at an advantage....

Pask follows Messick in arguing that learning styles are value differentiated. For some tasks a serialist strategy will be more effective, and for others a holist one. Students who have a strong preference for one or other style will find it difficult to shift strategy between different kinds of task. Pask showed that students learn more effectively from materials deisgned to match their particular learning style, while it seems likely that teachers adopt a teaching style which reflects their own learning style (Entwistle, 1981, pp. 231-9, 250). There thus seems to be a considerable advantage in being able to adapt readily to different presentational styles, adopting what Pask describes as a "versatile" style. It may be possible to help students to become more versatile by requiring them to practice their weaker style in some tasks, but stylistic preferences have been found to be unexpectedly strong (Pask, in Schmeck, 1988).

The strength of these preferences has been explained either in terms of underlying personality traits (Messick et al., 1976) or in terms of cerebral dominance of left (serialist) or right (holist) hemispheres of the brain, which differences Sperry (1983) has argued are inherited. This latter hypothesis underpined the extensive work on learning styles conducted by Torrance, again with its origins in work on the measurement of Creativity. Torrance has argued that creative thinking depends on the effective use of both sides of the brain, and sees learning styles as reflecting the predominant use of one or other

side.

Research on brain hemisphericity indicates that the left cerebral hemisphere is specialized primarily for verbal, analytic, abstract, temporal and digital operations.... [while] the right hemisphere [is] primarily specialized for non-verbal, holistic, concrete, spatial, analogic, creative, intuitive, and aestheitc functions.

(Torrance & Rockenstein,in Schmeck, 1988, p.278)

Torrance's inventory identifies people in terms of whether they rely on right or left brain functions, or whether they integrate those functions by using both hemispheres equally. Torrance, like Pask, spells out the dangers of over-reliance on one style and suggests teaching methods designed to encourage "whole-brain creativity."

Elsewhere in the psychological literature there is a less enthusiastic endorsement of the link between hemispherically and distinguishable thinking processes, with the current view being that there is no absolute specialization of function (Cohen, 1983). However, the parallels which emerge between the descriptions of Pask and Torrance suggest that whatever their physiological roots, there is good evidence that learners prefer to rely on distinctively different styles of learning, and that these styles have cognitive, motivational, and personality components. There have been some recent attempts to provide an integrated view of learning styles.

Integrated Descriptions of Learning Styles

In work paralleling that of Biggs (1987), Entwistle and Ramsden (1983) have described very similar factors derived from an inventory. They had, however, included scales of Pask's learning styles in their inventory and were able to show connections between approach, style, and motivation. The intrinsic deep approach was associated with a holist style while the anxious surface approach was linked with a serialist style. There was an additional fourth "non-academic" factor with low levels of motivation, negative attitudes, and with extreme learning styles linked to learning pathologies, as anticipated by both Pask and Torrance. Entwistle and Ramsden were also able to relate the two main approach/style factors to measures of cognitive style and personality to show that deep holists were likely to be impulsive "thinking introverts" with a theoretical outlook, aesthetic interests,

and complex conceptualization, who did well on tests of divergent thinking. The surface serialists had low scores of several of these dimensions. Linked to a strong practical outlook.

Schemeck (in Weinstein, Goetz, & Alexander, 1988, p. 175) provides a tentative integrative model of learning style, which includes three distinctive learning styles, described by Schmeck as deep, elaborative, and shallow. This terminology had its roots in an inventory which contained items derived from the ideas on depth of processing in the memory, but subsequent research has persuaded Schmeck that there are important personality correlates as well. Essentially, the stable introvert is seen as adopting a *deep* style which develops effective schema and conceptions, through articulated or field-independent thinking processes involving predominantly analytic and synthetic reasoning. The stable extravert is more likely to use an *elaborative* style which involves an impulsive personalizing of knowledge through global or field-dependent thinking, making substantial use of examples and concrete instances, often drawn from real-world experiences. Finally, the anxious individual is likely to be categorized as having a *shallow* style which depends on memorizing, or repetitive rehearsal of information, leading to the literal reproduction of what was studied.

A similar but less developed model had been produced earlier by Entwistle (1981, p. 113). Now it seems worthwhile to try to combine the two models as a summary of the current state of thinking about learning styles. Table 1 presents that model. Again it has to be seen as tentative because the empirical findings are not sufficiently strong, or coherent, to allow it to be treated as more than a heuristic to guide thinking about learning styles, and future research. On the whole, this model fits quite well with that of Schmeck, with one important exception. Schmeck suggests that an elaborative style will be associated with stable extraversion. This style is, however, what Jung described as introverted thinking, and empirical findings show a positive correlation of introversion with holism, while there is a negative correlation with serialism (Entwistle & Ramsden, 1983, p. 235). The implicit association between inversion and inpulsivity, and caution and extraversion, contradicts previous research. Further research will be necessary to clarify this apparent contradiction.

There remains considerable debate about the nature, and even the existence, of learning styles, certainly in relation to their possible hemispehric origins. There is, however, substantial evidence that people do show strong and relatively consistent preferences for tackling learning tasks in distinctive ways. If that is what is taken to be learning styles, it seems important to take account of them in both research and educational practice.

3

Cognitive Development

This term refers to the changes which occur to a person's cognitive structures, abilities, and processes during the course of that person's life-span. The cognitive changes which occur during adult aging are dealt with elsewhere in this volume the present entry will focus instead on the cognitive changes which occur during childhood.

Piaget's Stage Theory of Cognitive Development

The most comprehensive theory of children's cognitive development to have been proposed is that put forward by Piaget (e.g. Piaget & Inhelder, 1960). Piaget interpreted cognitive development as consisting of the development of logical competence. He argued that the development of this competence proceeds through a sequence of four major stages (sensori-motor, preoporational, concrete operation, and formal operational). Piaget proposed that, within any given stage, cognitive performance is not task specific but is homogeneous and structurally equivalent across a range of different tasks, with transitions from one stage to the next being marked by relatively abrupt but generalized qualitative changes in cognitive structures and performance.

However, more recent studies have cast some doubt upon Piaget's account of cognitive development. In particular, it has been found that children's cognitive performance is not homogeneous in the manner proposed by Piaget; instead, performance can vary considerably as a function of changes in task materials and

administrative procedures. For example, children's performance on visual perspective taking tasks varies as a function of the stimulus array which is used; performance on conservation tasks varies depending upon whether the child is asked to produce either a single judgement or two judgements in close succession; performance on drawing tasks varies as a function of the wording of the verbal instructions which are given to the child; and performance on transitive inference tasks varies depending upon whether or not the child is given the opportunity to memorize the premisses properly. Furthermore, on all of these different tasks, it has been discovered that young children are capable of functioning at a level which Piaget claims is possible early at a later age, as long as the appropriate task materials and procedures are employed.

Task Analysis

This finding, that children's performance varies as a function of task factors, implies that much greater attention needs to be paid to these factors if an adequate account is to be given of children's cognitive development. Task analysis also helps to elucidate the range of cognitive faculties and abilities which the child requires in order to perform successfully on any given task.

Piaget's classification task can be used to illustrate this point. In this task, the child is shown a set of flowers, some of which are primroses, and is asked the question "Are there more flowers, or more primroses?" (The original aim of this task was to test whether the child can conserve the quantity of the superordinate class when comparing it with the quantity of the subordinate class). It should be noted that, in order to answer the question correctly, the child must be able to: (a) comprehend the syntax, semantics, and pragmatics of the question which is being asked; (b) store the question's intent in memory; at least for the duration of the execution of the task; (c) perceptually discriminate the primroses from the other flowers; (d) assign the discriminated primroses to the category that was designated by the word "primroses" in the question; (e) quantify the number of priomroses; (f) hold the number of primroses in working memory; (g) assign all of the objects in the array to the category that was designated by the word "flowers" in the question; (h) quantify the number of flowers; (i) hold the number of flowers in working memory;

(j) compare the two numbers which are stored in working memory; (k) apply an appropriate decision rule (such as "Choose the larger number unless both numbers are equal"); (l) encode the output of that decision rule in a linguistic format for verbal production.

In short, for a child to succeed on this kind of problem, he or she must possess sufficient linguistic knowledge, adequate perceptual abilities, a large enough working memory, basic cognitive processes (such as storage and retrieval), suitable decision rules, and sufficient real-world knowledge (such as natural object categories). And notice that failure on this task could be due to a failure of any one of these subcomponents (and need not be due to a failure to conserve the quantity of the superordinate class). Thus, task analysis implies that cognitive development should be construed not merely as the acquisition of an abstract logical competence (as Piaget proposed) but as the acquisition and development of language, perception, working memory, cognitive processes, decision rules, and real-world knowledge. Language Acquistion and Perceptual Development are dealt with elsewhere in this volume; consequently, the remainder of this essay will focus upon the other four topics in this list.

Working Memory

It is known from use of the digit-span task (which entails asking children to listen to a sequence of numbers, and then to repeat them immediately back to the tester) that children's working memory improves considerably as they get older: 3-year-olds only have a digit span of two items, 5-year-olds have a span of four items, 12-year-olds have a span of six or seven items, while students have a span of about eight items. Pascual-Leone (1970) has interpreted this increase as reflecting an increase in the size of the child's "central computing space M." According to Pascual-Leone, M space is divisible into two components, e and k: e is the space which is necessary for storing the task instructions, and for scanning the perceptual array; k is the space which is available for storing information while executing the task. It is hypothesized that the size of e remains constant during development, while the size of k increases with age; hence older children are able to handle tasks that require more information to be stored than younger children.

An alternative point of view has been advocated by Case (1985).

He also argues that the child's total processing space is divisible into two components: operating space (which is that space which is devoted to executing current cognitive operations) and shor-term storage space (which is where the products of previous operations are stored). Unlike Pascual-Leone, though, Case argues that total processing space remains constant through development. Instead, what changes with age is the child's operational efficiency; that is, cognitive operations are executed more efficiently as the child gets older. Consequently, the operating space nees to occupy less of the total processing space with increasing age; as a result, more of that total space becomes available for use as short-term storage space.

Case (1985) presents strong evidence to support the notion that operational speed increases with age. In addition, there is good evidence that the acquisition and practice of mnemonic strategies such as rehearsal produce substantial improvements in children's short-term memory performance. However, it should be noted that the conclusion that operational efficiency improves with age does not rule out the possibility that total working memory space also increases with age, and that the superior performance of older children is a consequence of both increased capacity and superior efficiency.

Cognitive Processes

Basic cognitive processes such as storage and retrieval are already functional at birth. This conclusion is supported by the fact that neonates exhibit habituation; that is, if a neonate is repeatedly presented with the same stimulus, he or she will spend progressively less time looking at it, and if a different stimulus is then shown, attention will recover. Notice that for habituation to occur, the neonate must have stored the initial stimulus, and must be retrieving it and comparing it with the current stimulus on each subsequent presentation. When there is a match between the previous and the present stimulus, habituation occurs; when there is a noticeable discrepancy between the two stimuli, dishabituation occurs. Thus, these types of basic cognitive processes are present from the very outset of infancy, and are probably inherent in the cognitive system itself.

However, this is not to say that basic cognitive processes do not exhibit any changes during the course of development. As we have already seen, Case (1985) presents evidence which suggests

that efficiency and speed of execution improve with age. Such improvements could be caused either by biological maturation or by practice and experience.

Decision Rules

The acquisition and development of decision rules in children have been studied extensively by Siegler (e.g. 1976), who proposes that cognitive development in some problem-solving domains can be explained entirely in terms of the progressive acquisition of increasingly powerful decision rules. This proposal can be explained most clearly by reference to the balance-scale task, which Siegler (1976) has investigated in detail. The apparatus for this task consists of balance scale, which has four positions to which variable numbers of weights can be attached on either side of the fulcrum; the arm of the balance tips either left or right or remains horizontal, depending upon the precise arrangement of weights on either side of the fulcrum. On any particular trial, the arm of the balance is locked into a horizontal position, an arrangement of weights is placed on both sides on the fulcrum, and the child is asked to predict how the balance will behave when the lock is released.

Siegler postulates that peformance on this task can be based upon one of four different decision rules. Rule I takes into account only the number of weights on each side of the fulcrum: if they are the same, the balance will remain horizontal; if they differ, the side with the greater number of weights will descend. Rule II also relies exclusively on weight if the two sides have different number of weights; however, if the numbers of weights are equal, then the distances of the weights from the fulcrum on each side are also taken into account. Rule III considers both weight and distance, and solves the problem correctly if one or both are equal; however, if one side has more weights but the other side has its weights further from the fulcrum, then the child is reduced to guessing. Rule IV involves the calculation of torque (weight × distance) for each side, and predicting that the side with the greater torque will descend.

By presenting a child with a range of different problems, and by examining the pattern of correct and incorrect answers which are given across this range of problems, it is possible to work out whether or not the child is following one of these four rules and, if so, which

particular rule is being followed. Using this procedure, Siegler (1976) found that 90 percent of 5- to 17-year-olds perform in accordance with just one of the four decision rules, and that there is a trend toward the use of the more complex rules with age. In addition, Siegler found that children's explanations of how they make their decisions is consistent with the rule which is postulated on the basis of their actual performance.

Siegler's rule assessment approach has since been applied to a wide variety of different problem domains (e.g. the projection of shadows, probability, conservation of liquid, and the Tower of Hanoi), and it has been found that the cognitive-developmental changes which occur on all of these different problems can be explained in terms of a progression through a sequence of increasingly powerful decision rules.

Real-World Knowledge

During the course of their development, children acquire a vast range of real-world knowledge, including knowledge of objects, events, relations, people, social institutions, and so on. Children's knowledge of objects and events constitutes two areas that have received particular attention from developmental psychologists; the present discussion will be limited to these.

Knowledge of Events

Nelson (1986) is the most influential theorist to have examined children's event knowledge. She argues that, during the course of their development, children acquire and elaborate mental representations of entire events. These event representations (ERs) have four main characteristics. First, they specify sequences of acts that are linked together temporally (and possibly causally as well) which are appropriate to particular contexts (e.g. a meal-time ER might specify the acts of "sitting," "being served," "eating," and so on). Second, ERs specify the people who are involved in the constituent acts, and the roles which these people occupy in those acts. Third, ERs specify the objects which may be involved in the constituent acts. Finally, ERs may contain slots which can be variably filled by different people or objects, in which case they also specify the range of alternative slot-fillers for each particular slot.

This theoretical formulation has generated a variety of studies. In one study, for example, Nelson (1978) examined 4- to 5-year-old children's descriptions of the events which occur when eating at home, at a McDonalds, and at a daycare centre. These descriptions contained a regular temporal sequencing of the constituent acts, and they included mentions of both obligatory and optional elements. In another study, Gruendel (1980) studied ERs in children aged between 4 and 8 years of age. This study revealed that as the children got older, they included more optional alternative elements and paths in their event descriptions, indicating that development entials the incorporation of greater complexity and flexibility into ERs.

Studies investigation children's memory for stores have shown that children's ERs can influence their recall of information. Nelson and Guendel (1981), for example, presented 4- to 5-year-old children with stories that followed a familiar sequence of acts, which the children then had to recall. In some of these stories, one act was presented in the wrong temporal position. During recall, this act was either omitted or put back into its proper position in the sequence. This suggests that the children's recall was being influenced by their ERs, which provided them with normative information about correct temporal sequencing.

Finally, it has been found that even 1-year-old children have already acquired some ERs. This conclusion emerges from studies which have investigated the very early linguistic productions of 1-year-old children. These studies reveal that many early words are produced only in the context of highly specific events. This type of early word use can be explained in terms of the children having mapped these words on to ERs, and consequently confining their use of these words to situations in which those represented events occur.

Knowledge of Objects

The main influence upon the study of children's knowledge of objects has been Rosch's prototype theory. This theory postulates that an object category is represented mentally by means of a prototype, which is a specification of the most typical and clearest example of the kind of object which is included within the category. Actual objects are then included within that category if they share attributes with the prototype. Depending upon how many attributes are shared with the

prototype, the object will be a more or less typical member of the category. Because category membership is determined by the number of attributes which are shared with the prototype, and because more typical category members share many attributes with the prototype while less typical members share fewer attributes with the prototype, this theory predicts that the degree of typicality exhibited by a category member will influence the cognitive processing of the object. In addition, the theory predicts that while typical category members (which all share many attributes with the prototype) should have many attributes in common with each other, atypical category members (which each have relatively few attributes in common with the prototype) need not have any attributes in common with each other. Thus, categories should display a family resemblance structure (rather than a structure defined by a set of necessary and sufficient features).

A variety of studies has been conducted with children in order to test the first of these two predictions (that children's cognitive processing of objects is influenced by their degree of typicality). These studies have produced strong support for this prediction. Mervis and Pani (1980), for example, found that it is easier for 5-year-old children to acquire a new category from exposure to a typical category exemplar (which bears a high degree of resemblance to the prototype) than from exposure to an atypical category exemplar. Similarly, Kuczaj (1982) found that when 2-year-old children are asked to select an object (e.g. "a ball") from an array of objects, they initially select highly typical exemplars, and only subsequently go on to choose less typical exemplars if they are asked to select further category members.

The second prediction, that children's object categories should display a family resemblance structure, has also been confirmed in several studies. Barrett (1982), for example, has shown that the full range of objects to which a 1- to 2-year-old child refers by means of an object name does not necessarily display any common attributes which are shared by all of the category members; however, all of the objects within the category have at least on attribute in common with one of the child's initial referents (the prototype) for that object name. Similarly, Kuczaj (1983), in a detailed scrutiny of the range of real and possible objects that 2- to 5-year-old children are prepared to include within the category "airplane," found that these objects are not linked together by a set of necessary and sufficient attributes; instead, this

category displays a family resemblance structure.

Constraints On Cognitive Development

Finally, it should be noted here that, in recent years, several authors have proposed that the course of cognitive development is held within particular bounds by a set of very specific constraints. It is argued by these authors that children sometimes master enormously complex cognitive domains with such apparent ease that their progress must be facilitated by constraints which eliminate, right from the outset, many of the potential errors which these children might otherwise make. To give a specific example drawn from the domain of word learning, Markman and Hutchinson (1984) point out that, when an adult directs a child's attention to a particular object (e.g. a black dog) and utters a word (e.g. "dog"), there is an indefinite range of conclusions which the child could in principle draw from this event (for example, that "dog" is the name for dogs in general, that "dog" is the proper name of this particular dog, that "dog" means "black," that "dog" means "four-legged," and so on). However, Markman and Hutchinson argue that children only consider the one correct conclusion (that "dog" is the name for dogs in general), the reason for this being that they are limited by a taxonomic constraint which restricts them to the assumption that words must refer to categories of similar objects. Other constraints which have been proposed in the domain of word learning are the natural category constraint, which leads children to assume that individual words must refer to either objects or events but not both, the principle of contrast, which leads children to assume that all words must contrast in meaning, and the principle of mutual exclusivity, which leads children to assume that each object has one and only one name. In addition, some authors (e.g. Markman & Hutchinson, 1984) suggest that these constraints may not be learnt by the child during the course of development, but might be innate.

However, Nelson (1988) has drawn attention to several problems with this approach to cognitive development. For example, she raises the question of whether these proposed constraints actually constrain the child, or whether they represent instead just the pro-cessing biases and strategies of the child. True constraints should be manifested as all-or-none behavior; biases and strategies would be manifested instead merely as statistical trends. In fact, the evidence indicates the presence

of trends rather than all-or-none behavior; that is, all of the proposed constraints are actually violated with considerable frequency by children during the course of their development. Also, Nelson points out that if these constraints are to be of maximum benefit to the child, they ought to be operative very early in development when the child requires the maximum assistance. However, the evidence suggests that these proposed constraints only become operative after the child's development has already commenced by other means, a finding which is more consistent with the view that these constraints are merely processing biases or strategies which are acquired during the course of development.

Nevertheless, the issue of constraints raises some crucial questions about the nature of children's cognitive development, questions about the extent to which the child's cognitive development proceeds via domain-specific processing or via the application of more general cognitive processes, and the extent to which there may be specific innate inputs to children's cognitive development. These questions are currently the focus of much debate by development psychologists.

4

Cognitive Dissonance

As originally proposed by Festinger (1957), cognitive dissonance theory deals with the consequences for behavior and attitude change of a person experiencing a sense of inconsistency or contradiction ("dissonance") between two or more cognitions. Dissonance is conceived of as a noxious state that motivates cognitive change aimed at eliminating or reducing such inconsistency. Where cognitive dissonance theory differs most from other attitude theories that incorporate notions of cognitive consistency is in its concern with cognitions related to behavioral decision. Thus, the main issue addressed in cognitive dissonance research has been that of how people deal with thoughts or information implying that they have made a wrong decision. In terms of the theory, dissonance arises in this context from a person holding the belief "I performed action A" at the same time as the belief" Action A was inconsistent with my attitude." Since it is typically less easy to deny a previous action than to find justification for it, dissonance may be resolved by changing the latter belief to "Action A was consistent with my attitude" or at least "Action A was excusable."

A large number of experiments have examined this hypothesis by using a procedure termed "induced" or "forced compliance." The idea is to induce subjects to perform an action assumed to be inconsistent with their prior attitudes, and then observe any shift in their attitudes following their action. Thus, Festinger and Carlsmith (1959) had subjects perform a boring task and then inform another

person apparently waiting to perform this same task that it was enjoyable. Where subjects were offered only a small reward for giving this misinformation, they gave higher ratings of their own enjoyment of the task than did others who had been offered a large reward. The interpretation is that the former group resolved dissonance by bringing their previous attitudes into line with their subsequent behavior. For the latter group, the high reward may have led them to regard their behavior as more excusable, or simply have provided a compensation for the unpleasantness of the experienced inconsistency, thereby reducing its motivational force. Whatever the interpretation, the prediction is that *lesser* rewards produce *greater* attitude change within this paradigm. A parallel prediction is that attitudes will be held more strongly when associated with more rather than less costly commitment, as in the case of members of a cult prophesying the end of the world who persisted with their proselytizing even after the day of doom had safely passed (Festinger, Riecken, & Schachter, 1956).

These predictions only hold, however, under certain limiting conditions, the most important of which is that subjects should see their decision as having been freely made. Under conditions of free choice, one finds greater attitude change (in the direction of reduced inconsistency with behavior) with lesser rewards; under conditions of low choice, greater rewards produce greater change (Linder, Cooper, & Jones, 1967).

Subsequent research has questioned Festinger's assumption of dissonance as a noxious motivational state, produced by inconsistency and removed (typically) by attitude change. Bem (1967) has argued that subjects' reports of their own attitudes are not derived from any such motivational processes, but are merely self-descriptions of behavior that take into account situational factors such as incentives and constraints on free choice, and are no different in principle from the descriptions that might be offered by another person observing their behavior. An observer would infer a closer match between a person's attitudes and behavior if the behavior were performed for a small rather than large reward. Objections of a different kind have been raised by Nuttin (1975) who proposes that, in forced compliance studies, dissonance is *produced* by inequitably low rewards rather than *reduced* by high rewards.

There appears to be a renewal of sympathy for Festinger's theory

in contemporary attitude research, helped particularly by evidence that procedures designed to influence dissonance produce changes in physiological arousal, and that arousal that cannot be attributed to some other cause (e.g. a drug) can lead to attitude change (Cooper, Zanna, & Taves, 1978). The idea of dissonance as a state of tension or unpleasant arousal is therefore quite plausible. On the other hand, there is less support for the idea that what produces such tension is merely the experience of an inconsistency between different cognitions. According to Cooper and Fazio (1984), dissonance should be thought of more as an emotional reaction to the knowledge that one has been responsible for an action that has produced unwanted consequences. Cognitions that reduce this sense of responsibility (e.g. that one acted under compulsion, or that the consequences could not have been anticipated) will reduce the extent to which one "feels bad" about one's behavior, and hence the motivation to re-examine one's beliefs. However, this "feeling bad" depends on the consequences of one's behavior rather than on the holding of contradictory beliefs.

Cognitive Maps

A cognitive map is a mental representation of some area of one's environment. It might be a representation of the layout of the interior of one's house, one's neighbourhood, or even one's country or the world. It is sometimes useful to speak of cognitive maps because people can use their mental representations of their environment in ways similar to their use of maps. Thus, for example, in a well known environment, people can plan their routes in advance, take short cuts, and, in general, move around in a sensible and economical way. However, the term cognitive map is potentially misleading. It is very unlikely that the processes which allow us to navigate around familiar environments employ any of the same processes as are involved when we navigate with a real map.

When we move to a new town we begin to acquire spatial knowledge about our environment as we move around and develop memories of the routes that we have followed. Landmarks within the town help us to interlink these routes so that we gradually develop the ability to represent our environment two dimensionally. We begin to know how the separate routes are interrelated and the ways in which one would take short cuts between them (Appleyard, 1976; Evans,

Marrero, & Butler, 1981). Despite the success with which people familiar with a location can navigate around it, there is considerable research that suggests that the mental representations that they use are not as exact as those to be found on Ordnance Survey maps. Byrne (1979), for example, found that long-term residents in a town give longer estimates for routes within the town center and for those routes with many changes in direction. He argued that the number of landmarks and turns contributed to the judgement of distance. He also found that when people were asked to draw the angles with which roads joined, they tended to draw them as metting at right angles even though this could involve an error of 30°.

The tendency to think of roads as meeting at right angles, found by Byrne, may reflect a general tendency for mental representations of the environment to be simplified into north, south, east, west alignments. When asked which is the further west, Bristol or Edinburgh, most British people say Bristol since they know that Bristol is on the west coast while Edinburgh is on the east coast. Most people's mental representations of the British Isles ignore the fact that Scotland leans to the west of England so that Edinburgh is, in fact, west of Bristol (Moar, 1979; Tversky, 1981). Similarly, Stevens and coupe (1978) found that Americans make, on average, an error of 67° when asked to indicate the direction of Reno, Nevada, from San Diego, California. California appears to be imagined as stretching north-south with Nevada to the east. In reality, California lies in a north-west to south-east angle so that Reno is north-north-west of San Diego.

Cognitive Neuropsychology

Cognitive neuropsychology studies the effects of brain injury on such things as language, memory, perception, and attention. Most of the research is done with adult subjects who have suffered a cerebral vascular accident (cva or stroke), tumor, missile wound, or other form of brain injury, but work is also done on developmental disorders much as developmental dyslexia.

Cognitive neuropsychology has two main aims (Ellis & Young, 1988). The first aim is to explain the patterns of impaired and intact performance seen in brain-injured patients in terms of damage to one or more of the components of a theory or model of normal cognitive functioning. Thus, a cognitive neuropsychologist interested in reading

would study patients with acquired dyslexia (impaired reading following brain injury), and would seek to explain the different patterns of reading breakdown in terms of damage of different components of an information-processing model of skilled reading. Similarly, a cognitive neuropschologist interested in Aphasia or Amnesia would endeavor to explain the different patterns of language and memory breakdown in terms of impairment to different components of a psychological theory of language processing or learning and memory respectively. To the extent that one can speak in cognitive neuropsychology of "localizing" the deficit in a given patient, that localization is within an information-processing model of cognitive processes and takes the form of saying that the patient has suffered damage to the direct visual procedure for word recognition, syntactic processes, episodic long-term memory, or whatever. This does not mean, though, that cognitive neuropsychologists are necessarily disinterested in which brain structures mediate which aspects of cognition.

The second aim of cognitive neuropsychology is to draw conclusions about the nature and organization of normal, intact cognitive processes from the patterns of impaired and intact capabilities seen in brain injured patients. In other words, cognitive neuropsychologists assert that studying neurological patients is a valuable alternative or adjunct to carrying out experiments on normal subjects where the objective is to understand how the normal mind works. Thus, it is common in cognitive neuropsychological papers to see claims to the effect that it would not be possible to observe a particular pattern, or set of patterns, of impaired and preserved cognitive performance following brain injury unless the cognitive system in the normal, intact brain were organized in a particular way.

Modern day cognitive neuropsychology bears quite a close resemblance to work done in the late nineteenth and early twentieth century by neurologists and aphasiologists such as Wernicke, Lichtheim, Charcot, and the young Sigmund freud. In recent times, cognitive neuropsychology was revived in the late 1960s. The work of Marshall and Newcombe at the Radcliffe Infirmary, Oxford, and Shallice and Warrington at the National Hospital for Nervous diseases, London, was particularly influential in convincing the world that cognitive neuropsychology is a worthwhile enterprise. Shallice and Warrington (1970) studied a patient whose pattern of preserved and

intact abilities could, they maintained, be explained if one argued that he had suffered damage to his auditory-verbal short-term memory, while other forms of short-term memory and access to long-term memory were preserved. The patient had a very reduced digit span and performed badly on all tests of short-term memory involving letters, digits, or words, yet his immediate recall or nonverbal material (e.g. shapes or sounds) was intact, as was his ability to commit lists of words to long-term memory. Marshall and Newcombe (1973) described three different forms of acquired dyslexia, and argued that their occurrence was compatible with some models of reading but not others.

Much of the strength of cognitive neuropsychology derives from the fact that many of its practitioners entered the area after training as experimental cognitive psychologists. This means that cognitive neurospychological studies employ the same experimental paradigms, and analyze their data using the same statistical methods, as are seen in mainstream cognitive psychology. Indeed, many cognitive neuropsychologists alternate between case studies of neurological patients and laboratory experiments with normal subjects, arguing that some questions about how the mind works are best answered by case-studies while other questions are best answered by laboratory experiments.

Traditional neuropsychology studied groups of patients considered to belong to the same "syndrome" group (e.g. "Broca's aphasia" or "Wernicke's aphasia"), usually comparing the group's performance with that of either another patient group or normal subjects without brain injuries. The widespread belief among cognitive neuropsychologists is that the conventional syndrome groups are too broad, encompassing patients who differ one from another in theoretically important ways. The trend in cognitive neuropsychology has been to carry out intensive investigations of single cases, or at least to pay close attention to individual differences between patients.

Like all approaches in psychology, cognitive neuropsychology rests on a set of assumptions (Caramazza, 1984; Shallice, 1988). It is assumed, for example, that the mind is composed of a large number of semi-independent processing components or *modules* which are separately represented in the brain and are therefore capable of separate impairment. It is also assumed that the pattern of impaired and intact

performance seen in a patient, including the pattern of errors, will provide a reasonably transparent guide to the nature of the underlying deficit. Finally, it is commonly assumed that although modules may be lost through brain injury, new modules are not developed in the mature brain.

Any assumption can, of course, be wrong, but one of the attractions of cognitive neuropsychology is that its assumptions are to some extent different from those upon which conventional cognitive psychology rests. Where research in the two approaches converges upon the same conclusion regarding an aspect of the mind's structure and function, one can be doubly confident in the validity of that conclusion.

5

Mood Disorder and Cognition

The major mood disorders are clinical anxiety and clincial depression. Of course, these are broad categories, and it is possible to subdivide them. Clinical anxiety has subcategories such as generalized anxiety disorder, social phobia, obsessive-compulsive disorder, panic disorder with agoraphobia, panic disorder without agoraphobia, agoraphobia without panic disorder, and post-traumatic stress disorder. Clinical depression has been subcategorized in various different ways. However, two of the most influential distinctions are between unipolar and bipolar depression and between endogenous and non-endogenous depression.

There are major differences in symptomatology among the various subcategories of clinical anxiety and depression. However, it is likely that there are important similarities among different types of clinical anxiety, with the same also being the case with respect to clinical depression. As a consequence, the broad categories of clinical anxiety and clinical depression will often be used in the subsequent discussion.

Patients with mood disorders have been found to differ from normal controls in several ways. These differences occur at the physiological, behavioral, and cognitive levels. Are there any persuasive reasons for assuming that there is particular value in focusing on the functioning of the cognitive system in mood disordered patients? Lazarus (e.g. Lazarus, Kanner & Folkman, 1980) has consistently claimed that cognitive appraisal is an essential ingredient in the

experience of mood or affect. If that is correct, then that would appear potentially to give the cognitive system an extremely important role in describing and explaining mood disorders. However, Zajonc (1980) denied that affect or mood necessarily occurs only after cognitive operations such as recognition and classification, and he cited experimental evidence which appeared to support his position. Close examination of such evidence reveals that he considered only cognitive processes of which there is conscious awareness, and he failed entirely to consider the possibility that affect or mood depends to some extent on prior preconscious processing. In sum, the available evidence is entirely consistent with the notion that cognitive processes precede the experience of mood or affect.

There is a somewhat puzzling state of affairs that exists when we consider congnitive factors in mood disorders. On the one hand, there are numerous cognitive therapists (e.g. Beck & Emery, 1985) who regularly treat mood disorders from the cognitive perspective. On the other hand, there was until recently remarkably little hard evidence available about the ways in which the cognitive functioning of mood-disordered patients differs from that of normals. In other words, Cognitive Therapy for mood disorders has developed largely on the basis of clinical experience rather than being built on a solid foundation of research evidence. However, there is increasing interest in the possibility of studying cognitive functioning in anxious and depressed patients under laboratory conditions. The results of such studies, as we will shortly see, offer hope for enriching cognitive therapy in the future.

A further reason for considering cognitive functioning in depressed and anxious patients is to clarify the relationship between these two kinds of mood disorder. Anxiety and depression are often found to coexist in patients, but it is still usually claimed that they are conceptually distinct. As we will see, a comparison of the cognitive profiles of depressed and anxious patients provides substantial new evidence concerning the differences between depression and anxiety.

Basic Findings

Much of the research on cognitive factors in depression has been influenced by the theoretical views of Beck (1976). He argued that depressed patients possess various cognitive schemas (schema

= organized collection of knowledge stored in memory). These schemas primarily concern the self, and encompass themes such as personal worthlessness, self-blame, and guilt.

This schema theory has been investigated primarily by an experimental paradigm in which subjects make self-referent judgements (e.g. describes you?) about a series of adjectives, some of which are negative and others of which are positive. It is assumed that the existence of a relevant self-schema enhances recall, perhaps by providing a rich network of associations which can facilitate the retrieval process. Since depressed patients allegedly have several negative self-schemas, whereas normal controls do not, it is predicted that depressed patients should recall relatively more negative words relevant to their negative self-schemas than normal controls on this self-referent task. However, if the same words are judged in terms of how well they apply to someone else, then the depressed patients' self-schemas should not be activated, and a negative recall bias should not be found.

There have been numerous studies testing these predictions from Beck's (1976) schema theory. While there are some exceptions, the general finding is that depressed patients have a negative recall bias on self-referent tasks but not on other referent tasks. Sometimes depressed patients differ from normal controls more in terms of reduced recall of positive self-referent adjectives than enhanced recall of negative self-referent adjectives, but there is clear evidence that the balance between positive and negative is shifted in the negative direction in depressed patients.

Another expectation from Beck's (1976) schema theory is that the interpretation of ambiguous information would tend to be influenced by an individual's underlying schemas. This means that depressed patients would be more likely to provide a negative interpretation of ambiguous information than would normal, especially if the information were of personal relevance. The results generally conform to expectation. There have been a number of studies using the Cognitive Bias Questionnaire, which provides a description of events together with four possible responses. The response which is selected may be considered to reflect the extent to which the event in question was interpreted in a negative fashion. Depressed patients consistently select more negative alternatives than normal controls. However, this self-report measure provides only a rather indirect assessment of the interpretative process, and it fails to exclude the possibility of response bias.

Some studies have considered the implications of Beck's schema theory for attentional and perceptual processes. The general expectations are that negative self-schemas in depressed patients would enhance attention to, and initial processing of, relevant negative stimuli. Despite the fact that there have been various experimental attempts to investigate these predictions, there is very little (if any) convincing evidence in their support.

In sum, the available evidence indicates that depressed patients differ in their memorial functioning from normal controls, and probably also in their interpretation of ambiguous information. However, they do not appear to show systematic differences from normals in their attentional and perceptual processing.

The evidence on cognitive functioning in anxious patients (primarily those with generalized anxiety disorder) and normal controls has been discussed at length by Eysenck (in press), and only a succinct account of the general trend of the research will be given here. An orienting assumption is that anxious patients are hypersensitive to threatening stimuli, and so differences in cognitive functioning between anxious patients and normals are more likely to be found with respect to the processing of threatening stimuli than with respect to neutral stimuli. Since physical health threats and social threats are the major areas of concern for generalized anxiety patients, most of the research has used threat-related words selected from these two domains. Anxious patients typically perform equivalently with physical health and social threat words, and so the findings will usually be described simply in terms of threat.

Anxious patients have fairly consistently been found to differ from normal controls in attentive and preattentive processes. For example, if a threatening and a neutral stimulus are presented concurrently, then anxious patients show a selective bias in favor of the threatening stimulus. There is additional evidence that anxious patients devote extra processing resources to threat-related stimuli even when these stimuli are presented below the level of conscious awareness.

Anxious patients also differ from normal controls in terms of

distractibility. Parients with generalized anxiety disorder are more distractible than normal when a target stimulus must be detected as rapidly as possible in the presence or absence of distracting stimuli. The effect is there whether threatening or neutral distracting stimuli are used, but it is somewhat greater with threatening distractors.

Anxious patients differ considerably from normal controls in their interpretation of ambiguous stimuli. For example, one paradigm requires subjects to write down the spellings of auditorily presented homophones having a threatening and a neutral meaning (e.g. die; dye). Anxious patients consistently interpret such stimuli in a more threatening fashion that normals.

The major focus in studies of memorial functioning in anxious patients has been on the negative recall bias which has been shown so clearly in depressed patients. There have been numerous attempts to demonstrate a negative recall bias in patients with generalized anxiety disorder, but these attempts have been uniformly unsuccessful. The reason for this is not known, but it is possible that anxious patients adopt the strategy of inhibiting processing of negative personality relevant material.

In sum, anxious patients differ from normals in their preattentive and attentional processing of threatening stimuli, and in their interpretation of ambiguous information. However, they show no evidence of a negative recall bias along the lines of depressed patients. In general terms, non-normal cognitive functioning in anxious patients primarily relates to the early stages of information processing, wheras non-normal cognitive functioning in depressed patients relates more to later stages of processing. The findings for anxious patients are consistent with the view that the function of anxiety is to facilitate the early detection of potential threats.

The Causality Issue

The fact that anxious and depressed patients differ from normal controls in several different aspects of cognitive functioning indicates that a cognitive approach to the mood disorders is variable. However, the findings which have been discussed so far fail to resolve an important theoretical issue. In essence, it is possible that non-normal cognitive functioning is found in anxious and depressed patients as a consequence of their current mood state. Alternatively, it may be that

non-normal cognitive functioning plays some role in the development of mood disorders. In other words, there may be a cognitive vulnerability factor, in which non-normal cognitive functioning increases the probability that a given individual will subsequently develop a mood disorder. Realistically, of course, it may be the case that some aspects of non-normal functioning in mood-disordered patients reflect their current mood state, whereas others reflect a cognitive vulnerability factor.

It is very important theoretically and practically to attempt to solve the causality issue. The reason for this is that one must understand how and why mood disorders develop in order to improve attempts to prevent such disorders happening and to improve methods of treatment. The optimal way of investigating the causality issue would be by means of a prospective study. This will be discussed in relation to clinical anxiety, but precisely the same logic is applicable to clinical depression. Very large samples of normal individuals would be given a range of cognitive tasks. If those normal subjects who subsequently developed clinical anxiety resembled anxious patients in their cognitive functioning more than did those normals who did not subsequently develop clinical anxiety, then this would strengthen the argument that a cognitive vulnerability factor is involved in the development of clinical anxiety. In contrast, if those normal subjects who subsequently developed clinical anxiety were comparable to those normal subjects who did not subsequently develop clinical anxiety in their cognitive performance, this would be evidence against a cognitive vulnerability factor.

Since most prospective studies are extremely expensive and time-consuming, alternative approaches have generally been used. An approach which has been used extensively (discussed at length by Eysenck, in press) involves recovered patients. This will be described with reference to anxiety, but is equally applicable to depression. In essence, the approach requires thorough assessment of cognitive performance in three groups of subjects: currently anxious patients; recovered anxious patients; and normal controls. The major assumption which is made is that the cognitive functioning of the recovered anxious patients is comparable to their cognitive functioning before they became clinically anxious. This assumption may not be entirely warranted. It recovered patients are comparable to current patients in their cognitive

performance, then it is possible to argue that this as occurred simply because they have not really recovered. Alternatively, it may be that suffering from clinical anxiety has a long-lasting effect on the cognitive system, so that it may never return to precisely the state it was in before the onset of the mood disorder. It is argued that those aspects of cognitive functioning reflecting a cognitive vulnerability factor should distinguish the normal controls from both of the other two groups, whereas those aspects reflecting anxious mood state should distinguish the current anxious patients from the recovered patients and the normal controls.

The evidence so far as clinical anxiety is concerned is somewhat equivocal. Let us consider first studies of the interpretation of ambiguous stimuli having both a threatening and a non-threatening meaning. The proportion of homophones interpreted in a threatening fashion by recovered anxious patients is intermediate between that of current anxious patients and normal controls, suggesting that interpretation of ambiguity reflects a combination of cognitive vulnerability and anxious mood state. However, it was found in another study that current anxious patients recognized relatively more of the threatening interpretations of ambiguous sentences on a memory test than did normals, and that the recovered anxious patients closely resembled the normals in their performance. Those findings suggest that the tendency of anxious patients to perceive and to remember the threatening meanings of ambiguous material is primarily a reflection of anxious mood state.

Somewhat clearer findings have emerged from the study of attentional processes related to anxiety. It will be remembered that anxious patients have been found to be more distractible than normal controls, and that this enhanced distractibility is more in evidence with threatening distractors than with neutral ones. Recovered anxious patients have been found to resemble controls when neutral distractors are presented, but they are significantly more distractible than controls when threatening distractors are presented. An important theoretical implication of these findings is that the tendency for attention to be captured by emotionally threatening stimuli may constitute part of a cognitive vulnerability factor. It is entirely reasonable that those indivisuals who cannot avoid processing mildly threatening environmental stimuli should tend to be vulnerable to generalized anxiety

disorder. More research is needed to see whether other aspects of cognitive functioning contribute to a cognitive vulnerability factor.

So far as depression is concerned, most of the available evidence indicates that non-normal cognitive functioning in depressed patients reflects their depressed mood state rather than a vulnerability factor. Consider, for example, a large prospective study that was carried out by Lewinsohn, Steinmetz, Larson, & Franklin (1981). They utilized questionnaires to assess a range of negative attitudes and cognitions. Their findings suggested that altered cognition in depressed individuals was due to being in a depressed state:

Prior to becoming depressed, ... future depressives did not subscribe to irrational beliefs, they did not have lower expectancies for positive outcomes or higher expectancies for negative outcomes, they did no attribute success experiences to external causes and failure experiences to internal causes, nor did they perceive themselves as having less control over the events of their lives.... People who are vulnerable to depression are not characterized by stable patterns of negative cognitions.

There have been rather similar questionnaire studies of cognitive functioning is recovered depressed patients. It has usually been found in these studies that recovered depressed patients have fewer and less intense negative cognitions than current depressed patients, and they often appear to resemble normal controls.

One of the most consistent differences between depressed patients and normal controls in their cognitive functioning is that only the former group demonstrates a negative recall bias, i.e. depressed patients have a relative tendency to recall negative information encoded in relation to themselves. There have been attempts to decide whether or not recovered depressed patients have the same negative recall bias as current depressed patients. The findings are somewhat variable, but appear to indicate that recovered depressed patients either have no negative recall bias or a smaller one than current depressed patients. It is probably reasonable to conclude that there is as yet little support for the view that a vulnerability factor is involved.

The evidence which has been discussed appears to point to the conclusion that there is no cognitive vulnerability factor in depression. Such a conclusion would be premature. Most of the relevant research has made use of questionnaire assessment of cognition, and thus

excludes any consideration of preconscious processes. Since controlled processes (or the products of such processes) have been investigated in most studies, rather little is as yet known about whether any automatic or preconscious processes might form part of a cognitive vulnerability factor for depression. In other words, no definite conclusions about the existence or otherwise of a cognitive vulnerability factor in depression can be reached until the cognitive system has been explored much more systematically.

In sum, most of the non-normal cognitive functioning in depressed and anxious patients which has been identified so far appears to reflect current mood state rather than a vulnerability factor. However, some indications of a cognitive vulnerability factor are present in the anxiety data, and studies of depression are too limited to justify any sweeping conclusions.

6

Pereception

Pereception has been increasingly thought of in cognitive terms, as is exemplified by the following definition from Roth (1986, p. 81): "The term perception refers to the means by which information acquired from the environment via the sense organs is transformed into experiences of objects, events, sounds, tastes, etc." In spite of the fact that perception can be studied in each of the sensory modalities, it is actually the case that most research and theory have focused on auditory and visual perception. It seems natural to assume that perception involves conscious awareness, but there is increasing evidence that extensive perceptual processing can occur in the absence of conscious awareness.

An important theoretical issue is the relative contribution to perception of bottom-up or data-driven processes and top-down or conceptually driven processes. Some theorists (e.g. Gibson, 1972) have argued that the stimulus is a very rich source of information, and typically provides the information required to produce veridical perception. In contrast to Gibson's (1972) emphasis on data-driven processes, constructive theorists have argued that conceptually driven processes play a major role in perception. Such theories provide an explanation of Visual Illusions, but appear to predict that perception should be more error prone than is actually the case.

While some perception theorists focus on general principles of perception, other theorists have concentrated on the perception of relatively specific categories of stimuli which are of major importance

in everyday life. For example, speech is obviously a very significant form of auditory stimulus and human faces are important visual stimuli.

Over the last 15 years or so, there has been increasing emphasis on considering perception from the perspectives of Cognitive Neuopsychology and Conitive Science. Among the conditions studied by cognitive neuropsychologists are Agnosia and Blindsight (Weiskrantz, 1986). The study of brain-damaged patients has provided a test bed for cognitive theories of perception. Cognitive scientists (e.g. Marr, 1982) have constructed numerous computer programs designed to reflect the ways in which human perceptual processing occurs. These programs have illustrated many of the lacunae in previous theories of perception put forward by cognitive psychologists.

Problem Solving

Definition

"Problem solving" is cognitive processing directed at transforming a given situation into a goal situation when no obvious method of solution is available to the problem solver. This definition involves four basic ideas. First, problem solving is cognitive - that is, it occurs inside the mind or cognitive system of the problem solver so that its existence can only be inferred indirectly from the behavior of the problem solver. Second, problem solving is a process - that is, it involves manipulating knowledge in the problem solver's mind or cognitive system (i.e. performing cognitive operations upon internal symbolic representations). Third, problem solving is directed—that is, it is intended to produce a solution to a problem. Fourth, problem solving is personal—that is, the difficulty of transforming a given state of a problem into a goal state depends on the existing knowledge of the problem solver.

A "problem" exists when a situation is in a given state, the problem solver wants the situation to be a goal state, and there is no obvious way of transforming the given state into the goal state. Hayes (1978, p. 177) summaried this definition as follows: "If you want to do something but do not know how, then you have a problem"; and Duncker (1945, p. 1) wrote that a problem arises when a problem solver "has a goal but does not know how this goal is to be reached." It follows that problem solving is what happens when a problem solver tries to reduce the difference between the given and goal states

of a problem. Polya (1968, p. ix) characterized problem solving as "finding a way out of a difficulty, a way around an obstacle, attaining an aim that was not immediately"; and Duncker (1945, p. 1) noted that "such thinking has the task of devising some action which may mediate between the existing and the desired situations."

For example, the left side of Fig. 1 shows the "nine dot problem" in which the problem solver is asked to draw four straight lines that pass through all nine dots, without raising the pencil from the paper (Adams, 1974). In this case, the given states is the nine dots, the goal state is having four straight connected line through the dots, and the allowable operators are to draw four straight lines without raising the pencil from the paper. A solution to the problem appears in the right side of the figure below, and requires realizing that the drawn lines may go outside the area bounded by the dots.

Problems may differ in how specifically the given state, goal state, and allowable operators are presented. In *well defined problems,* such as playing a game of chess, each of these components is clearly presented; in *ill-defined problems,* such as finding a solution to the energy crisis, one or more of these components is not clearly presented. *Routine problems* occur when a problem solver possesses a pre-existing strategy for solving the problem, whereas *non-routine problems* occur when a problem solver does not. For example, for normal adults, a long-division problem such as $7891 \div 13 =$ — , is a routine problem; whereas establishing a budget for household finances is a nonroutine problem. The distinction between routine and non-routine problems may be best viewed as a continuum in which the degree of required creativity is varied; creativity in problem solving involves generating and evaluating novel alternatives about how to represent the problem or how to design a solution plan. An extremely routine problem, in which the solution method is obvious to the problem solver, would not be classified as a problem at all because there would be no obstacles.

To place problem solving within the context of cognitive psychology, it should be noted that the former is a type of thinking (namely, directed thinking rather than undirected thinking – Gilhooly, 1982) and that thinking is a type of cognition (namely, a cognitive process involving manipulation of existing knowledge). Further, problem solving subsumes reasoning, including both inductive and

deductive reasoning.

Types of Problems

Greeno (1978) has described three types of problems. *Problems of inducing structure* occur when a problem solver is given a series of instances and must discover a pattern or rule, such as series completion and analogy problems. *Problems of transformation* occur when a problem solver is given a problem in an initial state and must determine a sequence of operations that produces the goal state of the problem, such as water jar problems (Luchins, 1942) and Tower of Hanoi problems. *Problems of arrangement* occur when the problem solver receives all the parts of the problem and must arrange them in a way that solves the problem, such as anagram problems and cryptarithmetic problems.

Gilhooly (1982) has distinguished between *adversary problems* such as the games of chess, Go, and poker, and *non-adversary problems* in which a single problem solver or group attempts to solve a problem. Further, reasoning problems include inductive reasoning problems, in which a problem solver must extrapolate a rule based on limited information, and deductive reasoning problems in which a problem solver must apply logically correct procedures to given information in order to derive a proven conclusion. Inductive and deductive reasoning problems can be categorized as problems of inducing structure and problems of transformation, respectively.

Processes in Problem Solving

The process of problem solving can be analyzed into four interrelated processes: representing, planning, executing, and monitoring. The representing process occurs when a problem solver translates the given problem into an internal mental representation of the givens, goals, and available operators in the problem (Mayer, 1983). Polya (1945) refers to this process as "understanding the problem" and Hayes (1978) calls it the "understanding process." The planning process involves establishing a hierarchy of subgoals for solving the problems, a process that Ploya calls "devising a plan". The executing process occurs when the problem solver implements the plan by carrying out a series of actions, a process that Polya calls "carrying out the plan." Finally, the monitoring process occurs when the problem

solver anlayzes his or her progress in solving the problem and determines how each executed action fits into the solution plan.

For example, consider the following problem (Riley, Greeno, & Heller, 1983):

Joe has 8 marbles.

He has 5 more marbles than Tom.

How many marbles does Tom have?

To represent the problem, one might mentally construct three sets of marbles, arrange the sets in part-whole relation to one another, and assign values to each set.

To devise a plan, one might mentally construct the following plan based on the equation. (Tom's marbles) = (Joe's marbles) – (Difference set):

1. To get the value of Tom's marbles, subtract the value of Joe's marbles and the value of the difference set.

2. To carry out goal 1, find the value of Joe's marbles and the value of the difference set.

The problem solving process depends on the knowledge that the problem solver brings to the problem solving task (Mayer, 1985). For example, representing the problem depends on linguistic knowledge such as knowledge of the English language and factual knowledge such as knowing Tom and Joe are boy's names. Planning depends on schematic knowledge such as knowing problem types similar to the marble problem and strategic knowledge such as knowing how to establish subgoals. Execution depends on algorithmic knowledge such as knowing how to add and subtract whole numbers. Monitoring depends on the problem solver's evaluation skills such as determining whether two facts are consistent with one another. Finally, the problem solver's beliefs about problem solving may affect the amount of effort put into solving the problem.

Theoretical Approaches to Problem Solving

Historically, there have been three major theoretical approaches to the scientific study of problem solving: associationism, Gestalt psychology, and computer simulation.

According to the associationist view, problem solving can be described as the exercise of a chain of pre-existing links in the problem solver's associative network. The major activity in problem solving is

to apply actions determined by the strengths of existing connections until by trail and error the problem is solved (Thorndike, 1898). For example, anagrams that make common words such as Ugars (Sugar) are easier to solve than anagrams that make less common words Obrac (Cobra), presumably because the associative links are stronger to common words (Mayzner & Tresselt, 1966).

According to the Gestalt view, problem solving involves reorganizing a problem and can be described as a search for structural insight—that is, a search for understanding of how and the elements in the problem fit together to satisfy the requirements of the goal. The major creative act in problem solving is to mentally represent the elements of the problem within the context of the goal.

The major insights in this problem are recognizing that the functional requirements of the goal are such that each stick must be part of two triangles and that to accomplish this, three-dimensional space rather than two dimensional space must be used.

The Gestalt psychologists produced several lasting contributions to our understanding of problem solving. First, Gestalt psychologists distinguished between *reproductive thinking* - applying pre-existing solution procedures based on past experience to a new problem - and *productive thinking* - generating a novel solution to a new problem (Wertheimer, 1959). Second, Gestalt psychologists introduced the idea theat stages of problem solving involve successive reformulations of the problem, including increasingly more specific representation of the given state and goal state of the problem (Duncker, 1945). Third, Gestalt psychologists provided evidence that past experience can create rigidity in problem solving, including the inability to use an object in a novel way to solve a problem (Luchins, 1942; Duncker, 1945).

According to the computer simulation view, problem solving involves applying a set of strategy rules to guide the search of a problem space. The problem space is a representation of the given state, goal state, and all possible intervening states produced by applying the operators to each state. A widely used search strategy, "means–ends analysis" (Newell & Simon, 1972), works like this: if there is an obvious action to solve the problem, it is carried out; if not, the problem solver establishes a subgoal of reducing the difference between the current state and goal state; if the problem solver finds an action that

can be directly applied, it is carried out; otherwise, the problem solver establishes a subgoal of removing any constraints on applying the action, and so on. In Newell and Simon's (1972) computer simulation of means-ends analysis, the search rules serve to establish each of three major types of subgoals: (1) "transform state A into state B," in which the problem solver compares the current state of the problem to the desired state in order to determine the differences between the two states; (2) "reduce difference D between state A and state B," in which the problem solve identifies an operator that can be applied to one of the states in order to reduce the difference between the states; and (3) "apply operator Q to state A," in which either a new state is created or the problem solver determines a new constraint that precludes applying the operator.

More recent advances in Cognitive Science – particularly theories in the field of Artificial Intellegence that build upon or are stimulated by the pioneering work of Newell and Simon (1972) - have added new perspectives on how to characterize problem solving in both humans and machines. In particular, the advances most relevant for problem solving are techniques for representing knolwdge and planning processes in novice and expert problem solvers.

Pattern Perception

This expression refers to the process of detecting and discriminating significant recurring combinations of spatially or temporally distributed energy in environment in order to make appropriate responses. This essay will examine the mechanism which may underly such discriminations, especially discrimination based on combinations of environmental features.

Animal behavior provides many examples of special pattern recognition systems, which (usually) enable appropriate inbuilt responses to occur to specific sensory inputs (McFarland, 1985). These inputs may be relatively simple, characterized by one or very few features. The male stickleback, for example responds with attack to any red astimulus, including a red bus passing the window. A single molecule of the phermone bombykol released by the female silkworm moth can trigger approach by a male.

These simple dependencies are likely to be controlled in the brain by a receptive cell, or more likely a group of cells, which will

fire only to a particular external stimulus and prime a set of possible responses, the precise response emitted being determined by other aspects of the environment. There is little ability to discriminate appropriate from inappropriate stimuli because in the natural environment this is largely unnecessary.

In more complex cases some combination of stimulus characteristics is required. The herring gull reacts to an egg outside the nest by reaching beyond the egg with its bill and retrieving it. Size, shape, background coloring, and speckles all play an independent part in eliciting the response, which also occurs (though less strongly) to a range of variations around the optimal egg pattern. The frog responds to moving spots by orientation and striking with the tongue. Neurons have been located in the frog's optic nerve which fire only to black moving dots. In order to respond only to a moving dot of the right size within a certain range, the detector cell in the frog's optic nerve must receive a complex combination of inputs over time from the retina.

In species higher in the evolutionary scale, inbuilt responses to particular stimuli are relatively rare. More important is the ability to learn to discriminate and identify complex patterns - that is, combinations such as shape and color or frequency and temporal duration, and learn new responses to them. The average human can immediately identify 50,000 words, if reasonably literate, or several thousand faces, and in specialist areas like chess an expert with extensive experience (around 10,000 hours in the case of International Grandmasters) can achieve similar feats for chess game positions. Apes have been taught to identify signs of American Sign Language and respond with sings of their own and to recognize plastic colored shapes and place them in combination to make requests for food and other requirements (McFarland, 1985). In the visual modality much pattern perception involves the identification of objects in a three-dimensional environment, often referred to as Form Perception. However, there are other important aspects of visual pattern perception in human beings, such as Face Recognition, perception of pictures, and reading, while Speech Perception is the most important human form of acoustic pattern perception.

These pattern recognition abilities are obviously much more complex and varied than those discussed at the beginning of this

essay. They cannot depend on specialist innate detectors for specific stimuli but rather on the ability to discover critical differences and ignore irrelevant ones. A number of theories have attempted to explain how these feats might be possible and evidence has been drawn from neurophysiology, psychological experiments, and attempts to program computers to recognize patterns. A will emerge, most of the theories have been designed to cope with two-dimensional visual patterns varying in shape.

Some theories assume that a master copy or "template" exists in the brain for a pattern which is to be recognized and the input is compared as a whole to this template. This might be done (using a visual example) by a single neuron collecting evidence for the presence and absence of light over a certain area of the retina (as the unit in the frog's optic nerve collects evidence for the presence of a spot). Other theories assume intervening levels of analysis, whereby particular "feature detectors" first operate on the input, identifying straights lines and curves of different orientations, the presence of which is then signaled to a higher order unit which only fires when a particular combination of features is present. There are also theories which assume no structured analyzers operating on the input but simply an interconnected network of identical units, of which different subsets respond to different patterns. Recently there has been a revived interest in sophisticated versions of such parallel distributed processing (PDP) models.

Before discussing these theories in more detail, it should be noted that they assume that either the input is presented as an already isolated unit such as a letter or face on a uniform background, or that a preliminary process of segmenting the input and isolating a distinct unit has occurred (this is no less true for speech where the gaps we "hear" between words are not present in the physical input but are creations of our speech perception processes). The preprocessing required is far from simple, particularly in the case of natural scenes or pictures of such scenes. There are no ready-made boundaries between the significant elements; contours are blurred, hidden, or non-existent and false contours are present due to shadows and other "noise." Marr (1982) has demonstrated how complex the processing is likely to be in visually segmenting natural scenes and implemented some possible algorithms in computer assimilations. Some of the critical

cues which induce segmentation in two-dimensional patterns have been identified, such as change in brightness or the orientation of texture (Pomerantz, 1985).The Gest Principls Of Perceptual Organization also embody principles of scene segmentation.

Template Theories

Template theories are not taken very seriously by most theorists because it is unclear how they could cope with variations or imperfections in the input. You can recognize a picture of Winston Churchill, even a cartoon consisting of a few lines, when you have not seen it before, because it is similar to pictures you have seen before, but templates are usually taken to require exact matches. Yet even frogs and herring gulls can cope with variation in the stimulus. Elaborations of template theories have been suggested in which inputs are preprocessed to "normalize" position, size, and orientation, but it is unclear how the necessary normalization can be determined without already knowing the template to which the input is to be matched. At best these operations could only cope with stimuli of limited variability. Cells in the monkey's visual cortex which respond to visual presentation of a monkey's hand appear to offer some support for the notion of templates (Gross, Rocha Miranda, & Bender, 1972). However, such cells can respond to a variety of hand patterns and the details of their make-up are unclear.

Feature Analysis

Such problems encourage development of a feature analysis theory, whereby a complex pattern is decomposed into simpler elements which are detected by similar processes to those envisaged for templates (features are in fact miniature templates). Another modification is the assumption that the stored representation of each template or feature consists, not of an exact copy, but of a "prototype," or typical instance derived by averaging over instances which have been encountered. A prototype can be defined more precisely as the sum of the averaged values of all the component features. This approach can in principle deal with examples like recognition of Churchill's face, but it should be noted that no precise account is offered as to how variability around the prototype is processed in the nervous system.

Hubel and Wiesel (1959) reported evidence for units in the cat's visual cortex responsive only to edges or bars of a particular orientation. Subsequently they also found units repsonsive to corners. Such unit were originally thought to be basic feature detectors (though no units responsive to other important features such as curves have been found), but theorists such as Marr (1982) have argued more plausibly that these units are involved in the basic operation of locating contours at an early stage of input processing to create the "primal sketch". Psychological evidence for use of features in identifying simple patterns comes from experiments such as those in which subjects search for a target letter consisting of curved lines in a background of straight letters. Search times increase markedly if the background letters are changed to a set consisting of curved lines but not if the new set also consists of straight lines (Rabbitt, 1967).

The best examples of reasonably plausible models of feature theories involve language processing (e.g. Speech Perception). Human speech in produced by forcing air through the vocal chords. Movement of different parts of the vocal tract reinforces distinct frequency bands in the resulting sound, which are known as "formants." Each vowel is typified by a particular combination of two main formants (though three or four are present). The formants are clearly apparent when the frequencies present at each instant are plotted on a graph over a period of time. Consonants appear as brief bursts of wide-frequency noise but have few invariant characteristics. Instead, their characteristics vary according to the succeeding novel. However further cues are provided by the "transitions" - that is, the changes in the formants which occur between the consonant noise burst and the following vowel which show typical patterns for different consonants and can produce perception of a particular consonant even when the burst of wide frequency noise due to the consonant is removed. Removal of the transition, on the other hand, often eliminates consonant recognition. Thus, the important units seem to be consonants plus vowels (that is syllables) rather than single sounds or phonemes.

The most elaborate feature model for word recognition is that of McClelland and Rumel hart (1981). Fig. 1 shows the essentials. Input excites feature detectors which are connected to letter detectors with an excitatory connection where they form part of the letter and an inhibitory connection otherwise. Similar connections exist between

letter detectors and word detectors and feedback also occurs in the reverse direction. The word units will response correctly to partly obscured words because the correct solution receives some excitation from all the letters. Feeback from words to letters explains how letters can be more easily indentified when presented in words when presented on their own.

There are, however, many difficulties with feature analysis theories. While it may be possible to specify adequate features for restricted pattern sets, it is difficult or impossible to do this for many other patterns such as everyday three-dimensional objects and pictures of them. Second, the relevant features often depend on the current context and purpose. Deciding in the context of numbers whether a smudged item is a 3 or an 8 requires different information from deciding in the context of letters whether it is an s or a g. Third, the Gestalt psychologists argued that it was impossible to specify patterns adequately in terms of components parts, since complex patterns often produce an experience which is not predictable from the nature of their parts. Visual Illusions are the most striking examples. Later research has found other examples. Discrimination of the pattern (from) is much easier than discrimination (from) even though the same element has been added to both, giving no extra discriminatory power on its own. The addition produces enclosed shapes which in effect provide a new feature (often called an "emergent feature"), so it is impossible to account for all pattern preoccupation with a limited set of basic features (Pomerantz, 1985).

The most crucial objection to treating patterns as lists of features, however, is that *relations* between features are vital. L and T both have a vertical and a horizontal bar, K and Y both have a vertical bar and one left-inclined and one right-inclined bar; it is the spatial relations between these elements that differ. Hence, more elaborate theories are needed in which such relations are specified. Such theories are known as "structural description theories." These may appear more promising, but in practice they are still vulnerable to all the difficulties of specifying critical elements and dealing with variability and imprecision which have been described above.

Many theorists have argued that some of these problems of pattern perception can be reduced if it is assumed that the type of analysis depends on the context. If printed language is anticipated,

certain feature analyzers will be engaged, while others, when required, will cope with handwritten language or faces. Furthermore, recently processed information can guide processing of later input. This is known as "top-down" or conceptually driven processing as opposed to "bottom-up" processing in which invariant and possibly inbuilt processes are applied to all inputs. Marr (1982) has argued that bottom-up processing is adequate in natural environments to produce a viewer-centered representation of objects in depth. Such a view is not, however, incompatible with a belief in the occurrence of top-down processing in other situations, especially where patterns created by humans are involved. For example, spoken words in sentences are identified more easily than those in isolation, though predicting ahead seems to play only a very restricted role in reading (Rayner & McConkie, 1977).

Parallel Distributed Processing Theories

None of the theories discussed so far offers an explanation of how features and feature combinations might be learned. It is clearly unlikely that preformed pattern detectors could exist for all the variations which can be processed and the great strength of pattern recognition in higher species is the ability to cope with new discriminations. The appeal of the newly developed connectionist theorise of memory of PDP models is that they offer a viable approach to such learning as well as to the classic problems of variability and imperfection in the input. These theories do not envisage single specialist units which gather evidene for specific patterns. Instead, large numbers of identical units are connected up in specific ways. Each unit has a threshold and excitatory or inhibitory links, which vary in strength, with other units. A unit fires if it receives sufficient net excitation from others, and then it passes a signal to other units. Thus, Input A will exile one combination of receptor units which pass signals on through the system, and Input B excites another combination (which may include some of the same units). With some additional assumptions, which there is no space to describe here, it has been shown that such systems can be taught to discriminate different inputs by adjusting the strengths of the connections according to whether a desired output is achieved or not. When an error occurs, excitatory connections are weakened and inhibitory connections strengthened. Removal of some of the

units or part of the input does not seriously impair performance. Hence, these systems offer impressive parallels for pattern discrimination in living organisms. This is not to say that all pattern discrimination learning is of this gradual nonanlytic nature. A child may learn the critical difference between E and F by explicitly attending to the relevant feature.

As stated earlier, feature theories have tended to concentrate on shape discrimination and PDP theories provide the best approach to shape discrimination by bypassing the problems of defining suitable features (though they are not limited to a specific form of input). However, many patterns have features other than shape, such as color, size, texture, and movement. Neurophysiological evidence suggests that separate specialist subsystems in the brain are responsible for processing different feature classes (Zeki, 1978). The processing of combinations of these feature classes has been investigated in psychological experiments by Treisman (1988) and her associates. Treisman has examined particularly the processing of color-shape combinations and has argued that color and shape are computed by independent processes and the results are combined in an "object file" which collects the information for a specific spatial position. A parallel scan can be carried out across the whole input field for a single feature value such as green, since a green item can be found just as quickly whether it is hidden among 10 or 20 blue items. The combining process, however, is sequential and requires direction of attention to a specific location (spatial or temporal), since time increases when searching for a green T among green Xs and blue Ts as the number of irrelevant items increases. Treisman argues that empirical criteria, such as whether parallel processing of the above kind is possible, can be established for identifying separate features. Clearly this work implies that the passive implicit learning embodied in PDP models could not provide a complete picture of the processes involved in pattern perception.

Summary

Several different mechanisms responsible for processing input patterns have been suggested, including purpose-built innate systems for detecting specific stimuli, flexible PDP systems which learn discriminations in a nonanalytic way, and synthesis of different feature

classes present at a specific location or time to form a complete object. It is therefore unlikely that a single all-embracing theory of pattern perception is viable.

Perception, Computation Theory of

The logical foundations for claiming that a computational appraoch is appropriate in psychology rest on Turing's proof that "a language capable of defining 'effective procedures' suffices, in principle, to solve any computable probelm .... [so that] .... If a psychological science is possible at all, it must be capable of being expressed in computational terms" (Boden, 1988, p. 259). At a practical level, computer simulations of cogntive processes have provided a brutally rigorous way of weeding out flawed ideas, or of refining valuable but rather vague notions into more precise formulations, and of directing attention to hitherto neglected theoretical problems. The end result has been working models of various cognitive processes. Taking vision as an example, visaul systems capable of guiding sophisticated hand-like robot grippers in picking up and manipulating an object have now been demonstrated. To be sure, such systems are presently to limited scope compared with human vision in the range of objects and scenes they can deal with. But the fundamental change in outlook prompted by having at least some working models in the areas classically the concern of perceptual psychology is felt by many to have had a revolutionary impact on the subject. Computational ways of addressing problems have brought far greater clarity, precision, and detail into theories of perception than existed hitherto and the same is true of cognitive theories in general. Indeed, it seems fair to say that the point has now been reached where it is almost inconceivable for a cognitive psychologist intent on worthwhile theory development to have much chance of success without deep familiarity with the associated computational literature (often mathematical in nature in the case of perception).

The general character of the computational approach to perception will be outline here by describing the conceptual framework advanced by David Marr. His work is chosen for illustrative purposes because he has had a larger impact on cognitive psychology than any other single figure engaged in developing computational models of perception. It should be realized, however, that he is far from being a

lone voice: the computational literature on perception is an explosively increasing one with no sings of levelling-off at present, and Marr's own work is but a fragment of it. His work has been influential, partly because it was closely linked to studies of biological perceptual systems and partly because it offers a general framework for cognitive science. To ensure that its implications are brought out in sufficient detail in the limited space available, just one specific topic will be discussed – the stereo correspondence problem. That is fitting because it is a problem to which Marr gave considerable attention and which he used himself as a paradigmatic illustration of his approach.

The point has already been made that the power of the computational approach derives in large part from the discipline of creating a computer implementation of the theory of interest. This soon leads to the shortcomings of existing theory being revealed and, as adjustments are made to improve it, to an ever more thorough analysis of the nature of the task to be solved. At this level of analysis, called the "computational theory" level by Marr (1982; Marr & Poggio, 1976), the questions raised are twofold: first, what exactly is the goal of the computation?; and second, what method(s) can be devised for achieving the required goal? In applying the computational approach to perception, the answer to the first question is frequently an abstract specification of a desired mapping from one kind of information to another, and usually an ambiguity that must be resolved in order to make the mapping is identified. The anser to the second question is couched in terms of (a) *constraints* flowing from the nature of the viewed world and its projections into images, or sometimes arising from the logic of the task and (b) *demonstrations of methods* (preferably with a precise mathematical treatment) showing that these constraints are adequate in principle for achieving the desired mapping. Given this content, it is questionable whether Marr's choice of the term "computationai" for this level of theoretical analysis is a helpful one but it has become established in the literature.

Exploiting the computational theory in a practical system requires attention to a second level of analysis, called by Mark the "algorithm level." Here the issues are to do with achieving a workable implementation of the abstractly defined method(s) specified in the computational theory. This forces choices about the particular input and output representations to be used and the detailed sequence of

processing steps needed to achieve the required input-to-output transformation. Marr (1982, p. 20) defines a representation as a formal system for making explicit certain entities or types of information, together with a specification for how the system does this. He calls the result of using a representation a "description" of an entity in that representation. By an "explicit" description is meant one which makes the information required available for immediate use – that is, without any need for further work by subsequent processes using the description.

Marr's third level of analysis concerns the hardware: how can the algorithm be realized physically? Often the same algorithm can be implemented in quite different technologies. The choice will usually have much to do with practicalities such as availability and cost. For example, Marr (1982, p. 24) observes that wires are rather cheap in biological architecture, because they can grow individually and in three dimensions. In conventional [computer] technology, wire laying is more or less restricted to two dimensions, which severely restricts the scope for using parallel techniques and algorithms; the same operations are often better carried out serially.

This approach will now be illustrated by computational work on the stereo correspondence problem, for which the goal is to take two stereo images and map them into a single representation making explicit the binocular disparities of the various features in each image. The major difficulty, brought to the fore by Julesz's work using random-dot stereograms, is generally regarded as resolving ambiguities of matching: which point in the left eye's image should be matched to which point in the right eye's image? The kind of theory provided by the computational approach is the identification of constraints flowing from the nature of the viewed world and its projections into stereo images. Numerous constrainsts have now been proposed. Marr and Poggio (1976) suggested that it is generally reasonable to assume that the visual world is made up of matter that is separated into objects whose surfaces are smooth compared with their distance from the viewer. In other words, the visual world is *not* usually made up only of a cloud of identical dust particles: the matching problem would be intractable if it were. They called this the "cohesivity constraint" and they used it to justify their *continuity* binocular matching rule: *prefer possible matches that could have arisen from smooth surfaces.* Marr

and Poggio also identified the "uniqueness constraint": a scene entity cannot be in two places at the same time, from which they derived the binocular matching rule: *each item from each image may be assigned at most one disparity value.* Marr (1982 , p. 114) added to these two constraints a third called the "compatibility constraint" which leads to the binocular matching rule: only allow matches between descriptive elements derived from the two images which could have arisen from the same physical surface marking.

From these three constraints, Marr postulated the fundamental assumption of stereopsis: *If a correspondence is established between physically meaningful primitives extracted from the left and right images of a scene that contains a sufficient amount of detail, and if the correspondence satisfies the three matching constraints, then that correspondence is physically correct.* He stated that ".... to isolate this fundamental assumption and to establish that it is valid is precisely what I mean by the computational theory of a process" (Marr, 1982, p. 115). What this amounts to is developing the assumption in more precise terms and then proving that the constraints do indeed force unique correspondences (Marr noted that phrases like "scene contains a sufficient amount of detail" and "physically meaningful primitives" are too imprecisely stated for mathematical proofs). He went on to show that, given the constraints, matches falling in same-disparity planes will be denser than on any others and hence preferred by the matching rules.

Marr and Poggio (1976) implemented their two matching rules in an algorithm that utilized a cooperative neural network. The nodes of the network represent potential matches and these nodes exchange excitation if they have the same disparity. This mutual facilitation implements the continuity constraint by giving preference to matches that could have arisen from smooth surfaces. The uniqueness constraint is implemented by inhibition exchanged between nodes lying along the same line of sight, the effect of which is to suppress more than one match for each primitive. This network could be realized in a variety of hardware architectures, varying from simulations on standard serial von Neumann machines through to specialized parallel processing devices – and, plausibly but speculatively, in biological neurones. For a tutorial introduction to the Marr and Poggio (1976) algorithm, as well as an introduction to the computational approach in general.

As an illustration of what constitutes theory development using the computational approach, it is worth noting that it soon became evident that the form of facilitation used by the Marr and Poggio (1976) stereo algorithm is needlessly restrictive: insistence that only units with the *same* disparity exchange support limits the algorithm strictly to dealing with surfaces which lie locally in planes perpendicular to their viewing direction (e.g. frontoparallel planes near the fixation point.) A broader interpretation of smoothness is implemented in the so-called PMF stereo algorithm which allows pairs of matches with different disparities to exchange support as long as the gradient of disparity between the pair is not too large. Pollard, Porrill, Mayhew, and Frisby (1986) prove that the use of this "disparity gradient limit" constraint imposes a form of scene-to-view and view-to-view continuity during stereo matching which can be informally described as allowing "jagged but not too jagged" surfaces to be dealt with. This proof admirably satisfies Marr's requirements for a computational theory, it extends and tightens up the mathematic expressing the nature of the smoothness constraint, and as a result it produces a more effective and more widely applicable stereo algorithm (Frisby, 1990). Such a development is an example of an advance in computation theorizing in perception.

Yet another use of the surface smoothness constraint is Mayhew and Frisby's (1980) "figural continuity" matching rule: because cohesive objects generate surface edges and surface markings that are spatially continuous, *prefer matches that preserve figural continuity.* This is mentioned here to being out the point that diverse methods can often be drived from a broadly stated constraint. Indeed, the smoothness constraint is used more widely than just in stereo, playing a role as a so-called "regularizer" for what are technically known as "ill-posed problems" (i.e. mathematically over- or underdetermined problems) in algorithms capable of dealing with extracting surface shape from motion, shading, and other depth cues (Peggio, Torre, & Koch 1985; this paper provides a good entry point to the computational literature on low-level vision).

Marr and Poggio (1979) implemented the same two constraints, of smoothness and uniqueness, in a quite different *non*-cooperative stereo algorithm that used multiple spatial frequency tuned channels

and a coarse-to-fine matching strategy. The key idea here was to exploit the fact that only a few edge points, and hence only modest ambiguity problems, arise in very coarse channels. Matches obtained in these can then be used to guide matching in more finely tuned channels. Ambiguities in all channels are reduced almost to zero by suitable coupling of spatial frequency tuning to the disparity range allowed for matching. This requires that the high spatial frequency channels have very narrow disparity ranges and that they therefore need to be "put in the right place to look" if they are to find the correct matches. The algorithm does this by generating appropriate vergence eye movements driven by the coarser channels. It is unlikely that human vision implements this algorithm in that disparities beyond the range allowed by the theory for high spatial frequencies can be fused without need of eye movements (Mayhew & Frisby 1980). This experiment is an example of a psychopysical study being driven by computational theorizing.

The viability of their second algorithm depends, Marr and Poggio (1979) argued, on the same surface smoothness constraint used to underpin their first one. The details of this argument can be challenged (on the grounds that surface smoothness in three-space does not necessarily imply useful continuity across different spatial frequencies: Mayhew and Frisby, 1980; Prazdny, 1987). Nevertheless, Marr's claim that both algorithms depend on the same underlying constraints is a nice example of his belief that it is both useful and possible to distinguish different levels of analysis:

These three levels of analysis [computational theory, algorithm, hardware] are coupled but only loosely: The choice of algorithm is influenced for example, by what it has to do and the hardware in which it must run. But there is a wide choice available at each level, and the explication of each level involves issues that are rather independent of the other two.

(Marr, 1982, p. 25)

The need to distinguish different levels of discourse is accepted without question in computer science in the design and analysis of human-made complex information-processing systems. Many technical concerns at the level of digital hardware are quite irrelevant to the business of devising good algorithms to run on that hardware. This is reflected in the important notion of a "virtual machine" in

computer science, defined as a set of information-processing operations:

A physical mechanism (a calculator, a computer, or brain, perhaps) may instantiate a particular virtual machine which can be used as a basis for implementing other virtual machines (using programs which define operating systems, compilers, interpreters, and so on).

(Sloman, 1980, p. 403)

Indeed, although Marr's particular prescription for the kind of analysis required at the computational theory level is distinctive, insistence on the general importance of this separate explanatory level does not originate with him. The whole enterprise of artificial intelligence rests on Newell and Simon's assumption that "A physical symbol system has the necessary and sufficient means for general intelligent action", which itself can be traced to Turing's analysis on the notion of "computability". In other words, many different "physical symbol systems" can support intelligence, not just biological ones. This is also the central, though not unchallenged, claim of the functionalist approach to the philosophy of mind. Also, Gregory's (1973) emphasis by way of developing Helmholtz's dictum that "perception is unconscious inference", on the need to know the *strategies* embodied in a mechanism before one can be said to understand it, is yet another example of the widespread recognition that there is a need for at least one more level of discourse over and above that of hardware.

Marr's achievement was not therefore a new insight that different levels of analysis are required for understanding complex information-processing systems, but a specification for what those levels are and what each should be concerned with. It is of course too soon to tell whether his formulation will endure but Boden (1988), in reviewing the evolving body of computational and psychopysical studies deriving from his work, judged that

We have here, then, the germ of what Lakatos called a "scientific research programme": a progressive body of hypothesis and experimentation, generated by a central theory that is amended as research proceeds.

(Boden, 1988, p. 75)

Nevertheless, neuroscientists do not always find Marr's three-levels distinction a comfortable one. This may be because neural

structures genuinely require of quite different form of analysis, one perhaps in which (at least) the algorithm and hardware levels are inextricably bound up together. Alternatively, it may simply be that a preoccupation with studying the neural mechanisms and phenomena of biological brains tends naturally (if regrettably) toward a neglect of the computational theory level of task analysis, whereas this level is manifestly important when faced with the task of trying to build an "artificial brain" using computers. For computational theorists, the simple demonstration of a "grandmother cell," the classic *reduction and absurdum* in debates about stimulus encoding by highly specifically tuned neurones, is irrelevant to their principal concern, which is "an answer to how you, or a cell, or anything at all, does it" (Mayhew, 1983, p. 214), not just knowledge of how the concept of grandmother is finally encoded in the brain. Yet another additional factor may be that the training and skills of biologists often render the computational literature rather inaccessible to them because of its mathematical leanings. However, the bridge between the computational and biological literatures is now being traversed more and more often. Indeed, the term "computational neuroscience" has been coined for the difficult but potentially richly rewarding enterprise of playing psychophysical and physiological knowledge about neural systems within a computational framework that provides a detailed analysis of the nature of the task(s) being solved by those systems.

In an entertaining and provocative epilogue with an imaginary interlocutor, Marr (1982) expresses in trenchant terms his view that analysis at the computational theory level has been neglected in much cognitive psychology and also in much artificial intelligence. For example:

> As a computing mechanism, a production system exhibits several interesting ideas – the absence of explicit subroutine calls, a blackboard-like communication channel, and some notion of a short-term memory. However, just because production systems display these side effects does not mean that they have anything to do with what is really going on [in human congnition]. For example, I would guess that the fact that short-term memory can act as a storage register is probably the least important of its functions. I expect that there are several "intellectual reflexes" that operate on items held there about which nothing is yet known and which will eventually be held to be the crucial things about short-term memory. Studying our performance in close relation to production systems seems to me a waste of time, because it amounts to studying a mechanism, not a problem. Once again, the mechanisms that such research is trying to penetrate will be unravelled by studying the problems that need solving, just as vision research is progressing because it is the problem of vision that is being attacked, not neural visual mechanisms.
>
> (Marr, 1982, p. 348)

As might be expected, such criticisms have not gone unchallenged. For example, Sloman (1980) believes Marr's three-levels assumption is "confused" because it rests on the mistaken belief that the topmost computational level of theorizing can in general be separated from the level of algorithms and the study of representations. The point at issue here may prove to be a deep one or it may be no more than a matter of terminology. Marr (1982, p. 23) uses the term representation in the context of (level two) discussions of practical design questions about how to devise an effective set of procedures for implementing a set of constraints. An example of such a question is: should a node in the Marr–Poggio neural net stereo algorithm "represent" a dot, or an edge point, or whatever? Others might wish to use the term "representation" in different (more abstract?) senses in (level one) debates about what constraints are available and what mathematical proofs can be demonstrated regarding the implications of those constraints.

This raises a further complaint from Sloman (1980): pursuing Marr's set of levels is likely to divert attention from difficult and messy problems in psychology to relatively simple mathematical problems:

> Many of the most important issues in AI have been concerned with the study of trade-offs between .... space and time, efficiency and flexibility, completeness and speed, clarify and robustness. It is possible that such tade offs are the key to much of the complexity of human and animal psychology, and ultimately neurophysiology. If so, it may be a serious impediment to scientific progress to advocate an oversimple methodological stance .... The rigidity of function of a typical calculator makes it unnecessary for our understanding of it to involve consideration of many layers of implementation or the kinds of trade-offs and mixtures of levels found in human psychology. By

contrast, when we study *human* arithmatical expertise (acquired after many years of individual learning), most of the mathematical theory of numbers is an irrelevant digression. Instead we have to consider issues of storing many "partial results", indexing them, linking then methods of recognising situations where they are applicable, associating them with monitoring processes for detecting slips and mistakes, and so on.

(Sloman, 1980, p. 403)

Marr's answer to this would probably be Yes and Not (One can say only probably, because Man died from leukemia at the age of 35). Sloman's list of trade-offs is important and they are exactly the kinds of things that need to be understood at the algorithm and hardware levels. But it is premature to consider their significance via tasks such as mental arithmetic (or chess) which are:

.... problems for which human skills are of doubtful quality and in which good performance seems to rest on a huge base of knowledge and expertise. I would argue that these are exceptionally good grounds for *not* yet studying how we carry out such tasks. I have no doubt that when we do mental arithmetic we are doing something well, but it is not arithmetic, and we seem far from understanding even one component of what that something is. I therefore feel we should concentrate on the simpler problems first, for her we have some hope of genuine advancement.

(Marr, 1982, p. 348)

The controversy continues . Marr's book is an excellent and largely nontechnical introduction to his approach over a wide range of vision problems: he and his colleagues did much more than tackle the stereo correspondence problem used as illustration here. Boden (1988) provides an impressively detailed and illuminating discussion of the unsefulness both of Marr's strictures and of the computational approach in general to the whole field of cognition. Her book brings out with exception force how the language of computation has provided a new way of speaking about cognition, and a conceptual framework which offers the promise of unified treatment of diverse cognitive phenomena. For a more introductory account of cognitive science and the way it has been stimulated by the advent of the digital computer, consult Johnson-Laird's (1988) excellent review.

7

Perceptual Constancies

Perceptual constancies are to be found in various sense modalities; they refer to the fact that our perception of objects, sounds, and so on remains relatively constant in spite of substantial variations in sensory input. If our perception of the properties of objects were completely based upon the information that our sensory receptors receive at any moment, then our world would seem as chaotic as the Wonderland that Alice found at the bottom of the rabbit hole. Consider our visual impressions, for instance. A piece of white paper ought to appear black when viewed in the moonlight, since the amount of light in the image on the retina of your eye is no brighter, under these conditions, than than from a piece of coal viewed in normal room light. Since the retinal image is larger when objects are closer, as we move the piece of paper nearer to us it should appear to increase in size. This same piece of paper would appear to change shape continually as we tilt it, since the retinal image would vary from rectangular to trapezoidal depending upon the inclination. The color of the white paper should also appear to fluctuate, appearing yellow under incandescnet lighting, blue in fluorescent light, and white in sunlight. Fortunately, our visual perception of objects is much more *constant* that it would be if the only information available to consciousness was the retinal image itself. Thus, the piece of paper appears to be a white rectangular object with a fixed size, even though you might sense the fact that the light falling on it has changed in intensity or color, or that its distance from you or angle of tilt have changed.

These examples illustrate a basic principle of perception, namely that *the properties of objects tend to remain constant in consciousness even though our perception of the viewing conditions may change.*

There are many different types of perceptual constancies. One set includes specific object properties, such as an object's size, shape, color, or the whiteness of its surface; another set pertains to the location of the object relative to the observer's body. Each is named on the basis of the aspect of the object that remains unchanged. Thus, "size constancy" refers to the fact that the size of the object remains unchanged in perception despite the fact that the retinal image size changes, "color constancy" refers to the fact that object color remains unchanged despite changes in the color of the illumination, and so forth.

To understand how constancies come about it is important to recognize that each constancy may be analyzed into two major stages of operation. In the first stage we have "registration," a function in which changes in the sensory input are encoded for further processing. The individual is not necessarily consciously aware of this registration process. The second stage of processing involves "apprehension," or the actual subjective experience. Apprehension refers to the conscious component of your perceptual experience that is available for you to describe. Under normal conditions, registration is oriented toward a "focal stimulus," which simply refers to the object that you are paying attention to. While you are registering the focal stimulus you will also be registering many of the stimuli that are nearby, or stimuli that reach your sensory receptors around the same time. These additional stimuli form the "stimulus context". During apprehension you become aware of two different classes of properties. The first involves the "object properties" of the focal stimulus. These include the size, shape, and color of the object, which are the properties that tend to remain constant in consciousness. The second set involves "situation properties," which include the more changeable aspects of the environment, such as your distance from the object or the amount or color of the available light. Situation properties are derived from cues found in the context.

All of this may sound a bit complicated in theory, but in practice it is really quite straight forward. To see how the various categories of stimuli interact, refer to Table 1. It describes the various sets of

stimuli for several different constancies and is based upon Coren and Ward's (1989) treatment of this problem. To understand how to use this table let us consider "size constancy" as an example. Notice that the focal stimulus (what you pay attention to) is the size of the retinal image of the object. The context stimuli are all of the surrounding cues that indicate the distance of the object. Although you process these stimuli during the registration phase, you are generally not consciously aware of them. During the apprehension stage, you become aware of two aspects of the object, one that changes and one that does not. In size constancy it is the object's size that remains invariant while its distance changes.

The context is very important for each of the constancies. Thus, for size constancy, if we remove the cues that indicate the distance of the object, constancy breaks down. Now changes in distance result in the perception of the size of the object changing as the retinal image size increases or decreases (e.g. Harvey & Leibowitz, 1967; Chevrier & Delorme, 1983). Furthermore, if the viewing conditions are strange, and the context is unfamiliar or unusual, constancy fails (Day, Stuart, & Dickinson, 1980). To see your size constancy break down you need only climb to the top of a tall building and you will see that people below appear to be tiny dolls and automobiles appear to be toy cars. This is because the far distance, and unusual viewing angle (from the top) make the context less interpretable, and size constancy can not be maintained.

The perceptual consultancies, then, are a complex set of "corrections" based upon the availability of context stimuli. They take into account the ongoing conditions and allow us to extract a stable set of object properties from the continuous flow of changing sensory inputs at our receptors. Were it not for the perceptual constancy corrections, objects would have no permanent properties in consciousness at all, but would change in size, shape, lightness, and color, with every move that we make or with every change in environmental conditions.

Perceptual Development

This term is concerned with the systematic changes in perceptual abilities and processes manifested by children as they develop. "Perception" and "cognition" are often distinguished in that the former

refers to the basic detection and processing of sensory information, while cognition refers to the ability to extract sense, order, and meaning from this information – i.e. what sense does the infant make of what is seen of heard? While this distinction, in practice, is often difficult to make, the theme "from sensation to cognition" is a focus of this entry, which concentrates particularly on visual perceptoin. This bias is justified on the grounds that vision is the modality whose development has been most explored, and it can be argues that it is the most important of the senses. Coordination between the senses, and the senses of taste, olfaction, and audition, are discussed later.

Basic Visual Capacities

As we might expect, the visual information detected by the newborn is very impoverished when compared with that detected by the adult. Sensitivity to contrast differences is poor. A black and white pattern gives a contrast approaching 100 percent, and under good viewing conditions adults can discriminate between shades of gray giving contrast values of less than 1 percent; a contrast value of 30-40 percent is close to the newborn's threshold of detectability. Visual acuity, the ability to detect fine detail, is also poor. The most commonly used procedure to measure visual acuity is the visual preference method, where black and white stripes (gratings) are shown to the infant paired with an equal luminance gray patch: the width of the stripes is progressively reduced and when the infat no longer looks at the stripes in preference to the gray patch it is assumed that the acuity threshold has been reached. Acuity estimates for newborns measured in this way are about 1 cycle per degree, equivalent to the ability to detect stripes 2 mm wide shown at a distance of 30 cm from the eyes. This level of acuity is, curiously, not too different from that for the adult cat (about 3 cycles per degree), and acuity improves quickly in the postnatal months so that adult levels (30-40 cycles per degree) are reached sometime after six months of age.

Other basic visual capacities are present at, or develop shortly after, birth. Accommodation, the ability to focus on objects at different distances, improves along with changes in acuity, so that from two months, or earlier, all normal infants alter their accommodation in the appropriate direction as the distance of a visual target changes.

Newborn infants have some degree of color vision. Adams,

Maurer, and Davis (1986) reported that newborns differentiated gray from green, from yellow, and from red: for each of these colors they preferred color and gray checkerboards to gray squares matched for overall luminance. However, the newborns showed no evidence of discriminating between gray and blue. While we do not yet have a detailed account of its development, it seems likely that adult-like color vision is present by about three months of age.

By being physically separated in space the two eyes provide slightly different images of the perceived world. Detection of these differences, or disparities, provides the basis for at important binocular cue to depth, known as stercopsis, and several studies suggest that stereoscopic depth perception emerges around three months of age. However, the ability to detect disparity differences does not necessarily mean that they specify different depths to the infant, although, of course, they may do so. Other studies, described later, have given more definite evidence of depth perception in infants.

Early Visual Perception

From birth, infants make eye movements which regularly shift fixation from one part of the visual world to another, and the patterning and direction of these eye movements suggest that at all ages infants must be considered to be active seekers of stimulation. Newborn infants, when shown pairs of stimuli, will display consistent preferences between them, in the sense of looking more at one member of the pair: for example, they will prefer moving to stationary, large to small, three-dimensional to two-dimensional, high contrast to low contrast stimuli (Slater, 1989).One procedure that has proven to be particularly useful in uncovering infants' discriminatory abilities is "habituation": if one stimulus is presented repeatedly over a period of time infants will spend progressively less time looking at it, and they will often subsequently "dishabituate" (show recovery of attention) when a different, novel stimulus is shown. This procedure works with infants from birth. For example, after newborns had been habituated to one of four simple geometric shops (a square, cross circle, or trinagle), they gave novelty preferences when the familiar shape was paired with one they had not seen before. While this is evidence of early shape discrimination, the basis of the discrimination is unclear since any two shapes will differ in features such as orientation of lines and

angles, overall size, density of contour, enclosed versus open, and so on. These sorts of "low order" variables are discriminated by newborns and it is likely that they are discriminating between configurations on the basis of these features rather than on the basis of "true" form perception. An experiment by Cohen and Younger (1984) illustrates this. Six- and fourteen-week-old infants were habituated to a simple stimulus consisting of two connected lines which made either an acute (45°) or obtuse (135°) angle. On subsequent test trials the six-week-olds dishabituated to a change in the orientation of the lines (where the angle remained unchanged), but *not* to a change in angle alone, while the fourteen-week-olds did the opposite in that they recovered attention to a change in angle, but not to a change in orientation. This suggests that shape perception in infants six weeks and younger may be dominated by attention to lower order variables such as orientation. However, from two months of age infants are able to perceive angular relationships, which may be the basic elements or building blocks of perception. Beyond about two months of age infants begin to demonstrate the ability to perceive wholes, rather than parts of visual stimuli and to classify and categorize stimuli on the basis of perceiving similarities and differences between different examples – for example, recognizing that several different rectangles are members of a class separate from a square or a circle.

We have seen earlier that sensitivity to binocular disparity, stereopsis, as a cue to depth, appears around three months of age. Motion carried, or "kinetic" depth cues are responded to even earlier. Newborn infants will selectively fixate a three-dimensional stimulus in preference to a photograph of the same stimulus, even when they are restricted to monocular viewing and the major depth cue is motion parallax.Appreciation of pictorial depth cues – those cues to depth that are found in static scenes such as might be found in photographs – has been found from about five months. One such cue is relative size, the larger of two otherwise identical figures usually being perceived as the closer. Yonas, Cleaves and Pettersen (1978) used the "Ames window", a trapezoidal window rotated around its vertical axis. When adults view the two-dimensional Ames window monocularly a powerful illusion is perceived of a slanted window with one side (the larger) closer than the other. Yonas et al. reported that six-month-old infants wearing an eye patch (to remove binocular information) are

twice as likely to reach for the larger side of the distorted window than for the smaller side, suggesting that this depth cue is detected by this age.

Important organizational features of visual perception are the visual constancies, such as brightness, size, and shape constancy, and these are probably present at birth. Evidence of shape constancy at birth was reported by Slater and Morison (1985). They familiarized newborns (mean age 1 day, 23 hours) to a shape, either a square or a trapezium, which changed in slant during the familiarization trials. On subsequent test trials the newborns gave a strong novelty preference for a different shape when this was paired with the familiarized shape, the latter in a different orientation from any seen earlier. This experiment gives clear evidence of infants' response to distal cues from birth.

Finally in this section we will consider perception of the human face. The face is three-dimensional and contains regions of high contrast and when seen by the infant the face is usually animated or moving. These aspects of stimulation are highly salient to newborn and older infants, which ensures that the face will be attention getting and holding, but the interesting question is whether there is a predisposition to respond to the face other than as a collection of stimuli. The question becomes, "Do infants have an innate perceptual knowledge of the face?" There is no easy answer to this question, but certain lines of evidence suggest that we can offer a tentative "yes." Newborns quickly learn some of the characteristics of faces and, 49 hours from bith, they show a reliable preference for their mother's face when paired with that of an adult female strangers. Such evidence of early learning may be only a specific example of a more general learning ability, and we need to look elsewhere for convincing evidence of a special response to the face. Perhaps the most compelling evidence is the suggestion that newborns will imitate adult facial gestures. An early report of neonametal imitation of mouth opening, tongue protrusion, and lip pursing was that of Meltzoff and Moore (1977). Similar reports of newborn facial imitation have also appeared, which implies that infants have an inborn representation of the face on to which they can map their own facial movements.

Objects, Events, and Encounters

From the 1970s there has been a growing increase in the number

of studies which have shown moving and changing "dynamic" stimuli to their infant subjects. Much of this research is inspired by Gibson's theory of Direct Perception. According to Gibson, perception of basic stimulus invariants is direct in that it does not need enhancing as a result of particular experiences. In line with this approach many researchers have argued that visual perception is most meaningful, and is most meaningfully studied, under conditions of change. A few illustrative experiments are described here.

Kellman, Spelke, and Short (1986) described a series of experiments investigating young infants' perception of partly occluded objects. Four-month-old infants were habituated to a stimulus, usually a rod, which moved back and forth behind a block which occluded its center portion. Following habituation, the babies were shown two test displays without the occluder, one being the complete rod, the other being the top and bottom parts of the rod, with a central gap where the occluder had been. On these test trails the infants looked more at the *discontinuous* stimulus (the two rod pieces), suggesting that they had seen the object as being connected or complete behind the occluder in the habituation trials - i.e. they were "filling in" the unseen portions of the rod. Baillargeon, Spelke, and Wasserman (1985) reported that five-month-old infants appreciated the continued existence of a *completely* invisible object. Their babies were shown a solid block, which was then hidden by a screen, but which should have prevented or blocked a moving drawbridge from travelling through a full 180° rotation. They spent more time looking at the "impossible" complete 180° rotation that at a "possible" 120° rotation, suggesting that they were aware not only of the block's continued existence behind the screen, but that they "knew" that its presence constituted an obstacle to the draw bridge's movement.

These experiments investigated infants' understanding of occluded objects by using what have been called "events," defined as dynamic changes to the optic array which do not require activities of the infant observers. An "encounter" is where the observer is actively involved, and an example of infants' responses to encounters is in the "moving room" procedure, in which the whole of the visual environment moves relative to the infant. Butterworth and Cicchetti (1978) tested infants who were seated or held inside a small room in which the floor (and the baby) was stationary, but the three walls and

the ceiling were made to move toward and away from the baby. In these studies the babies (like adults) lose balance, and the loss of balance is always appropriate to the direction of movement (i.e. if the room moves toward them they sway, or lurch forward). Under normal circumstances the flow patterns of visual information that occur in the "moving room" correspond to those that occur when the baby moves, or sways backward and forward: their presence in the absence of the baby's movement disrupts the infant's normal, visually guided, postural control.

In summary, babies enter the world with an immature visual system, but the basic visual capacities develop quickly in early infancy. The immediate visual input is that which impinges upon the flat, two-dimensional retinae, but at no age do infants perceive a two-dimensional world, and from birth infants perceive *distal* cues in the sense of perceiving the world "out there" rather than responding only to proximal retinal cues. Visual perception is active and organized at birth, and babies soon learn to perceive shapes as wholes, and to classify and categorized visual stimuli. Studies of "dynamic" perception tell us that from birth infants use visual information to guide their own movements in the world, and that by four or five months the perceive a world of coherent, spatially connected, unified, whole, and permanent objects.

Coordination Between the Senses

There is evidence that the senses are coordinated from birth. We can distinguish at least two types of sensory coordination. One is where information from one modality specifies or has consequences for information from another. Examples are visually guided reaching, or turning in the direction of a sound source. With respect to the latter, Butterworth (1983) has argued that when newborn babies turn their eyes in the direction of a sound stimulus they *expect* the sound to have a visual consequence – i.e. they expect to see the thing that produced the sound. This suggests that, from birth, infants can differentiate the sensory information given by the different modalities. A second type of sensory coordination is where different modalities provide equivalent information. Meltzoff and Borton (1979) gave evidence for the detection of intersensory equivalence in a habituation experiment. One-month-olds were familiarized to one of two dummies

(pacifiers), the dummy being placed in the babies' mouths. One dummy had a smooth nipple while the other, nubby nipple, had protuberances on it. Following familiarization the babies were shown visual replicas of the two dummies and they showed a reliable visual preference for the one they had previously perceived orally. This experiment gives evidence of early intermediate matching, supporting the view that there is an innate unity of the senses.

Taste, Olfaction, and Audition

Research on taste and olfaction has concentrated on newborn infants, with a paucity of research in the later months of infancy. The newborn infant is as well equipped with taste buds as at any later time in life, and distinguishes between the four basic taste sensations – sweet, salty, sour, and bitter – as shown by differential sucking and ingestion of solutions, and by facial gestures. While taste has a number of "primary" sensations this does not appear to be the case for olfaction, and perhaps because of this infants' responses to a great variety of chemical all factory stimuli have been studies. Space does not permit a detailed description of these studies: suffice it to say that the sense of smell is clearly active from birth. There is also evidence that newborns will learn about and orient toward preferred odours: Macfarlane (1975) found that six-day-old infants would orient toward their mother's breast pad in preference to one from another lactating mother. The chemical sences are probably important in the early control of feeding and serve to enhance the intake of nutritious foods and to inhibit ingestion of non-nutritive, harmful, or toxic substances.

After vision, hearing is the modality which has received the greatest attention, and much of the research has focused on Speech Perception. While the foetus experiences an auditory environment that is quite different from that of the adult, there is good evidence both for auditory perception, and for learning about auditory stimuli, before birth. DeCasper and Spence (1986) had pregnant women recite a passage of speech aloud, in a quiet room, every day during the last six weeks of pregnancy. Shortly after birth (an average of 2 days 8 hours) the infants preferred the mother's previously recited passage to one spoken by a female stranger: in this experiment "preference" was indexed by the baby being more prepared to change sucking patterns to hear the mother's voice. While auditory thresholds at birth

are much higher than for adults, it is known that thresholds reduce during infancy, and that infants display good discrimination between stimuli at differ in intensity and/or pitch during the first year.

Infants also make discriminations between speech sounds which may suggest that they are innately attuned to the phonetic characteristics of human speech. For example, there is an auditory continuum known as voice-onset time (VOT) which separates pairs of phonetic units – i.e./ba/ and /pa/, /ga/ and /ka/, /ta/ and /da/ – and infants as young as one month discriminate these sounds across the boundaries of the continuum, but not within categories. More recently it has been discovered that other species, including chinchillas, rhesus monkeys, and Japanese monkeys (Macaques), give the same discrimination performance, perhaps suggesting that as language evolved it capitalized on general auditory perceptual mechanisms for sound discrimination, rather than developing specific mechanisms for processing speech (Kuhl & Padden, 1982).

The fine distinctions between general auditory, and speech-specific, mechanisms will exercise researchers in the years to come. It cannot be doubted, however, that human infants are uniquely predisposed to acquire speech. Since research on young infants has focused on their discrimination between isolated speech sounds, we do not know when they discriminate words as separate auditory segments. However, comprehension of speech sounds as clearly present in the second half of the first year, and the baby's first words are uttered toward the end of the first year, indicating responsively to word meanings from an early age.

Conclusions: From Sensation to Perception to Cognition

The senses are functioning from birth and, where opportunity allows, are functioning *in utero*. No modality operates at adult-like levels at birth, but such levels are achieved surprisingly early in infancy, leading to recent conceptualizations of the "competent infant." Infants' Cognitive Development begins before birth, and the senses provide information for the gradual organisation and construction of a coherent world. It is never easy to draw a distinction between perception and cognition, but a reasonable view is that early perceptual competence is matched by cognitive incompetence, and that much of the reorganization of perceptual representation is dependent upon the

development and construction of cognitive structures that give access to a world of objects, people, language, and events.

Perceptual Motor Coordination

This expression refers to our ability to generate appropriate muscular commands so that our limbs reach positions in space specified by our perceptual systems. We demonstrate this ability effortlessly all the time, but analyzing and apparently simple action like reaching out and picking up as glass allows us to appreciate the difficulties involved in perceptual motor coordination.

First, we must represent the position of the glass and the hand that is going to grasp it in terms of a common set of coordinates. The primary source of perceptual information about the position of the glass is the image on the retina. But the position of the image does not specify the position of the object in space because the eye can move in its orbit and the head can rotate about the trunk. To know where the glass is relative to the major body axes we must combine information about the retinal position of the image with propriceptive information about the rotation of the eye relative to the head, and the rotation of the head relative to the body. Further, propriceptive information is then required to identify where the various segments of the arm are, relative to the same body axes.

Second, we need to have learnt mapping rules which will specify what muscular commands will take the limb from its current position to the target. By combining the inputs about the position of the glass and the various limb segments with these mapping rules it is possible to generate a motor program which will move the different segments in such a way that the hand arrives at the glass. The program must also contain precise instructions to open and close the fingers at the right time, and in the right orientation, to ensure that the glass grasped rather than knocked over. During the movement visual feedback about the position of the arm may indicate that an adjustment is required to some program parameter. The program must be structured in such a way these modifications can be incorporated into the program during execution.

This complex sequence of events has been investigate with a wide range of the techniques available to experimental psychology. I will review a number of these approaches.

Prism Adaptation Experiments

One component of perceptual motor coordination is the mapping rules which specify the relation between positions in space generated by the perceptual system and motor commands which will move a limb to that position. These have to be learnt initially, and as limbs grow and change size, they have to be modified. One way of investigating this process has been prism adaptation experiments. In these an adult with fully developed perceptual motor coordination views the world through prisms which distort the sensory input. The process of learning to adapt to the new relation between visual input and the appropriate motor output (i.e. learning to modify the mapping rules) can then be studied.

If you view the world through a prism the apparent position of an object is displaced from its real position. In the earliest and most famous of these adaptation studies. Stratton wore prisms which inverted his view of the world. He found that he learnt to adapt to this distortion in the course of a few days and could carry out normal actions successfully. Modern studies tend to use a more controlled version of this task with a wedge prism which produces a simple lateral displacement of the viewed world. If one points at an object viewed through such a prism without being able to see the hand, it will point at the apparent position rather than the real position. With feedback about the error people quickly adapt (i.e.. adjust the mapping between sensory input and motor output) and point in the correct direction. If the prisms are then removed, initial movements show a compensatory error in the opposite direction, the result of learning to adjust for the distortion produced by the prism. The correct mapping is quickly relearnt if feedback is available about the error. It appears that adaptation takes place by adjustment of both the sensory and motor components of the mapping process. Following a well known series of experiments by Held in the 1960s it was believed that adaptation only took place if the subject actively moved his or her arm. It now seems that although active movement is the most effective way to produce adaptation, it can still occur if the limbs are moved passively.

The results of prism adaptation experiments suggest that sensory signals specifying target position, the instructions to muscles, and the consequences of the movements are all stored in

such a way that they allow mapping rules to emerge as an average consensus of past experience. It is clear that the rules are expressed in such a way that they can be used to generate plausible movements by extrapolation from similar situations experienced previously in novel situations. This form of storage demonstrates a general functional property of the nervous system for which Bartlett coined the term "schema". We store our past experience in such a way that we can benefit from any general regularities which it exhibits, not just from the exact repetition of past situations. Bartlett pointed out the advantage of such a system by analyzing tennis strokes. Each one is played in a circumstance which is similar but not identical to past examples. The ideal memory system must allow us to benefit from the similar experiences in the past, without being affected by the fact that the circumstances are not an exact match. The design of physical systems which demonstrate this property (i.e. the ability to generalize from similar but not identical instances) is a major topic of current investigation.

Development Studies

Prism adaptation experiments emphasize the fact that the mechanism which maps sensory space on to motor action is one which learns and adapts easily. It is not hardwired in humans as it is in some simpler organisms which never learn to adapt to distorted sensory input. However, experiments with neonates show that it has some wired-in structure even in humans. Newborns will follow moving targets with appropriate eye and head movements. Remarkably, they also have the ability to initiate certain facial gestures such as protrusion of the tongue. This requires a specific sensory to motor mapping as well as a complex level of visual analysis.

Visual experience quickly allows complex mappings between visual information and motor action to develop. Hofsten (1983) has shown that by the age of five months infants make appropriate preparatory orientation of the hand when about to grasp objects of different shapes; they can time the moment of hand closing to coincide with the moment when they will touch an object they are trying to grasp; and they can launch a successful anticipatory movement to intercept an object coming past them at 30 cm/sec.

Use of Visual Feedback

Prism adaptation experiments study the mapping between position specified by our sensory system and the motor actions used to program reaching movements toward those positions. Such movements can be modified by visual feedback once they have been launched. Two points have received much attention. The first is that there is a speed-error trade-off – faster movements are less accurate. The second is that a certain minimum time is required for feedback to be processed and modify the movement.

The most influential modern study of these effects was by Paul Fitts in 1954. His subjects had to move a stylus as quickly as they could from a start button to a target. Fitts varied the width of the target and the amplitude of the movement from the start to the center of the target. For a given degree of accuracy, movement time was greater of accuracy, movements or smaller targets. He found that the following relations between movement time (MT), target width (W), and movement amplitude (A) fitted the data:

$$MT = a + b\log_2 (2A/W)$$

where a and b are constants.

This relation is called Fitts' law. True to the spirit of his time Fitts assumed that the significant fact here was the logarithmic relationship between movement time and task difficulty and interpreted his result in terms of information theory. More recently various authors have found that the somewhat implausible application of information theory to movement control is not necessary because the data can be modeled by assuming that the movement is controlled by visual feedback. The relation between speed and accuracy is predicted by models which assume continuous correction to the movement, or intermittent corrections, or an initial ballistic movement and a single terminal correction. If the experimenter requires the movement to be made in a precise time, rather than ending in a precise place, then a linear trade-off between speed and accuracy is found in place of Fitts' logarithmic one. In this case the origin of the speed-accuracy trade-off appears to be the increased variability of the impulse required to drive faster movements.

Corrections take time to make. In many experiments it has been found that about 200 milliseconds elapse between an unpredictable

visual signal appearing and the production of an appropriate response. This lag in the human information-processing system can be readily seen in many games; a ball which swings late is very hard to hit in cricket, it is difficult to adjust a shot to a net-cord in tennis, an unpredictable jab by a boxer cannot be dodged. A review of explanations of Fits' law, and of the time to react to unpredictable stimuli, can be found in Chapter 3 of Jeannerod (1988).

Tracking and Manual Control

The lag in responding to unpredictable visual events also sets a limit on the performance of the human operator in any control system where rapid response is required to visual signals. With inputs which vary at more than about 1 Hz, human response lags so far behind the input that the controller would often do better if he or she put no input into the system at all. Studies of the human in the control loop have revealed a number of ways in which people can overcome this intrinsic limitation. Given either a statistically predictable stimulus, or preview of upcoming stimuli, people can prepare a response in advance and fire it off at the appropriate time, thereby overcoming their reaction time lag completely. Thus, for example, when approaching a corner a skillful driver prepares his or her movement of the steering wheel beforehand and executes it to coincide with arrival at the bend.

The picture which emerges from studies of the human operator as a system controller is of a hierarchy of levels of visual motor control. Given preview of the input to come, or memory of previous similar situations, plausible outputs can be generated in open-loop mode, i.e. without using the error signal present at the time. If this response is not quite correct any residual error will be removed by a closed-loop servo mechanism. A review of the visual motor performance of the human operator in manual control systems which does not rely on complex mathematical or engineering concepts is Pew (1974).

The Gibsonian Perspective

A different approach to studying the relation between vision and action was pioneered by J.J. Gibson. He emphasized the importance of the optic flow field at the eye as an individual moved through the

environment. Gibson maintained that the information required for many actions is specified by some property of the optic flow field. Thus, for example, if the optic flow field in peripheral vision streams forward one is probably falling over backward and should make an appropriate postural adjustment. (It has been shown experimentally that manipulating optic flow fields by moving walls of rooms does indeed make people stagger, i.e. it induces them to make inappropriate postural adjustments).

One of the goals of Gibson's followers has been to discover situations in which there are invariant properties of the optic flow field which are always sufficient to specify appropriate action. An interesting discovery of this sort is that as an object approaches an observer, its time of arrival is given by the ratio of the size of the retinal image to the rate of expansion of the image. The key point is that it is not necessary to know the object's distance or speed or size. Time of arrival is specified by the optic flow filed on the retina irrespective of any property of the object. Thus, for example, all objects arriving in 1 sec. will give the same ratio of retinal size to retinal velocity whether they are near and slow, or far and fast. For many acts relative to approaching objects the time of arrival is sufficient for controlling action. We either wish to move before it arrives (e.g. avoid it) or we wish to act at its moment of arrival (e.g. grasp it or hit it). Lee has suggested that the optically specified time-of-arrival variable controls a wide range of behavior from hitting approaching balls and controlling gait to deciding when to brake while driving a car. He has also suggested that it is used by diving sea-birds to ensure that they fold their wings before they hit the sea. These ideas are explored in more detail in Lee (1980).

Gibson and his followers have made some important and interesting observations about the relation between perception and action. However, the views of Gibson himself frequently seem somewhat obscure to readers who are not already convinced that he is right. And the framework in which Gibsonians operate deliberately avoids many of the concepts used by others studying perceptual motor coordination, so it is often difficult to relate work in orthodox and Gibsonian traditions. An introduction (by a non-Gibsonian) can be found in Bruce and Green (1985). Some of Gibson's own ideas can be found in Gibson (1970).

Patient Studies

It appears from studies of brain damaged patients that one region, the posterior part of the parietal lobe, plays a particularly important part in visual motor coordination. Patients with damage to this area often show difficulty in tasks such as pointing to an object, although they have no serious visual or motor deficits. The problem appears to be in *mapping* visual information on to motor commands rather than in the visual system or the motor system themselves.

Pateints with unilateral parietal lesions show normal pointing with the arm under control of the undamaged hemisphere. With the arm controlled by the damaged hemisphere, pointing without visual feedback is systematically distorted away from the target. But with visual feedback, pointing will often be accurate, although slow and with an abnormal pattern of acceleration and deceleration. This observation supports the separation suggested for normal movement control between a mechanism which generates a program by mapping perceptual space on to motor space, and a separate mechanism which can use feedback to correct a movement already underway. A further dissociation, between the mechanism required to point accurately and that required to grasp the object correctly once the hand has reached the target, is suggested by patients who have one skill but lack the other. It seems that these two parts of the reaching motor program are specified independently. A review of results found in patients with visuo-motor coordination problems, and animal work relating to them, can be found in Jeannerod (1988).

Eye Movements

In terms of the underlying neurology the best understood perceptual motor system is that which controls eye movements. One of the roles of this system is to keep an object of interest centered on the fovea, the part of the retina which can perform detailed discrimination, as either the observer or the object moves. The neural circuits which compute retinal slip as the image moves off the fovea are known, as are those which produce corrective signals for the muscles which move the eye. This system is discussed in detail in Miles and Evarts (1979) and Gouras (1985).

Personality and Cognition

Because substantial individual differences exist in cognitive functioning, any complete theory of cognition must account for these differences; among the numerous relevant dimensions of individual differences are those relating to personality. Although personality theorists have diverse views about which personality dimensions or traits are most important, they generally agree that personality traits not only are semi-permanent dispositions to feel and to behave in certain ways, but also relate to individual differences in the emotional, motivational, and intellectual spehres.

There are various valuable reasons for studying the relationship between personality and cognition. First, the study helps to account for and to understand individual differences in cognitive performance. If personality is not considered, then individual differences in cognitive performance are usually relegated to the error term in analyses of variance. The introduction of personality variables offers the possibility of accounting for at least some of the error variance.

Second, the study of personality and cognition can clarify theoretical controversies in cognitive psychology. Suppose, for example, that two separate tasks allegedly measure the same underlying theoretical construct. If a given personality dimension or trait is consistently related to performance on one of those tasks but not on the other, then the implication is that the two tasks are not equivalent measures of the same construct.

Third, it is often theoretically fruitful tc consider personality dimensions or traits from the cognitive perspective. For instance, individuals differ in their characteristic level of anxiety or trait anxiety (Spielberger, Gorsuch, & Lushene, 1970). Analysis of the differences in cognitive functioning between individuals high and low in trait anxiety is useful not only because it is relevant to a description of what is meant by having high or low trait anxiety, but also because it may contribute to an understanding of the development of the anxious or the non-anxious personality. If these differences do play a part, then they may be pertinent to explaining individual differences in trait anxiety.

Dimensions of Personality

While it is to not true that as many dimensions of personality exist as do personality traits, it sometimes appears to be nearly true.

Despite the myriad of personality dimensions that have been psotulated, many of them overlap substantially with one another (H.J. Eysenck & M.W. Eysenck, 1985). The three independent "superfactors" of introversion-extraversion, stability-narcoticsm, and psychoticism define a central three-dimensional personality space. The great majority of other personality dimensions or traits that other theorists have proposed can be related to this three-dimensional framework. Thus, for example, trait anxiety not only correlates highly with neuraticism, but also shows a modest negative correlation with extraversion.

Introversion–extraversion, neuraticism–stability, and psychoticism can be regarded as the major personality factors, at least so far as individual differences in emotional and motivational terms are concerned. However, other theoretical approaches are based on a broader definition of personality and attempt to identify so–called cognitive styles that distinguish one person from another. One of the best known of these cognitive styles is field dependence–independence (Witkin et al., 1954). In essence, field-dependent individuals are said to be less "psychologically differentiated" from others and from the external environment than field-independent individuals. Field-dependent individuals rely heavily on external cues from the environment, whereas field-independent individuals are more responsive to self-generated cues and are less dependent on external stimuli.

Pask (1976) proposed another kind of "cognitive style" that has been influential, one based on a distinction between serialism and holism. Serialists allegedly learn and remember a body of information simply as strings of terms, whereas holists store and retrieve the information as a whole.

In light of the foregoing discussion of some major dimensions of personality, consideration must now be given to their relevance to cognition. (Psychoticism will not be discussed, because there has been insufficient research relating this personality dimension to cognitive functioning).

Trait Anxiety

The relationship between trait anxiety and cognition has been studied extensively. Several differences in cognitive functioning between individuals high and low in trait anxiety have been identified,

especially when the testing conditions involve an element of stress (e.g. failure feedback or ego-involving instructions). Generally, those high in trait anxiety show increased attentional selectivity combined with decreased attentional control and high susceptibility to distraction. If attention can be likened to a beam of light, then those high in trait anxiety have a beam not only narrower but also more mobile than those low in trait anxiety.

Trait anxiety affects both attentional progresses and memorial functioning. At least when stressed, individuals high in trait anxiety have reduced short-term and long-term storage of information, and evidence suggests that retrieval efficiency is reduced.

Some of these cognitive difference can be understood by reference to a hypothesis about Working Memory. Specifically, when individuals high in trait anxiety are stressed, they have reduced available working-memory capacity, especially the central executive component of the working-memory system. A possible explanation for this reduce capacity in high anxiety is that "task-irrelevant processing relating to concerns about performance and negative self-evaluations may be preempting some of the resources of the central executive" (M.W. Eysenck & Mathews, 1987, p. 202).

This approach is limited in that it demonstrates little or nothing about the factors involved in the development of the anxious personality. The relationship between high anxiety and reduced available working-memory capacity presumably occurs because high anxiety reduces available capacity, rather than because reduced available capacity increases anxiety, although the causality possibly operates in both directions. By adopting a rather different approach based on assessing individual differences in the processing of threatening stimuli, more is likely to be discovered about cognitive factors involved in the etiology of the anxious personality. The fundamental assumption is that individuals high in trait anxiety will be more sensitive to threat-related stimuli (e.g. words referring to social or physical health threats) and will devote more processing resources to such stimuli than will individuals low in trait anxiety.

When only one stimulus is presented at a time, individuals high and low in trait anxiety do not differ greatly in their processing of threatening stimuli. However, substantial individual differences occur when threatening and neutral stimuli are presented concurrently: those

high in trait anxiety allocate more processing resources to the threatening stimulus than to the neutral stimulus, whereas those low in trait anxiety exhibit the opposite tendency. In other words, individuals high and low in trait anxiety differ in terms of selective bias.

A related line of research interprets ambiguous stimuli in either a threatening or a neutral fashion. As might be expected, individuals high in trait anxiety tend to interpret such stimuli in threatening ways, whereas those low in trait anxiety generally interpret them in neutral ways.

That high-anxiety and low-anxiety individuals differ in their processing of threatening and ambiguous stimuli is potentially relevant to the etiology of the anxious personality. An individual who typically engages in excessive processing of threatening stimuli at the expense of neutral stimuli and who focuses on the threatening interpretations of ambiguous stimuli is likely to find the environment much more subjectively threatening than do other individuals. Since the amount of anxiety experienced is presumably determined in part by how threatening the environment appears, it can be seen that individual differences in cognitive processing of threatening stimuli may partially determine individual differences in trait anxiety. However, to date no definitive evidence indicates that the causality operates in this direction.

Introversion–Extraversion

Most of the research on cognitive functioning in introverts and extraverts has taken as its starting point H.J. Eysenck's theoretical assumption that introverts are more cortically aroused than exraverts. The findings reinforce the viability of this view of the major difference between introverts and extraverts; they also suggest the value of a hierarchical conceptualization of the information-processing system.

The basic strategy of research in this area has been to compare the effects of introversion–extraversion on cognitive performance with those of other manipulations said to affect the level of arousal (e.g. white nose, and incentive). The term "arousal" is rather amorphous, but there allegedly exists a continuum of arousal ranging from deep sleep or coma at one extreme, to panic-stricken terror or great excitement at the other extreme.

In summary, introverts generally differ from extraverts in that they have greater attentional selectivity, superior long-term memory,

reduced speed but increased accuracy of performance, greater distractibility, and impaired retrieval efficiency. On the assumption that introverts are more cortically aroused than extraverts, it is difficult to determine the pattern of the effects of introversion–extraversion, since no two so–called arousing agents have precisely the same effects on cognitive functioning. However, the modal pattern associated with high arousal consists of increased attentional selectivity, reduced short-term storage capacity, improved long-term memory, increased speed but decreased accuracy of performance, and greater distractibility.

Although the above findings indicate that some similarities exist between the effects of introversion–extraversion and those of arousing agents on cognitive performance, there are a number of differences. Introverts fail to show impaired short-term storage capacity are slow and accurate rather than fast and inaccurate, and show impaired retrieval efficiency. These differences pose problems for an arousal theory of introversion–extraversion.

Revelle, Humphreys, Simion, and Gilliland (1980) identified further problems. First, the implusivity component of extraversion, rather than extraversion per se, is related to cognitive performance. Second, and more damaging, the relationship between impulsivity and arousal in the evening is the opposite of that in the morning: high impulsives are less aroused than low impulsives in the morning, but more aroused in the evening. Thus, it cannot be argued that some individuals are characteristically more aroused than others, since an individual's relative level of arousal is dependent on time of day.

Arousal theory cannot account for the differences in cognitive performance between introverts and extraverts. However, a different and altogether more cognitive theory would probably have greater success. In several studies, introverts and extraverts have performed cognitive tasks either under arousing conditions (e.g. after ingestion of caffeine, or while exposed to intense white noise) or under normal conditions. Typically, such manipulations affect the performance of introverts much less than extraverts, a finding that can be accounted for by assuming that introverts and extraverts differ in terms of an executive control system. Introverts are more likely than extraverts to use their executive control system to maintain performance in the face of varying levels of arousal, whereas extraverts are more affected by the prevailing level of arousal. Greater understanding of the

executive control system may be achieved by comparing the performance of introverts and extraverts in different situations.

Field Dependence

There have been numerous studies concerned with field dependence and cognitive functioning. Most of the studies have dealt with either perception or memory. From their work on individual differences in perception, Witkin, Dyk, Faterson, Goodenough, and Karp (1962) concluded that field-independent individuals are less likely than field-dependent individuals to show distortions and inaccuracies in perception under difficult perceptual conditions.

The findings from memory experiments have been more complex. Commonly, field-independent individuals show higher levels of intentional learning than field-dependent ones; however, the results for incidental learning are less consistent. When social stimuli are presented (e.g. human faces or words having clear relevance to social interaction), field-dependent subjects tend to exhibit more incidental learning than do field-independent subjects. This last finding may occur because field-dependent individuals need support and guidance from others and are therefore particularly attentive to social stimuli (Witkin et al., 1962).

Investigations of the relationship between field dependence and cognition are rather disappointing theoretically, in that they have taken little account of developments within cognitive psychology. For example, researchers have not seriously attempted to decide whether differences in memory performance as a function of field dependence result from individual differences in attention, short-term storage capacity, rehearsal strategies, efficiency of retrieval, or cautiousness in responding.

Moreover, some controversy exists about precisely what is being measured by tests of field dependence such as the Rod-and-Frame Test or the Embedded Figures Test. In particular, evidence increasingly indicates that field dependence correlates negatively with intelligence as assessed by standard intelligence tests; that is, field-independent individuals tend to be more intelligent than field-dependent individuals. To conclude that field-independent subjects generally outperform field-dependent subjects on perceptual and memorial tasks because of their higher level of intelligence rather than because of their cognitive style per se is therefore tempting.

Holism Vs. Serialism

Pask's (1976) theoretical distinction between holism and serialism, introduced previously, defines the precise nature of the distinction as follows:

Some students are disposed to act "like holists" (*comprehension* learners) and others "like serialists" (*operation* learners), with more or less success. There are also students able to act in either way, depending on the subject matter, and if they excel in both pursuits, we refer to those students as *versatile*. It is these distinctions which can, more appropriately, be referred to as learning style.

The essence of this theoretical approach lies in the notion that important qualitative differences exist in learning strategies. The preferred learning strategy or style of a given individual can be assessed by using introspective reports. Although the holistic strategy is generally more effective than the serialist one, both approaches have clear-cut limitations under severe time pressure. Serialists tend to present facts without an overview (which Pask terms "improvidence"), whereas holists tend to produce a personal conclusion without any supporting evidence ("globetrotting").

Reasonable evidence supports the holism-serialism distinction. However, being able to characterize correctly most individuals as either holists or serialists is a hazier distinction. The learning strategy adopted depends not only on the learner's initial preferences, but also on features of the learning situation (e.g. the nature of the task, or the time available for learning).

Conclusions

Investigations of the relationships between personality and cognition serve to increase the understanding of both personality and cognition. Studying how individuals of different personality types vary in their cognitive functioning can enrich personality theory, especial if some of those cognitive differences contribute to the development of personality. Taking account of the ways in which individual differences influence cognitive functioning can advance theories within cognitive psychology. Individuals differ emotionally, motivationally, and in their cognitive styles. Cognitive psychologists ignore these major aspects of individual differences at their peril.

8

Person Perception

Person perception refers to a cognitively oriented subfield of social psychology concerned with the ways that human beings use information about other people to form impressions, make judgements, predict behavior, and interact. The primary focus has been on the implications of various kinds so stimuli (especially personal appearance and behavior) for various kinds of judgements (especially trait attributions).

Person perception research is closely related to Social Cognition, a field directly concerned with the role that cognitive processes play in people's interpersonal involvements. However, person perception tends to focus less on memory organization or other basic cognitive processes than does social cognition. Moreover, whereas social cognition tends to adopt the fundamental concepts, theories, and methods of experimental cognitive psychology, person perception work frequently relies or somewhat autonomous approaches, as described below. Nonetheless, the issues of concern are similar in each realm, as well as in other cognitively oriented subfields of social psychology, such as Attitudes. It is therefore not surprising that these areas are becoming increasingly integrated in the published literature.

Accuracy of Impressions

An important issue in early person perception research was the accuracy of people's perceptions of each other. It seemed reasonable that some people might be especially skilled in judging others' personalities from various kinds of cues, and that such people might be more socially adept and more successful as leaders, personnel managers, or psychological counselors. Empirical efforts to identify such individuals were largely disappointing, and Cronbach's influential 1955 critique of the area brought a halt to most such work. Cronbach noted that the apparent accuracy of a person perception is affected by many factors that are unrelated to perceptiveness about individual's traits. For example, if a perceiver tends to project his or her own attributes on to others, the apparent accuracy of those impressions will fortuitously improve with increases in actual similarity between the perceiver and those being judged.

Another difficulty with studies of impression accuracy is the need for an objective criterion. A researcher who wishes to examine the accuracy of trait judgements, for example, must be able to determine the traits that objectively characterize the individual being judged. Yet the personality traits of common interest in such research are judgemental ones (e.g. kindness and honesty) for which objective measures would be difficult to create. Despite these difficulties, a resurgence of interest has occurred in the accuracy issue. Recent research generally takes care to differentiate various components of accuracy focuses on situational determinants of accuracy rather than on individual differences, and defines accuracy in terms of verifiable criteria such as behavior.

Impression Formation

The preponderance of research on person perception simply identifies the nature of impressions formed from different kinds of information about people, without addressing the accuracy or inaccuracy of those impressions. The following overview is presented with the caveat that many findings are qualified by such factors as the other kinds of information simultaneously available, and the characteristics and goals of the individual forming the impression.

Facial appearance appears to influence impressions primarily through perceivers' reactions to configurations of features, rather than through the implications of individual features. A few exceptions include scars or deformities (derogation is common), hair color ("blonds have more fun"), and spectacles (which lead to inferences of intelligence). Configurations of features may convey age, gender,

and ethnicity (all of which can imply a variety of stereotypical attributes), physical attractiveness (which generally leads to a wide range of positive inferences), and facial "type" (e.g. "baby-facedness" implies childlike qualities such as submissiveness and naivety; Berry & McArthur, 1986). As long as the overall reaction to a face is not negative, merely being exposed to it repeatedly can increase positive feeling about the person.

Facial expressions communicate people's emotional states and their truthfulness or sincerity. The kinds of expressions indicative of many emotions (including happiness, surprise, anger, disgust and sadness) are essentially the same across most cultures. As a rule, positive expression (e.g. happiness) produce more favorable impressions in observers. Facial expression can also "leak" indications that a person is being deceptive or insincere, though even better indications are provided by body movements, voice tone and rhythm, and incongruencies between these several sources of information (Zuckerman, DePaulo & Rosenthal, 1981).

Style of dress, body type, and posture reliably elicit an assortment of inferences. Dress can imply attributes such as self-confidence or non-conformity, and can also prime stereo-types of various occupations or subcultures. In many cultures, height is associated with status, and girth with good-naturedness. Body posture and physical gestures (sometimes jointly referred to as "body language") can convey approachability or decpetiveness. For example, crossed arms are often viewed as a defensive potsure, suggesting wariness, hostility, or tension.

Other subtle or minimal behaviors also have implications for impression formation. A person (especially a female) who makes eye contact and touching gestures will often be viewed positively, unless these behaviors are extreme (e.g. staring), inappropriately intimate, or seemingly manipulative. Similarly, and individual's perceived likeability is increased by closer physical proximity, within narrow limits prescribed by different cultures and different kinds of relationships.

Various nonsemantic aspects of spoken language (often termed "paralanguage") contribute to impressions: lower-pitched voices are often viewed as indicative of strength and composure, while higher-pitched voices and irregular or halting speech rhythms are viewed as

indicative of nervousness; faster speech sometimes implies self-confidence and expertise, though it can also suggest nervousness or insincerity; and dialects or accents can evoke ethinic and national stereotypes.

Naturally, the contents of verbalizations also affect impressions. A person who compliments a perceiver will be more positively evaluated, perhaps even when the compliment is known to be insincere. A person who boasts can sometimes influence perceptions related to the boasted attribute, though frequently at the expense of perceived modesty. "Self-disclosures" of personal secrets often enhance others' evaluations, unless such disclousures seem too personal to be appropriate for a given relationship, or unless they create an uncomfortable demand on the perceiver to make similar disclosures.

More substantive behaviors also influence perceiver's impressions of the actor. However, reactions to such behaviors may reflect complex reasoning processes. Investigators examining such influences have therefore constructed special approaches, called "attribution theories," to explain the impact of observed behavior on person perceptions.

Attribution Theories

The two major attribution theories – Jones and Davis's (1965) correspondent inference theory and Kelley's (1967) ANOVA model – imply that perceivers do not, and should not, always perceive an actor's behavior as indicative of underlying personality characteristics. The theories suggest that an occasional reluctance to draw inferences from people's behavior follows from the multiplicity of potential determinants for any specific act. In correspondent inference theory, perceivers discount socially desirable acts (sometimes characterized as role consistent or expected acts), and acts that produce numerous outcomes that would not otherwise have occurred ("noncommon effects"). The former might reflect social norms or pressure and the latter create too much uncertainty regarding the actual intentions or motivations of the actor. "Correspondent inferences" regarding and actor's traits are predicted to follow most readily from socially undesirable behaviors that cause relatively few noncommon effects.

In Kelley's ANOVA model, perceivers discount "high consensus" behaviros, which are typical of how most people would act in a

situation, and "high distinctiveness" or "low consistency" behaviors, which are atypical of how the actor him - or herself would act in other situations or at other times. The "ANOVA" label reflects the premiss that perceivers evaluate the causes of behavior by attending to patterns of behavior across actors, situations, and time – a process analogous to the statistical technique of analysis of variance (ANOVA).

Although different in approach, these theories both derive from work by Fritz Heider (1958) – especially his discounting and covariance principles. Writers have criticized these theories for ostensibly implying that people make attributions logically and after considerable cognitive effort. Whether or not such criticisms are entirely justified, they did lead the critics to identify logical "errors" in people's attributions, and eventually to develop more "automatic" conceptions of the attributional process.

A wide variety of attributional errors and biases was identified, demonstrated, and labeled with special terms. (These "errors" might be more properly understood simply as deviations from the logic of the major attribution theories.) Researchers showed that people often make the "fundamental attributional error" of inferring actors' traits even from coerced behaviors. Research on "actor–observer differences" demonstrated people's willingness to make trait inferences from others' behaviors that they would not make from their own. Perhaps some "ego-defensive" biases are introduced into self-attributions by people's motivation to think and communicate positive things about themselves. "False-consensus biases" occur because people tend to think that most others will act as they themselves would. This may explain why the kind of "consensus information" to which Kelley referred often seems not to affect attributions as his theory predicted. Finally, some "errors" may occur because people are prone to fixate on information that is especially salient in the environment, accessible in memory, or consistent with expected patterns.

Reference to casual schemata, or sometimes to Connectionist Models Of Memory, is characteristic of recent arguments that trait inferences are made with less deliberation than classic attribution theories seem to imply. Instead, some theories argue, people may derive trait impressions from perceived behaviors through relatively effortless and non-conscious Automatic Processing. For example,

features of a behavior or situation may simply be associated in memory with causes of the events and characteristics of the participants. (The associative structures relating behaviors to traits and traits to other traits are sometimes called "implicit personality theories.") Alternatively, certain patterns of events and circumstances may instantiate schemata which have implications about event causes and participant characteristics. These newer approaches to attribution have their own critics, but the issues they raise illustrate the complexities involved in trying to understand attributions based on behaviroal information.

Combinatorial Processes

Additional issues arise from the multiplicity of cues that can contribute to person perception – cues including appearance, body language, tone of voice, and overt behavior, to remention just a few. One such issue concerns the way in which so many sources of information are combined into a unitary impression. Information integration models suggest that the evaluative implications of individual pieces of information are added or averaged together. People's actual impressions often seem to be more influenced by extreme or negative information than simple integration models would predict. These widespread "extremity and negativity biases" have been explained in terms of several different models, including attribution theory and categorization processes.

Other approaches to combinatorial processes view various pieces of information as interacting, the implications of each piece changed by the others in a manner consistent with Gestallt Principles Of Perceptual Organization. (One example of such interaction is the "halo effect"– the common tendency for a person known to have one positive attribute to be evaluated more positively on other attributes as well.) Research generally confirms that people attend to larger confingurations of cues, rather than to isolated pieces of information, in making impression judgements.

Another issue concerns the way that impressions are affected by the sequencing of cues about a person. Early research suggested a "primacy effect" such that the first information encountered has the most profound influence on people's impressions (even though recent information is often remembered best). An explanation consistent with

the Gestalt viewpoint is that early information creates a general impression that then affect people's interpretation of subsequently encountered cues. Subsequent research confirmed the importance of first impressions, but also suggested that recency effects can occur if informational cues imply that the person being perceived may have changed.

As the Gestalt explanation for primacy effects implies, the processes involved in perceiving information about people can be selective and interpretative. In other words, it is the subjective encoding of cues, rather than their objective implications, that influences impressions. For example, research suggests that ambiguous information is often misperceived as more consistent with other things known about a person, and that consistent information often contributes most to an evolving impression. (However, under certain circumstances, it is inconsistent information that is best remembered). Situational demands and interaction goals also affect the way that information about people is perceived: for example, we tend to think more highly of people we know we will have to interact with.

People are clearly capable of forming impressions from diverse sources of information, and of using these impressions to respond to a variety of different kinds of judgement questions. However, it is unclear exactly what kinds of impressions, and what kinds of judgements, people ordinarily make on their now. Some theorists argue that people spontaneously from trait impressions ("She's kind and smart"), whereas others believe that people normally assign others to categories ("He's an absent-minded professor"). Recent therories of person perception hold that impressions comprise multiple implications and representations at the same time (cf. various chapters in Srull & Wyer, 1988). The implication is that many different kinds of integration and judgement processes may be involved in forming and reporting our impressions of others.

Person Vs. Object Perception

A recurrent debate in person perception concerns the similarity or dissimilarity between the processes involved in person perception and those involved in the perception of inanimate objects. Some theories argue that person perception is not fundamentally different from object perception, whereas others view person perception as a special

instance of object perception – one that brings more complex and specialized perceptual processes into play.

Although one may question the extent to which fundamentally different processes are involved, there do seem to be some characteristic differences between person and object perception. First, in person perception that which is being observed has important animate properties that are frequently absent in object perception: people tend to be active, complex, difficult to fully understand, aware that they are being observed, and capable of misrepresentation. Second, the observer may have different characteristics when engaged in person perception than when engaged in object perception: the observer may be emotioanlly involved or actively interacting with the person, may need to be circumspect in examining that person, and may be aware that the person is looking back simultaneously forming impressions of the observer. As a consequence, then, person perception may often require more cognitive analysis, produce more emotional involvement, and involve more self-reference than does object perception.

In any case, the area of person perception has much in common with object perception and with other areas of cognitive psychology. At the same time, it has made independent contributions to the study of cognitive processes because of the unique issues that arise in studying how people perceive each other.

Piaget, Jean (1896–1980)

Jean Piaget, a Swiss biologist and psychologist, originate by far the most influential theory in developmental psychology. He produced a reasonably simple theory of development which could explain an astonishingly broad range of behavior in children. He was a gift observer of the oddities of children's behavior and realized that these might provide the key to understanding children's intellectual and social development. A third reason for his pervasive influence is that he possessed, and used to good effect, a wide knowledge of many different subjects, being a biologist, and an expert in philosophy and in some aspects of mathematics.

The central idea in Piaget's theory is that logic develops. He thought that there are three major periods in this development of logic. The first, called "the sensori-motor period," covers the first two years of life. During this period the child acquires a basic understanding of

space and time (Piaget, 1954): he or she learns for example that objects exist even when not perceived, that one object can only be in one place at a time, and that particular events and actions have particular consequences. Piaget's best known observations during this period concern the understanding of "object permanence." He reported (a report amply confirmed by others) that the six month baby loses interest in an attractive toy a soon as it is put, for example, under a cover even though the baby could in principle easily lift the cover off and retrieve the object. Piaget's explanation was that the baby does not realize at first that objects go on existing when he or she can no longer perceive them and thus lives in an apparently impermanent and capricious world. Older babies (nine to twelve months) solve this problem but make a curious mistake with another task. If at first they see the object hidden in one place (A) and retrieve it but then see it hidden in a second place (B), they often continue (quite wrongly) to look for the object in the first place (A): this is called the AB error, which Piaget attributed to babies thinking that the action of going to the first location (A) creates the object. This phenomenon led Piaget to the idea that children remember spatial position in terms of their own actions or movements: their idea of space, he argued, was at first egocentric and much of the development during this period consists in shedding this egocentrism. Although the developmental changes during these first two years are very striking, Piaget insisted that the children were acquiring only a "practical intelligence."

The second major stage, "the concrete operations period," begins during the preschool years and ends roughly speaking at the beginning of adolescence. Its two most important features are (1) that the child begins to be able to solve logical problems, and (2) that this logical development happens because of an underlying change in the child's ability to "manipulate" perceptual information. At the beginning of the concrete operations period the child is dominated by his or her immediate perceptual input. By its end her or she is able to imagine what something in his or her environment would look like under changed conditions. Piaget called this new ability "reversibility" and he linked it to a wide variety of logical moves. An example is a task with three sticks (A,B,C) in which children are shown first that A>B then separately that B>C, without seeing A and C together, and are then asked how A and C compare. Young children, Piaget claimed,

cannot make this "transitive" inference because they cannot combine the separate premises: older children succeed because they can manipualte their memories of the separate premises and can imagine the three sticks in an A>B>C series.

Another instance where the development of reversibility is said to lead to the solution of a cognitive problem is Piaget's well known conservation task (Piaget, 1952: Piaget & Inhelder, 1974). Children are given two perceptually identical quantities (A and B), say two identical glasses holding the same amount of water, and then see the liquid in one glass (B) poured into a much narrower container (B1). They are then asked whether there is the same amount of liquid in A as in B1. Younger children tend to give the wrong answer. Piaget concluded that they think that the change in the appearance of the transformed quantity signifies that its actual quantity had changed too. Older children answer the question correctly, and Piaget claimed that they do so because, as the liquid is tipped from one container into the other, they can imagine what it would look like if it were tipped back into the first one: this ability to cancel out a perceived change (reversibility) leads to the proper understanding of the principle of invariance.

Piaget also claimed (Piaget & Inhelder, 1963) that during this period the child begins to be able to work out what a scene seems like to other people looking at it from a different point of view than his or her own. This is described as a change from "egocentrism" to "allocentrism". Piaget extended the idea of egocentrism to social and moral problems and argued that young children, being egocentric, are unable to work out the intentions of another person and therefore judge what that person does in terms of the consequences of his actions rather than of his intentions.

The final period, "the formal operations period," roughly covers the time of adolescence. It is the period of "operations on operations": adolescents begin to be able to think about thinking. This leads them to be able to work out how to test hypotheses, and allows them to see, for example, the value of testing one variable at a time while holding all the others constant. Piaget did establish that there were striking changes in children's success in testing their hypotheses during this period, but he admitted later on that "operations on operations" are not universal. A sizeable number of reasonably intelligent adults

may never achieve formal operations.

This account has dealt only with Piaget's description of developmental changes. Another part of his theory concerns the causes of these changes. Piaget (1975) argued that children learn mainly through their own informal experiences. He claimed that children seek intellectual "equilibrium" which depends on their being able to explain events to themselves satisfactorily. However, they find, from time to time, that they have two mutually incompatiable ways of explaining the same events. Such internal conflicts lead to internal "disequilibrium" and stimulate children to find a new internally consistent way of understanding the events in question. In order to do so the child must employ more sophisticated intellectual strategies, and thus the resolution of the conflict and the return to cognitive equilibrium sparks off intellectual changes.

Piaget's ideas have had their critics. Three kinds of doubt have been expressed. The first is that the numbers or mean ages of the children concerned are rarely given, and some of his most influential work was only carried out on his own children. However, most objections of this sort can be ruled out because his work has been replicated with striking success in systematic research in other universities. Second, it has been suggested, most notably by Donaldson (1978), the Piaget ignored the effects of the social context in which his experiments took place. In the conservation experiment, for example, the child may be misled by the apparent importance which the experimenter attaches to the change in the object's appearance into thinking that it must be right to say that its quantity has changed too.

Finally, Piaget's work has been criticized for faulty experimental design. Most of his experiments have little or nothing in the way of controls. Children's failures in transitive inference tasks may be due to their forgetting the premises rather than to an inability to combine these premises inferentially. The design of the conservation task can also be criticized on the grounds that it tests a great deal more than just the understanding of invariance, which means that children may fail in the task for several different reasons. The task involves memory, and one also needs to make an inference to reach the correct answer, as well as to have grasped the invariance principle. Much of the work on infants can be interpreted in more than one way: the baby's failure

to retrieve a hidden object may simply signify that he or she does not know how to do so rather than that he or she thinks that the object no longer exists (Bremner, 1988).

Boden (1979) pointed out that Piaget was ahead of his time in that all his work was done before Artificial Intelligence (AI) made any noticeable impact on psychological theory. Piaget would certainly have given AI models his enthusiastic attention, and it is a reason able assumption that he would have used them to strengthen the least convincing part of his theory – his causal hypothesis. Nevertheless, the inventiveness and the power of Piaget's theory and of the considerable empirical evidence that he amassed for it demand our respect. He remains the most significant developmental psychologist.

Positron Emission Tomography and Cognition

Positron emission tomography (PET) is a technology which is based on the use of positron emitting radiotracers (typically Fluorine 18 and Oxygen 15) to label brain function. Elegant cognitive models can now be tested with direct, regional brain functioning assessments made with the powerful technology of PET. Questions concerning localization of function are particularly amenable to PET analyses.

The emitted positrons used with the PET technology produce gamma rays which travel in opposite directions 180° apart. When the head is placed in a ring of gamma ray detectors (i.e. the PET scanner), every simultaneous detection at two points in the ring 180° apart is counted. A large number of such detections from all points around the ring results in a mathematical reconstruction showing a slice view of brain metabolic activity. Cognitive activity requires energy from glucose delivered in the blood. Greater neural activity requires more glucose and greater regional bold flow to deliver oxygen and remove carbon dioxide. The tracer uptake period can vary from 40 seconds (using O_{15}, as a blood flow tracer with a half-life of 123 sec.) to 35 minutes (using 18-flurodeoxy-glucose, FDG, as a metabolic tracer with half-life of 110 minutes).

The first use of PET to study lexical processing was reported by Petersen et al. (1988). Normal volunteers (11 females, 6 males) were scanned in four conditions.

In the first condition, subjects were instructed only for visual fixation with no lexical task. In the second condition, words were

presented passively, in either a visual or auditory way. In condition 3, the subject spoke each word presented and in condition 4, the subject said a use for each word presented. Data analyses were based on subtracting scans of each subject in two conditions to see the added brain activity as the conditions required more operations. Three sets of subtractions were reported. Condition 2 minus condition 1 highlighted brain areas used in sensory input and word-form processing. Condition 3 minus condition 2 highlighted output coding and motor control. Condition 4 minus condition 3 highlighted areas of semantic processing.

Group results indicated a small number of highly localized areas of brain activity for each cognitive component addressed in the subtraction sequences. Visual and auditory sensory input activated different areas (condition 2—1) in occipital and temporal cortex, respectively. Output and semantic conditions overlapped between visual and auditory presentations. Output tasks (conditions 3-2) activated rolandic, sylvian, and premotor cortex areas. Semantic tasks (conditions 4-3) activated prefrential and cingulate areas. The results support multiple, parallel routes for sensory specific components of lexical processing, at least for this combined sample of righthanded males and females of normal language ability.

LaBerge and Buchsbaum (1989) used PET to test a specific model of attentional filtering which predicts involvement of the pulvinar nucleus of the dorsal thalamus. Seven normal, right-handed women performed a visual identification task with a filtering display in either the right or left visual field. Each subject was scanned twice, on different days, alternating the side of the filtering display. FDG with an uptake period of 35 minutes was suged to label brain glucose metabolic activity. As predicted, when the filtering display was in the right visual field, the left pulvinar showed more activity than the right pulvinar and this reversed when the filtering display was is the left visual field. A control area for the thalamus, the mediodorsal nucleus, showed no such reversal. No general hemispheric effect using frontal white matter for contrast was found with analysis of variance, nor was there a hemispheric effect in 20 areas of the occipital lobe.

A different attention paradigm using the degraded Continuous Performance Test (CPT) has been used to study normal/psychiatric group differences (Buchsbaum et al., 1982, 1989). In general, areas

in the right frontal and parietal cortex are elevated in normals doing the CPT compared to schizophrenics and compared to other normals doing a control task (CPT stimuli only with no task).

Complex reasoning was studied with PET by Haier et al. (1988) using the Raven's Advanced Progressive Matrices (RAPM). Eight normal right-handed males were compared to matched groups doing the degraded CPT and the CPT control task (stimuli only). Left posterior cortical areas were more active during the RAPM than the other two conditions, consistent with neuropsychological models. Since the RAPM was designed to highlight individual differences in performance, scores on the RAPM were correlated with glucose metabolic rate in all cortical areas. Significant inverse correlations (i.e. the better the cognitive performance, the less glucose used) were interpreted as consistent with brain efficiency models of Intelligence. Individual differences in CPT performance showed much smaller, non-significant, positive correlations with cortical glucose.

In a similar study, Parks et al. (1988) confirmed inverse correlations between glucose use in frontal, temporal, and parietal cortex and verabal fluency scores. However, Chase et al. (1984) reported positive correlations between Weschler Adult Intelligence Scale scores and cortical glucose use in a group of mostly Alzheimers patients.

Additional PET studies of memory, aphasia, brain organization, and other neuropsyschological tasks have been reported. As the PET technology becomes more readily available, more elegant research designs with new radiotracers and with larger sample sizes can be expected. The spatial resolution of PET scanners will improve to about 2 mm so smaller brain areas can be studied.

Several core ideas and methods in cognitive psychology have originated with Michael Posner, who has been affiliated with the University of Oregon since 1965. His work has focused on elementary cognitive processes involved in perception, action, and attention. Two of the most important concepts from his work are mental codes and the distinction between controlled and automatic processing. The methods Posner has used have been as important as the principles he has uncovered.

Information can be used for different purposes. Its visual appearance can be admired, its sound can be enjoyed, and so forth.

This suggests that the brain may be able to represent information in terms of distinct cognitive "codes." Much of what we know about such codes has been revealed through a task developed by Posner and Mitchell (1967). People indicate whether two linguistic stimuli match on some dimension – for example, whether two letters are both vowels. Posner and Mitchell found that when the two letters have the same name (e.g. "A" and "a") the decision is made more quickly than when they have different names (e.g. "A" and "e"). The decision is made still more quickly when the two letters are physically identical (e.g. "A" and "A"). Based on these and related results, obtained both in Posner's laboratory and elsewhere, it has been established that the same stimulus can be transformed into visual codes, name codes, and semantic codes, and that these codes are available at different times.

By presenting a prime stimulus prior to the two items to be compared, Posner and Snyder (1975) obtained evidence for a distinction between controlled and automatic processing. If the prime does not match one of the test items, the time to compare the test items increases, but this is true only when the prime reliably predicts one of the test items. The result suggests that people can deliberately ignore the prime when it is unreliable – an instance of *controlled* processing. If the prime *matches* one of the test items, performance is enhanced no matter how reliable the prime is. Thus the prime can also facilitate processing of the test item *automatically.* A great deal of research on the distinction between automatic and controlled processing has followed up this elegant experimental demonstration.

Another domain in which Posner has studied automatic and controlled processing is the shift of attention from one spatial location to another. By cuing different locations with stimuli of varying reliability and by measuring the speed with which people can detect stimuli at those locations or make eye movements to them, Posner and his colleagues (e.g. Posner, Inhoff, Friedrich, & Cohen, 1987) have identified several processes involved in the control of spatial attention. These include general alerting, disengaging attention from the currently attended site, moving attention to the new target site, and engaging attention at the new target site. Posner, Cohen, and Rafal (1982), Posner, Walker, Friedrich, and Rafal (1984), and Rafal and Posner (1987) found that patients with lesions in specific brain regions show

impairments in one or more of these processes, suggesting that each of the processes is controlled by a particular part of the brain.

Posner, Petersen, Fox, and Raichle (1988) introduced another technique for investigating the localization of cognitive functions in the brain. They obtained positron emission tomography (PET) scans from the brains of volunteers engaged in tasks differing with respect to one or more processes, such as repeating a word or saying aloud a use for the object named by the word (e.g. "pound" in response to "hammer"). Semantic information is required only for the second task, which means that if a part of the brain "lights up" during the second task but not the first, that part of the brain can be identified as a locus of semantic information processing. Through this method, Posner et al. (1988) identified brain locations subserving different cognitive operations. The success of the technique points to close correspondences between mental operations and their biological substrates, and encourages cooperation between cognitive and neural scientists.

Priming

The term "priming" refers to a technique developed to study how context influences performance; the technique involves the measurement of responses to a second (target) stimulus as a function of its relation to a first stimulus (the prime). The prime thus serves as a context for the target. By varying the relations between the two stimuli, and the temporal interval separating the prime from the target, an experimenter can effect close control over the nature of the context effect due to the prime.

Priming has been widely adopted in studies of word and picture recognition. Different forms of priming can be distinguished, by varying the time interval between the stimuli and the probability that the prime and target are related. When there is a low probability that the prime and target are related (i.e. when primes are uninformative) subjects are not able to predict the target from the prime on any particular trail in an experiment. Under these circumstances, any effect of priming can be said to be "automatic." In contrast, when there is a high probability that primes and targets are related (i.e. when primes are informative), subjects can generate expectancies of targets from primes. Priming can then be said to be strategic or expectation

dependent. Automatic priming can occur even when there are short intervals between primes and targets (i.e. with short prime-target intervals, there are few effects of the informativeness of the prime). Expectation-dependent priming tends to be most pronounced with longer prime-target intervals (when the effects of the informativeness of the prime emerge). This difference in the "time-course" of the different effects suggests that subjects take some non-negligible time to generate expectancies of targets from contexts.

Primes can be informative both about the identity of targets and about their time of occurring (if the prime-target interval is constant across a block of trails). To distinguish temporal from identity priming, a *neutral* prime condition can be included, where the prime only serves as a warning signal for the occurrence of the target (e.g. the neutral prime might be a row of Xs while all the other stimuli in the experiment are words). The effects of identity priming can then be assessed relative to the neutral prime baseline. When this is done, both positive (facilitatory) and negative (inhibitory) priming effects can be identified. When the subjects generate expectations of targets from primes, and the expectation on a particular trail turns out to be incorrect (e.g. the prime and target are unrelated and not related as expected), responses are inhibited relative to the neutral baseline. When the expectations are confirmed (e.g. the prime and target on a given trial are related), responses are facilitated. However, under conditions of automatic priming, only facilitated priming tends to occur (when the stimuli are related). Relative to the neutral baseline, there are no costs when primes and targets are unrelated (Neely, 1977).

Thus, automatic and expectation-dependent priming can be distinguished using a number of operational criteria. Automatic priming is fact acting, unaffected by the informativeness of the prime, and produces benefits (when the stimuli are related) without costs (when the stimuli are unrelated). Expectation dependent priming is relatively slow acting, affected by the informativeness of the prime, and produces both benefits and costs (depending upon whether primes and targets are related or unrelated). By studying the relations between primes and targets that give rise to automatic priming effects, we can learn about the properties of stimuli that are processed automatically and are not dependent on particular strategies adopted by subjects.

When automatic priming has been examined various forms of

relationship between the prime and the target have been found to be important. For instance, priming effects occur when the stimuli have the same identity (i.e. the target is a repeat of the prime, e.g. Jacoby & Dallas, 1981), when they are associates of one another (e.g. Fischler & Goodman, 1978), when the target is a word following a highly predictive sentence (Stanovich & West, 1983), and so forth. Such effects are often attributed to the prime "preactivating" an internal representation that mediates target processing. When the target's representation is per-activated, less information is required from the target for a response to occur.

Interestingly, associative and repetition priming differ considerably in their time course. Associative priming tends to occur only when there are short intervals between primes and targets (e.g. 5 seconds or less; Meyer, Schvaneveldt, & Ruddy, 1975). Repetiton priming can last considerably longer, and can be measured when stimuli are represented a day or even a year later. Also, the effects of associative priming are additive with those of repetition priming, so that the benefits due to each alone sum to produce the overall benefit when the associates are represented in an experiment (Wilding, 1986). These results suggest that associative and repetition priming reflect different processes. Associative priming may reflect temporary activation within an associative or contextual recognition system. Repetition priming may reflect longer term persistence within a perceptual recognition system or a specific and relatively long-lasting memory of the perceptual aspects of the prime. Also, repetition priming tends to be much larger if stimulit are re-presented within the same modality (words as words, pictures as pictures), relative to when they are re-presented in different modalities (words as their pictorial equivalent, or vice versa). This last finding can be taken to indicate that perceptual recognition systems are modality specific, with there being separate perceptual recognition systems for pictures, words, and so on.

Priming techniques thus provide a general procedure for examining the kinds of internal representation mediating information processing. Recent work also suggests that priming can occur even when subjects are unaware that they have processed the primes (e.g. if primes are presented too briefly for conscious identification; Cheesman & Merikle, 1985), and it has similarly been adopted to

study the effects of brain damage on human performance – since patients may sometimes be primed by stimuli they fail to identify explicitly. In such cases priming provides a important technique for studying the kinds of information to which we gain access implicitly, or even without conscious awareness.

Proactive Interference (PI)

A term derived from interference theory which describes the tendency for prior learning to interfere with subsequent learning – sometimes known as proactive inhibition. In general PI will be a function of the similarity between information learned on different occasions: the more similar the material the greater the PI.

PI can be noted in everyday life when, for example, we mistakenly give our old address rather than the new address to which we have just moved. However, investigations of PI have centered on controlled studies of verbal learning. Underwood (1957), one of the principal proponents of interference theory, attempted to explain why students who had learnt a list of nonsense syllables showed so much forgetting over a 24-hour period. Because his subjects had done many other similar experiments, he reasoned that their forgetting might be due to PI caused by the number of previous lists they had learnt. His analysis supported this view but it was not the entire explanation of forgetting. Students who had not learnt any previous lists also showed substantial forgetting. Attempts to attribute this to other sources of PI were unsuccessful.

An alternative approach to the pehnomenon of PI comes from studies that have used the Brown-Peterson technique. This examines the ability of subjects to remember small amounts of information over shore periods of time. During these periods the subject performs a distracting activity to prevent rehearsal. If a subject is tested repeatedly there will be a build up of PI in that performance becomes increasingly worse as the number of trails increases. Such effects are more marked if the to-be-remembered information shares a common dimension (e.g. semantic category), but these effects can be ameliorated by increasing the interval between successive trails.

Studies of PI in the Brown-Peterson paradigm led to the view that PI is caused by the development of cue overload. As more and more similar information is presented, the cues of retrieving

information from any particular trail become associated with information from other trials. This theory was supported by a phenomenon known as release from PI. Again using the Brown-Peterson task, this manipulation showed that the gradual build up of PI caused by presenting similar information across trails could be destroyed by presenting information that was dissimilar (Wickens, 1973). According to cue overload theory, release from PI arises because subjects are able to encode the information in terms of a different cue to that used in retrieving information on other trails. However, an experiment by Gardiner, Craik, and Birtwhistle (1972) suggests that release from PI is due to factors that happen at retrieval rather than initial encoding. They showed that a change in the nature of information presented on different trails could still cause release from PI, even when the change was only indicated to the subjects at the point of recall.

PI is therefore a useful descriptive term when discussing the causes of impaired memory performance but, as yet, the mechanism or mechanisms underlying it are unclear.

Production Systems

These are computer programs that make use of If....Then rules or productions to model cognitive pehnomena computationally. Computational models of cognition based on production systems have been applied in a wide variety of topic areas in cognitive psychology; from problem solving and reasoning to learning and language understanding. Before reviewing some of these production system models, let us first examine the basic components of what can loosely be called "the standard production system."

Standard production systems have a Working Memory and a permanent of Long-Term Memory of productions (i.e. If....Then rules). A production rule consists of a "condition" and an "action": the condition specifies one or more conditions that must be met for the action to be carried out. For example, one could have a production for identifying a dog consisting of "IF something has four legs and is furry Then it is a dog." This production has the following condition and action:

Production for Indentifying a Dog
condition:

x has four legs
x is furry
action:
x is a dog

In most production systems, information input from the environment is added to working memory. On a given cycle of the system, the contents of working memory are matched against the conditions of all the productions in long-term memory. If a match is found to one of these productions, then it is fired and the effect of its action is added to working memory. So, if an object in the environment is encoded in working memory as having four legs and being furry, then the information that it is a dog will be added to working memory, when the dog-identifier production rule fires.

In this way, the contents of working memory are modified on successive cycles through the repeated application of the productions in long-term memory. Typically, over a number of successive cycles "rule chaining" occurs; that is, one rule adds something to working memory, that then calls a new rule on the next cycle. If we also allow working memory to contain goals, this rule chaining can easily model problem solving behavior. For example, assume that working memory has the following contents:

Goal: keep safe
x is furry
x has four legs
x is growling

and apart from the dog-identifier rule, there is also an "escape rule".

Production for Escaping
condition:
Goal: keep safe
x is a dog
x is growling
action:
run like hell

These rules can be chained to propose an action that will keep us safe. Initially, the dog identifier rule will add the information that "x is a dog" to working memory after matching on "x is furry" and "x has four legs." Then the escape rule will fire on the next cycle since

the information that matches its condition is present, suggesting the action of running away.

Finally, standard production systems also have special production rules called "conflict resolution rules." On any given cycle, several productions may have conditions that match the contents of working memory. However, most systems will only allow one rule to fire on any cycle. Therefore, conflict resolution rules must be applied to decide which one of rules should be fired. For example, assume we have a third rule for identifying cats:

Production for Indentifying a Cat
condition:
x has four legs
x is furry
x miaows
action:
x is a cat

and the contents of working memory are as follows:
x has four legs
x is furry
x miaows

Both the dog- and cat-identifie rules match these contents, so a conflict resolution rule must decide between them. Intuitively, we would want the cat-identifier rule to fire since it is more accurate. This intuition could be realized by a conflict resolution rule that prefers the production with more matching conditions. Typically, this conflict resolution principle is termed "specificity," as it favors rules that have a more detailed fit to the contents of working memory.

Apart from specificity, there are three other main conflict resolution principles used in standard production systems: refractoriness, recency, and production ordering. "Refractoriness" prevents a rule matching the same elements in working memory over and over again; hence, it is designed to stop looping. For instance, in the above escape example, without refractoriness the dog-identifier rule might match on the furry and four-legged elements again and again, adding to working memory the information that a dog is present repeatedly. "Recency" is a very powerful principle that prefers a rule with conditions that match those elements most recently added to working memory. Thus, in the escape example, on the second cycle

the dog-identifier rule and the escape rule would have come into conflict. If recency were the only conflict resolution rule available then the escape rule would be preferred because one of its conditions would match the most recently added "x is a dog" element in working memory. Finally, one can also decide between competing rules by simply ordering the set of production rules according to some criterion (e.g. order them in terms of importance). Then, when a conflict arises, that rule which is higher up the list is preferred to one lower down.

Although the basic ideas are very simple, production system models can become very complicated. As the number of productions increases, the number of possible interactions between the rules becomes very difficult to control and predict. In this sense, the conflict resolution principles are critical to the successful operation of the system. In artificial intelligence, where production systems are used in expert systems technology, it has been found repeatedly that large production systems will sometimes grind to a halt because of such unforeseen interactions.

Production System Models in Cognitive Psychology

The relationship between production systems and information-processing psychology is a bit of a chicekn and egg situation. It is not wholly clear whether information-processing ideas were formed in the image of production systems or production systems were a particularly good method for realizing information processing ideas.

The close relationship between production systems and information processing psychology can, to a great extent, be attributed to Allen Newell and Herb Simon. These two founders of information-processing psychology were among the first to model cognitive behavior using production systems. The first and most well known of these production systems was the General Problem Solver. GPS was applied to a variety of puzzle-like problems and was a vehicle for the development of Newell and Simon's ideas about the general-purpose heuristics (like means-ends analysis) that underlie human Problem Solving. Production systems, like GPS, instantiate several psychological features that have remained central to many subsequent psychological theories. First, the standard production system has a limited working memory that parallels the limitations of human working memory. Thus, by limiting the number of items working memory can hold, the effects

of information loss in a cognitive task can be simulated. Second, knowlede is instantiated as rules that manipulate symbols in working memory, so differences in knowledge between human information processors can be simulated by varying the number and/or the consent of the rules brought to a cognitive task.

Several further developments of production systems have been made since the initial work on GPS. Anderson's "Adaptive Control of Thought" theory (1976, 1983, 1987) has been realized in a series of production systems, the most well known of which is the ACT* system (pronounced ACT-star). ACT*, which has been used to model skill learning and other psychological phenomena, makes several changes to the standard production system architecture.

The first main modification to standard production systems is the addition of a second long-term memory. In ACT* there are two distinct long-term memories; a declarative memory and a production memory. The "declarative memory," modeled by a semantic network of interconnected concepts, contains "declarative knowledge" whereas the production memory contains a set of productions that make up the system's "procedural knowledge". The second main modification is that working memory is not a distinct entity with a number of slots for holding elements but is simply that part of declarative memory which is currently active.

Third, ACT* uses two main conflict resolution principles. In the first conflict resolution method, preference is given to rules that match the goals currently in working memory over those that do not. The second way in which conflicts between rules are resolved makes use of the differentail strengths of rules. In ACT*, whenever a rule is fired successfully its strength is increased by one unit. Thus, when a number of rules match the contents of working memory, that rule which is the strongest, or that rule which has been most successful in the past, is fired. These ideas about rule strengthening are used to model the ubiquitous effects of practice in cognitive skills.

Finally, unlike standard production systems, the ACT* system can also learn by applying rules to the contents of declarative and production memory. "Proceduralization" results in the formation of new productions from the contents of declarative memory and the trace of applying more general-purpose heuristic rules (e.g. learning by analogy) to a problem situation. "Composition" collapses a long

sequence of productions into a shorter sequence or a single production rule. In the process of "production tuning," existing productions can be modified by the generalization or discrimination of their conditions.

While ACT* is an impressive cognitive architecture, in the range of phenomena to which it has been applied, several other competing schemes have been proposed to challenge it. Recently, Holland, Holyoak, Nisbett & Thagard (1986) have modeled a variety of cognitive processes using a different type of production system. They argue that the most important difference in their system is that it allows many rules rather than a single rule to be fired on a given cycle.

Holland et al. point out that the restricition of firing only one rule on each cycle means that it is necessary to have a single rule for every step in a problem. This seems reasonable, accepting that one can specify all the necesary rules, but it also means that difficulties arise when the system has to deal with novel situations that do not match any useful rule. Recall the cat-identifier rule which states that "If something has four legs, is furry and miaows, then it is a cat." Suppose we have heard the animal miaow and seen that it is furry but we can only see its front two legs. In a single-rule-firing production system with this information present in working memory, the cat-identifier rule would fail to fire. However, clearly we would want the system to be able to "assume" that the entity had two other legs and make the conjecture that it might be a cat. The only way this can be done in a standard production system is to allow partial matching on the conditions of a rule, but this is not a very adequate solution to the problem. As Holland et al. point out, if a system has the following rule:

Production for Crossing a Lake
condition:
Goal: cross a body of water
body of water is about a mile wide
you are a strong swimmer
action:
swim across

and it wants to cross a large body of water, but is a weak swimmer, partial matching on the conditions of the rule will result in a dramatic decrease in the system's life expectancy. As a solution to this problem, Holland et al. allow multiple rules to fire on a given

cycle. By having this "limited parallelism" in rule firing, several different rules can summate converging evidence and can use multiple sources of weak support to arrive at a confident conclusion. Thus, there may be converging evidence for the cat-identifier rule from other rules fired simultaneously on other aspects of the situation. Like the ACT* system, Holland et al.'s system is concerned with learning and problem solving; the framework has been applied to several psychological phenomena – including analogy, reasoning, statistical judgements and scientific discovery – although implemented programs in each of these areas have not been carried out.

The final production system that deserves mention is, fittingly, the latest model to emerge from the work of Newell and his associates. This system, called SOAR (standing for State, Operator And Result), is also designed to be a cognitive architecture or a unified theory of cognition. Like more recent production systems, one of its central features is two learning mechanisms. The first of these is called "universal subgoaling," which can be viewed as another method for conflict resolution. In SOAR, whenever an impass is reached (e.g. when several rules can be applied to the contents of working memory), a new subgoal is established to resolve this impass. In deciding between conflicting rules, therefore, SOAR will set up a subgoal to choose one from the set. This subgoal might be achieved by extrapolating the various consequences of applying the alternative productions and then choosing one on the basis of considering the pros and cons of these consequences. The second learning mechanism used by SOAR, that is reminiscent of Anderson's compilation, is called "chunking". This mechanism replaces long sequences of rules with a single production that does the same task. SOAR has been tested more computationally than psychologically, in the sense that its proponents have used it to model established psychological phenomena rather than to generate new predictions for psychological experiments.

Whenever we find production system models in cognitive psychology we also invariably find a large-scale attempt to model psychological phenomena in a cognitive architecture. Thus, production systems have played a central role in many parts to cognitive psychology. The fate of them as computational modeling techniques for cognitive psychologists rests, therefore, to a large degree on the future of the research program to create cognitive architectures. Two

negative outcomes suggest themselves. First, accepting that unified theories of cognition are possible, production systems may not turn out to be the best vehicle for such an enterprise. Some other modeling scheme may be more appropriate. As Newell himself pointed out in the William James lectures, the great wave of connectionism may drown research efforts using production systems. Touretzky and Hinton (1988), for example, have recently shown that many production system features can be replicated in a connectionist scheme. Second, the current quest for a unified theory of cognition may be radically premature, in which case production systems may be tainted by the failure.

If the current production system architectures succeed, then they will remain a central part of cognitive psychology; if they fail then they will be the dinosaurs of the future. It remains to be seen which of these two rules will fire.

Prepositional Representations

These are abstract, amodal, language-like symbols that constitute the language of the mind; as mental representations they are a universal *mentalese,* the basic code in which all cognitive activities are proposed to be carried out.

One of the hallmarks of information-processing psychology is the attempt to specify the mental representations and processes that are involved in a particular cognitive task. A large part of this enterprise is, therefore, concerned with specifying the mental representations that are manipulated by cognitive processes. Given that people the world over, irrespective of the native language they speak, can understand one another and manifest broadly similar patterns of mental abilities, it has been proposed that there must be a universal code in which cognitive activities are carried out. Furthermore, given that each of us can know the world via a variety of modalities (e.g. sight, smell, and sound) and yet use that acquired knowledge in a modality independent manner, it has been argued that there must be a central code for representing information that is independent of its original source. Many have proposed that these representations are propositional in their format and that they are the medium in which all cognition is carried out.

However, this propositionalist view raises several important issues

about mental representations. First, there is the practical problem of specifying the nature of these representations; how can something which is universal to all languages and independent of the senses be specified? As we shall see, propositional representations are usually specified in terms of the predicate calculus. Second, while it may be admitted that propositional representations exist, it is reasonable to suppose that they are only one among a number of different codes used by the mind; that other modality-specific codes also play a role in cognition. We shall examine briefly both of these issues in the present essay.

Specifying Propositional Representations

Propositional representations characterize concepts in a format that is not specific to any language (whether it be Russain, Irish, or Sanskrit) or to any modality (whether it be vision, audition, alfaction, or touch). How, then, are we to talk about them? When theorists want to be explicit about propositional representations they usually specify them in terms of a logical system called the "predicate calculus." In logics like the propositional and predicate calculi, a proposition is defined as *an entity which represents a single idea.*

In the predicate calculus, propositions are detailed in terms of "predicates" and their "arguments." A predicate is anything which takes one or a number of arguments. Thus, the proposal that "Ronan hit Conor with the pole and the pole was hard" can be represented using the predicates Hit and Hard (note that the capitals denote that it is not the word that is represented but its conceptual content) and the arguments Ronan, Conor and Stick. These can be combined together in a bracketed notation, as follows:

Whenever one has a predicate and a number of arguments combined in this fashion the whole form is called a "proposition"; furthermore, any combination of several such forms is also considered to be a proposition (i.e. the whole of the above expression is also a proposition). It should be noted that predicates can take any number of arguments; they can take one argument, in which case they are called one-place predicates [like Hard (Pole)], or multi-placed [like Hit (Ronan, Conor, Pole)]. The predicate calculus comes in different versions called orders: first-order

predicate calculus, second-order predicate calculus. Fist-order predicates are predicates which take object constants as their arguments (like Ronan, Conor, and Pole), whereas second-order predicates can take propositions as their arguments. Thus, in characterizing the sentence "When Ronan hit Conor with the pole, Conor was hurt" we can use the second-order predicate Cause to predicate the two other propositions:

In conclusion, cognitive psychologists have used the predicate calculus to express *mental, propositional representation.* However, psychologists do not use all the strictures employed by logicians. The detailed axioms of the calculus are not often used and other aspects may be changed. For example, strict adherence to formal definitions of truth, falsity, and formal validity are often ignored or passed over. In short, typically, theorists merely use the notion that ideational content can be stated in terms of predicates taking one or more arguments. It is a moot point whether this is the best way to employ the predicate calculus.

However, one of the main benefits of using the predicate calculus is that any theory that assumes propositional representations can be modeled easily in a computer program. This is because many of the modern artificial intelligence computing language like LISP (Hasemer & Domingue, 1980) and Prolog (Clocksin & Mellish, 1981) have a close relationship to the predicate calculus. In fact, Prolog essentially *is* the predicate calculus with some computational additions. Thus, the use of propositional representations has allowed researchers to be very precise about their theoretical proposals, by permitting them to construct and run computer models of cognitive processes.

One Mental Code or Many?

There are actually two separate issues when we come to consider the question of whether there is one or more mental codes. First, there is the definitional issue of whether it is possible to categorically distinguish between different representational formats. Second, even if one assumes that different codes can be discerned, there is the remaining question about the role these different representations should play in psychological theories. We will consider each of these issues in turn.

Traditionally, propositional representations have been contrasted with "analogical representations." Analogical representations are images which may be either visual, auditory, olfactory, tactile, or kinetic; in short, they are modality specific. The prime example of an analog representation is a visual image, although several distinct types can be identified (e.g. auditory images). Propositional representations are, as we have seen, amodal.

However, beyond this distinction it has often been difficult to specify the exact sense in which the two types of representation differ. It has been proposed that analogical representations parallel the things they represent in a way that propositional representations do not; that analogical representations are *analogous* to what they represent, in the sense that they retain the same structure. For example, a visual image could be argued to retain the spatial structure of a set of objects in the world, in a way that a propositional representation of the same state of affairs would not. However, this distinction is not as cut and dried as it seems because the exact relationship between any mental representation and things in the world is very difficult to specify.

It has also been suggested that analogical representations are continuous and nondiscrete whereas propositional representations can be partitioned into discrete entities. That is, one can take an image and circumscribe any arbitrary part of it and it will still refer to something. A portion of an image of a cigarette will still represent a portion of the cigarette. Propositional representations cannot be divided arbitrarily and continue to represent entities in a similar manner. Rather, they consist of discrete conceptual entities that can only be divided up in ways that are determined by strict rules of combination. However, it is clear that both types of representations must have some rules of combination; perhaps the rules of combination for analogical representation are just more flexible

Setting aside the difficulties of finding definite distinctions between the two types of representations, the assumption that they *do* differ has formed the basis for one of the most controversial debates in cognitive psychology. This so-called, imagery-propositional debate has been prompted by assertions about the singularity of propositional representations in human cognition. In this debate the strong propositionalist view has argued that (a)

conceptions of analogical representations are ill-specified to the point of vacuity; (b) even if they were well specified, analogical representations are epiphenomenal in the sense that, even though they may accompany cognition, they do not play a central causal role in it; and (c) if a significant role can be attributed to analogical representations in cognitive behavior, that role can also be accounted for in propositional terms because analogical representations are reducible to propositional representations.

Several theorists from the analogical or imagery camp have argued against this view. They have made three main counter-claims: (a) that imagery is not a vacuous construct but can be well specified; in particular, it can be stated with sufficient precision to be modeled computationally; (b) that it is necessary to posit a separate, image-based processing and representational system in order to account for the results of numerous experiments on a wide spectrum of cognitive behavior; and (c) that visual images are represented in a spatial medium which has special emergent properties that are not manifested by propositional representations (Kosslyn, 1983).

Several empirical and theoretical attempts have been made to resolve this imagery-propositional controversy, usually by attempting to muster support for one view over the other. However, for several reasons, the issue has remained unresolved and has been abandoned by most researchers as an active topic of debate. This state of affairs can be largely attributed to the empirical intractability of the issue. That is, it is probably impossible to decide empricially between the two views because any cognitive task that is asserted to be the result of one type of mental representation acted upon by a given set of processes can be mimicked by another set of processes acting on a different form of mental representation. Furthermore, part of the controversy may proceed from fundamentally different philosophical stances on the nature of scientific knowledge and the means by which truth is to be achieved. Imagery researchers like Paivio take a strong empiricist stand on scientific knowledge whereas propositional theorists are more rationalist in their leanings.

A final twist to the controversy is that it is possible to admit a propositional view and yet upohold the need for an analysis of cognitive behavior in terms of analogical representations. Johnson-Laird (1983), for example, has made the point that an analysis of cognition solely in

terms of propositional representations would be too detailed. It would be like attempting to write programs solely in machine code rather than in some high-level programing language. People, the maintains, have higher-level analgocial representations, just as programers have high-level programming languages, to reduce the complexity to tasks and their information-processing loads. Psychological theories must, therefore, be pitched at a level that captures the use of these analogical representations, even though ultimately they may rely on detailed propositional representations.

Conclusion

Propositional representation are one of the most centrally used theoretical constructs in all of cognitive psychology. They are either implicitly or explicitly assumed in a wide range of cognitive theories. In language comprehension, they are used to characterize the underlying conceptual reprsentations of verbs and nouns (e.g. Norman & Rumelhart, 1975). In memory organization, they are assumed to underlie the representation of semantic memory, schemata, scripts, and other knowledge structures (e.g. Norman & Rumelhart, 1975; Schank & Abelson, 1977; Anderson, 1983). They are assumed to be central to problem solving and reasoning skills. The irony of this situation is that the extensiveness of their use is paralleled by a paucity of direct tests of their properties. The properties of propositional representations that are only tested indirectly; they are tested as they emerge when propositions are combined into more complex structures (like schemata). In contrast, analogical representations, on to which the burden of proof has often fallen, have had to show that their specific properties are revealed in a wide variety of experimental situations.

9

Reading Development

To some extent the way a child learns to read (i.e. its reading development) will be influenced by the particular language and writing system that is involved. Written English presents particular problems because it is orthographically irregular – that is, there is a very inconsistent relationship between spelling and sound. For example, the letters *ough* can be pronounced in several different ways. Such irregularity means that adult readers have to use a lexical procedure for reading many words rather than being able to rely solely on the use of rules for converting letters to sounds.

There are several different theories about how children learn to read English – and other orthographically irregular languages – but there is general agreement that several different stages or phases are involved. One of the best known models is that of Frith (1985) who proposes a three-phase theory of reading development in which the child uses a different strategy at each phase.

In the first phase, the child uses a "logographic strategy" in which words are recognized as holes. During this first phase, the child builds up a sight vocabulary of familiar words but is unable to make any attempt at reading unfamiliar words. By the second, "alphabetic phase," the child is able to convert graphemes into phonemes and relies on this ability when attempting to pronounce unfamiliar words. However, such attempts will only be successful with orthographically regular words and unfamiliar irregular words will be mispronounced.

Passing through the alphabetic phase is seen by Frith as a necessary precursor for attainment of the third phase of reading development, the "orthographic phase." In this phase, the child develops an orthographic strategy in which words are broken down into orthographic units - strings of letters. Unlike the alphabetic stage, these units are not converted to phonemes. The development of an orthographic strategy allows the child to attempt the pronunciation of irregular words.

One problem with Frith's theory is that she does not clearly explain exactly how reading occurs in the orthographic phase, or how this phase differs from the logographic phase. However, there is evidence for a phase in reading development where phonological decoding is paramount, and a later phase where children rely less on phonological coding. The evidence comes from an experiment by Doctor and Coltheart (1980) in which young normal readers at five age levels (six to ten years) were asked to read short sequences of words and decide whether or not they made sense. The meaningless sentences were of two kinds. Half were sentences like (1a) which, if read by a graphemephoneme conversion procedure, would sound meaningful. The other half were like (1b) and did not sound meaningful:

(1a) Tell me wear he went.

(1b) Tell me knew he went.

The purpose of presenting these two different kinds of meaningless sentence was to see whether children were using a grapheme-phoneme conversion procedure in reading. For children who were using such a procedure, sentences like (1a) and (1b) should seem very different whereas, for children who were reading using a more visual strategy which did not involve phonological recording, both types of sentence should be equally meaningless.

Doctor and Coltheart found that six-year-old children frequently thought that sentences like (1a) were meaningful although they correctly decided that sentences like (1b) were not meaningful. The suggests that they were relying on a grapheme-phoneme conversion procedure in reading. However, the older children in the experiment treated the two types of sentence as equally meaningless, suggesting that, by the age of seven years, most children are using a reading procedure that does not involve phonological recording.

An alternative to Frith's model has been put forward by Seymour

(Seymour & MacGregor, 1984), who argues for the existence of three broad and partially overlapping stages in reading development. Seymour uses the same names for these stages as Frith but there are important differences between the models in the two latter stages.

The initial, "logographic," stage is essentially the same as that proposed by Frith. The child builds up a limited sight vocabulary on the basis of purely visual differences between words. This is followed by the learning of simple grapheme-phoneme correspondences (and analogous phoneme-grapheme correspondences for spelling) in an "alphabetic" stage. These correspondences operate at the level of single letter-sound (and sound-single letter) and so are different from the correspondences occurring in Frith's alphabetic stage which involve larger units. In the final, "orthographic," stage, the child develops a sophisticated model of orthography which includes knowledge of morphology and the use of lexical analogies in reading new words. This latter ability allows the child to read unfamiliar words on the basis of their visual similarity to known words.

Seymour's main interest is in how these different stages of reading interrelate. Models of adult reading often make use of the notion of dual or triple routes. Seymour's model is of a dual lexicon. In this model, the child who is reading by use of a logographic strategy is setting up a logographic lexicon in which words are discriminated one from another on the basis of visual characteristics.

There has been considerable debate about the precise nature of these visual characteristics. One suggestion has been that, in this initial – sight-vocabulary – stage of reading, children recognize words i terms of their overall shape. However, Harris and Coltheart (1986) report data on the reading ability of four-year-old child, Alice, who was at this early stage of reading. Alice could read about 30 words and most of these had been "picked up," rather than being taught. One such word was the name "Harrods" which Alice had seen on the side of buses and on shopping bags from the Harrods store. This word was particularly interesting because Alice had almost certainly only ever seen this word written in one typescript – that used in the store logo.

Alice was presented with this word in an entirely different format, using a different typeface and alternating upper and lower case letters: 'hArRoDs.' If Alice were reading using information about visual shape, she should have had great difficulty in reading this visually unfamiliar version of the familiar word. However, this was not the case, and she immediately read the word as "Harrods."

If Alice's performance is typical of other children at the initial stage of reading development, it is clear that the basis of the visual recognition of words in a sight vocabulary is not overall word shape. This conclusion is supported by data from Seymour and Elder (1986) who suggest another possibility, namely that words are discriminated on the basis of some – or perhaps only one – of their letters. However, Alice demonstrated that she was able to pick out the word "Max" from the following items: mAx, rAx, mOx, mAv. This would only have been possible if she were taking account of all the letters in this word.

Whatever the precise nature of the discrimination involved in the first stage of reading, it does seem clear that it involves visual characteristics of words rather than their sounds; and hence gives rise to a visually organized – logographic - lexicon. In Seymour's model, the child also develops another lexicon. This is initially an alphabetic lexicon – arising from the alphabetic stage of reading – which is a rudimentary grapheme-phoneme convertor based on single letter-sound translations. This lexicon is alter superseded by an orthographic lexicon which contains the more sophisticated orthographic knowledge developed during the third stage of reading development. Once established, both the orthographic and the logographic lexicons remain through life.

Both Frith's model and that of Seymour have been used to explain why some children have difficulties in learning to read; and evidence concerning disordered patterns of reading development has been used to question the assumption that all children learn to read in the manner described by theories of normal reading development. One particular controversy concerns the tole of phonological awareness in reading development.

For normally developing children, learning to read occurs some time after learning to speak. This means that, when children first encounter written words, they already know a great deal about how words sound when they are spoken. Since one part of learning to read involves the development of grapheme-phoneme conversion skills, some researchers have suggested that children's ability to segment

words into phonemes will affect the rate at which they learn to read. This possibility was explored in a study by Bradley and Bryant (1983), who investigated the ability of prereading children to analyse words into their constituent sounds. In one test they presented children with sets of three one-syllable words (such as "hill," "pig," and "pin") and asked them to say which was the odd one out. (The correct answer is "hill" because the other two words both begin with the sound "pi").

Bradley and Bryant found that, when the children's subsequent reading performance was measured after three years of schooling, those children who had initially been good at analysing words into their constituent sounds were better readers than those children who had not been good at this task.

In another study, Bradley and Bryant gave various forms of training to prereading children who were poor at the sound analyzing task. One group of children was trained to analyze words in terms of their sounds; another was trained to analyze words in terms of their meaning, not their sounds. A third group of children was given the sound training and also taught that individual sounds could be represented by letters. Three years later the children's reading progress was assessed and it was found that the progress was assessed and it was found that the group who had received sound training had learned to read more successfully than the group who had been trained to classify words by meaning. The group who had received both sound and letter training had the highest reading scores.

The results of the Bryant and Bradley study suggest that the ability to divide words into their constituent sounds does facilitate reading development. However, there is also good evidence that learning to read increases phonological awareness (Stuart & Coltheart, 1988) so it would appear that reading development both facilitates and is facilitated by children's knowledge of the phonological structure of spoken language.

Reality Monitoring

The ability to discriminate between externally derived memories that originate from perceptions and internally derived memories that originate from imagination has been called reality monitoring. External memories are based directly on sensory information, and represent events that really occurred, objects that were really perceived, actions

that were performed, words that were neard or spoken, read or written. Internal memories are mental constructions, and represent events that have only been imagined or dreamed of, actions that were planned or considered, words that were thought but never uttered. The ability to distinguish between these two kinds of memory is to know the difference between facts and fantasies, and reality monitoring is obviously essential for competence in everyday life. However, the importance of this aspect of memory was largely unrecognized until Johnson and Raye (1981) coined the term "reality monitoring" and began an intensive study of it.

Traditional studies of memory have concentrated on the ability to retain information, but the recent trend toward studying memory in naturalistic real world situations has highlighted the fact that it is often important to remember the source of that information (Schachter, Harbluk, & McLachlan, 1984). For example, in social interactions or at work we may need to remember not only what was said, but who said it, and where and when. Source memory is the ability to remember the origin of information and usually involves distinguishing between a number of possible external sources. Reality monitoring is a special form of source memory involving the distinction between internal and external sources.

Although at first sight the distinction may seem to be obvious, it is, in fact, often a difficult one. This is because many memories are neither purely external, nor purely internal, but involve components from both sources. According to the constructivist view of memory, widely accepted today, an external memory is not a direct copy of the physical information received through the senses. Contructivist theories emphasize the role of elaboration, interpretation, and reconstruction based on prior knowledge in the form of stored Schemata, such that a memory is composed of externally derived sensory information integrated with internally derived knowledge. It follows that the distinction between internal and external memories is relative rather than absolute, and may sometimes be difficult to make. Childhood memories, or anecdotes that have been retold many times, often comprise this sort of mixture of perceived and self-generated material, with the original experience being imaginatively embellished or transformed until fact and fiction can no longer be distinguished.

Reality monitoring characteristically breaks down in

schizophrenia, in dementias such as Alzheimer's, and other mental disorders involving hallucinations or obsessions, and it is temporarily disrupted by delirium or intoxication. In these cases, people may be unable to distinguish between the real and the imaginary in current experience as well as in memory. It has also been suggested that reality monitoring is imperfect in the very young and the very old. And, although normal adults are generally able to distinguish competently between fact and fantasy, failures of reality monitoring in everyday life are not uncommon.

It is sometimes difficult to distinguish between the memory of performing an action and the memory of planning that action, or forming the intention to perform it. People may be uncertain, for example, whether they have actually locked the door, turned off the light, added salt to the soup, taken a dose of medicine, or only thought about doing these things. Similar confusions arise with respect to utterances. People may be mistaken about whether they have actually said something or only intended to say it. Uncertainties or confusions of this kind produce errors of performance, and the direction of confusion affects the kind of error that is made. When a planned or imagined act is mistaken of the memory of a real act, an omission error occurs because the actor concludes wrongly that the act has already been performed. Conversely, if the memory of a performed action is mistaken for an imagined one, the result is a repetition error. In this case, people are liable to take a second dose of medicine or to put salt in the soup twice over. Failures of reality monitoring may also underlie losing objects if the memory of putting objects, such as keys or spectacles, in a particular place is in fact the memory of an intention rather than a memory of an action. The quality of social interactions can also be adversely affected if errors of reality monitoring cause repetition errors such that the speaker tells the same person the same thing more than once.

Johnson and Raye (1981) have put forward a model to explain how reality monitoring decisions are made. They suggested that there are two ways in which internal and external memories can be distinguished. The first method consists of evaluating the features of the memory trace. According to the model, external memories are characterised by being richer in sensory attributes such as sound, color, and texture. They also are accompanied by more contextual

information linking them to a specific time and place of occurrence, and to preceding and succeeding events. Internally generated memories, by contrast, are more schematic and lacking in contextual details. They are also more likely to contain traces of the cognitive operations, such as reasoning, inferring, or imagining, that produced them. Thus, according to the model, the origin of a memory can generally be determined by evaluating the extent to which it possesses these characteristics. However, rules of this kind may not always yield a clear-cut decision. Internal memories can sometimes be unusually clear and detailed, or external memories may be blurred and vague. The failures of reality monitoring observed in everyday life confirm that the two kinds of memory representation are qualitatively confusable, and cannot always be differentiated. A second method for making the distinction is to employ the criteria of coherence and plausibility. If a memory is externally derived it ought to make sense and conform to our knowledge of the world. Internally derived memories can sometimes be recognized because they fail to fit this criterion. Dreams and fantasies, for example, may conflict with natural as or violate common sense.

These strategies may be successful in distinguishing between internal and external memories, but a different strategy may be needed for determining the particular source of an external memory. It is sometimes necessary to remember whether an action was performed by oneself or by someone else, or whether an utterance was made by oneself or by someone else. These decisions involve what is known as the "generation effect" whereby self-generated actions and utterances are remembered much better than other-generated ones. Relying on the fact that self-generated memory traces are reliably stronger, people assume that they would be able to recognize the origin of a memory if it were something they had done or said themselves (I'd know if I'd done it). If they are not confident that the memory trace was self-generated, they attribute the action or utterance, by default, to another person. The attribution of the memory is based on the degree of confidence.

This model of reality monitoring has been tested by experimentally manipulating the characteristics of internal and external memories and observing the effects on reality monitoring judgements. Johnson, Raye, Wang, and Taylor (1979) tested the prediction that people who

are unusually good at forming vivid and detailed visual images should be poor at reality monitoring because the vividness of their self-generated images would make them qualitatively similar to externally derived memories. As predicted, when asked to view objects and to image the same objects for a varying number of times, good images over-estimated the frequency with which they had really seen the objects. The judged frequency was inflated by the number of times they had imagined the object.

Other experiments have been designed to manipulate the amount of cognitive operations in the memory trace. For example, subjects were asked to generate words under conditions which made it either easy or difficult to retrieve the required words. Later, when they had to identify which words they had spoken themselves and which words had been spoken by somebody else, they were more accurate at identifying the words that had been difficult to generate. According to the model, these words were easier to recognize because the memory representation contained traces of more cognitive operations. The model predicts a gradient of confusability such that it is harder to distinguish between memories with more similar origins, and this prediction was confirmed in a study by Anderson (1984). She found that subjects were more likely to confuse two different kinds of self-generated memories, those for performed actions and those for imagined actions, than to confuse a self-generated memory with an other-generated memory of an action performed by someone else.

Experimental studies examining age effects have confirmed anecdotal reports that both young and old have more difficulty in reality monitoring. Young children have been found to make more errors in distinguishing between what they said and what they only thought, although they were successful at distinguishing between self-generated and other-generated utterances and between different external sources (Foley, Johnson, & Raye, 1983). A similar pattern emerged in children's memory for actions. Elderly people are also liable to confuse imagined and performed actions, and have shown a consistent bias in the direction of this confusion such that they tend to misidentify imagined events as real (Cohen & Faulkner, 1989). One possible explanation for these age effects is that children and elderly people are failing to encode enough of the sensory and contextual features that should characterize an external memory. Another possibility

is that the elderly adopt a strategy of "if in doubt decide it is real." That is, they may operate on the assumption that it is better to make the false positive error of deciding an event occurred when it was only imagined than to make the false negative error of denying that it happened when it really did.

Another method for investigating reality monitoring relies on what Johnson (1988) has called the experimental phenomenological approach. This approach involves asking people to complete a questionnaire giving ratings for the sensory, contenxtual, and cognitive characteristics of each memory. It is then possible to compare these ratings for memories of different origins, or to manipulate factors such as the retention interval or the amount of rehearsal and observe the effects on phenomenological aspects of the memory. The results have been consistent with the model, showing that memories of perceived events received higher ratings for sensory and contextual information and for the presence of supporting memories than memories of imagined events. Furthermore, it appears that the distinction is maintained as memories degrade over time by a differential loss rate. Imagined memories lose sensory details faster than external perceived memories so that they can still be distinguished after a lapse of time. Some kinds of REHEARSAL, especially rehearsing the thoughts and feelings associated with an event, appear to reduce the phenomenological difference between the two kinds of memory.

This kind of approach to the study of reality monitoring has also been applied to naturally occurring memories as well as to experimentally induced memories. For instance, ratings of autobiographical memories have shown that recent autobiographical memories can be readily distinguished from imagined events, whereas for childhood memories the differences were weaker. It has also been found that when people are asked how they know whether a particular remembered event is real or imagined, they justify a reality decision by reference to several kinds of evidence including perceptual detail, context emotions, consequences of the even and inferences about plausibility based on general knowledge. Diary studies recording everyday slips of action (Reason, 1979) have indicated that failures of reality monitoring occur most commonly with routine, frequently performed actions. The distinction between performed and planned actions appears to grow weaker when the action is often repeated.

This might occur because routine actions are performed automatically and would therefore lack traces of cognitive operations. This might then make it impossible to distinguish between the memory of locking the door and the memory of planning to lock the door. Thus the general principles of the model proposed by Johnson and Raye are supported by the objective evidence of error rates and by phenomenological evidence in both experimental and natural situations.

The principles of reality monitoring have real-world applications and are particularly relevant to Eyewitness Testimony, where it has been shown that the testimony of a witness may contain a mixture of veridical recall "memories," based on post-event suggestions, of nonexistent objects and happenings. Schooler, Gerhard, and Loftus (1986) found that, when the verbal descriptions supplied by witnesses were analyzed, judges could be trained to discriminate between real memories, and those based on suggestions, by employing the appropriate criteria. The same kind of methods that people use to evaluate their own memories can be applied to other people's memories. Another application of reality monitoring principles can be found in medical history-taking. It is recognized that patients's estimates of the frequency and intensity of symptoms or episodes of illness may be inaccurate. It is likely that patients, especially those who are depressed or solitary, who spend a lot of time mentally rehearsing these events may later be unable to distinguish between real and imagined episodes and so inflate their estimates of frequency of occurrence.

Reality monitoring should be considered as a form of Metacognition since reality monitoring decisions depend on the exercise of metamemory – that is, on knowledge about how memory works. As well as affecting speech and actions, reality monitoring also affects knowledge and beliefs. This is because our judgements about the source of our own knowledge, and the source of other people's knowledge and beliefs, influences the weight we attach to them. Knowledge based on our own real experiences or on information that is judged to come from a reliable external source will be more influential. We are less likely to be convinced by internaly generated beliefs or when an informant is judged to be speculating, exaggerating, or romancing. The ability to make these judgements therefore determines, to a considerable extent, what we

believe. Current research suggests that self-reports of the phenomenal qualities of memories can yield potentially useful data and help us to understand more about mechanisms of memory and metamemory and their functional significance. Another goal of this research is to gain more understanding of what is happening in cases of memory disorder when reality monitoring breaks down. Reality monitoring is an aspect of metacognition which is theoretically important and which is an essential component in the control of action, speech, and thought.

Reasoning

The rule-based process of logically deciding what to believe is known as reasoning. When people reason, they use information that is known or assumed to be true or probably true to support a belief. A distinction is often made between deductive and inductive reasoning. In deductive reasoning, the process begins with a set of statements, called premisses, that are used to infer if another statement, called the conclusion, is valid. Deductive reasoning is sometimes described as reasoning "down" from beliefs about the nature of the world to particular instances. For example, if we believe that winters are cold in Canada, we would use this information to decide that Vancouver, a city in Canada, is cold in the winter. In inductive reasoning, on the other hand, observations and experiences are used to support generalizations. Sometimes inductive reasoning is described as reasoning "up" from observations and experiences to create new beliefs about the nature of the world. For example, if we know that winters are cold in Toronto, Calgary, Saskatchewan, Ottawa, and Halifax, which are cities in Canada, we would generalize from this information to decide that all of Canada is cold in the winter.

Although deductive and inductive reasoning are theoretically distinct processes, this distinction is not particularly useful in understanding how people reason in most everyday contexts. The usual process of reasoning involves repeated alternations between deductive to inductive processes. Our hypotheses and beliefs about the world guide the observations we make, while our observations, in turn, modify our hypotheses and beliefs. Thus, most informal reasoning involves a continuous interplay of deductive and inductive processes.

Deductive Reasoning

When people reason deductively, they use their knowledge of two or more premisses to infer if a conclusion is valid. There are several models of deductive reasoning, each with its own set of rules: categorical syllogisms; linear reasoning; and conditional reasoning.

Categorical syllogisms. There are four types of assertions or moods that are used to describe category membership in reasoning tasks. These four different moods are formed by combinations of positive and negative statements with the terms "all" or "some." They are listed in the left hand column of Fig. 1. As you can see from Fig. 1, a statement is universal if it contains the terms "all" or "no" or "not"; it is particular if it contains the term "some"; it is negative if it contains "no"; and it is affirmative if it is not negative. Thus, it is easy to classify the mood of any statement by searching out the key terms.

Consider the following syllogism which has two premises and a conclusion:

All cognitive psychologists are intelligent.
Some intelligent people are professors.
Therefore, some cognitive psychologists are professors.

In syllogistic reasoning, the usual task is to determine if the conclusion is valid. A conclusion is valid if it must be true when the premisses are true. One way of determining if a conclusion is valid is with the use of circle diagrams which are spatial arrays that represent category membership. Venn diagrams, Duler diagrams, and ballantines are three different types of circle diagrams that are used in syllogistic reasoning tasks. Each of these types of circle diagrams has its own set of rules for combining circle representations of the premisses to determine if the conclusion is valid. Circle diagrams that correspond to each of the four moods are shown in the right-hand portion of the figure. When circle diagrams are used to determine the validity of a syllogism, the circle diagrams corresponding to each of the two premisses are combined. A conclusion is valid if the combined diagrams contain the information in the conclusion.

Syllogistic reasoning is common in everyday conversations. Of course, the statements are embedded in discourse and not labeled as premiss and conclusion. In natural language, terms like "only" and "each" signal a universal mood, terms like "many" and "few" signal a particular mood, and "none" and "no one" signal a negative mood.

Research on the way people solve syllogisms in real-world contexts has shown that there are three types of errors that are common. The most prevalent type of error is the confusion between a conclusion that is valid and one that is true. A conclusion may be both valid and false. Consider the following example:

All Americans maintain an allegiance to the Queen.
All New Yorkers are Americans.

Therefore, all New Yorkers maintain an allegiance to the Queen.

This is an example of a conclusion that is both valid (i.e. it must be true if the premisses are true) and false. It is false because the premisses are false. By contrast, a conclusion must be true when it is valid and the premisses are true.

Another common error in syllogistic reasoning involves transforming a premiss into a nonequivalent form. When most people read statements like "All A are B," they tend to believe that this is equivalent to stating that "All B are A." Transforming a premiss into a nonequivalent form is called an "illicit conversion." Illicit conversions are more likely to occur when the categories are abstract and represented by letters like "A" and "B". It is less likely to occur when natural categories are used. For example, it is easy to see that "All dogs are pets" is not the same as "All pets are dogs."

The third type of common error in reasoning with syllogisms involves belief biases, especially about issues that are emotional. Consider the following syllogism about abortion, a topic about which many people maintain highly emotional beliefs:

Some women who have abortions are remorseful.
Some remorseful women are psychologically disturbed.

Therefore, some women who have abortions are psychologically disturbed.

Although this syllogism is not valid (i.e. it is not necessarily true when the premisses are true), someone who agrees with the conclusion is more likely to decide that it is valid than another person who disagrees with the conclusion. This is an example of the way in which personal beliefs can alter the reasoning process.

Linear reasoning. When a reasoning problem involves an ordered array of terms, it is called a linear reasoning problem. Consider the

following simple example:

Nigel is talled than Roberto.
Henri is shorter than Roberto.
Who is tallest?

In this example, the three boys form a linear array with regard to height so that it is easy to determine that Nigel must be tallest. It is usually helpful to draw a line and mark the relative size of each term for more complex linear reasoning tasks.

Conditional reasoning. When the reasoning task involves "if, then" statements, it is called conditional reasoning because the occurrence of the second term is dependent upon or conditional upon the occurrence of the first term. For example, "If today is Joan's birthday, then she will have a party." The first part of this sentence (the part following the word "If") is called the antecedent, and the second part of the sentence (the part following the word "then") is called the consequent. When the antecedent occurs, the consequent must occur. This is called affirming the antecedent. Similarly, when the consequent does *not* occur, the antecedent must not be true. This is called denying the consequent. In other words, if Joan is *not* having a party, the it must not be her birthday. Errors commonly occur when people incorrectly assume that if she is having a party, then it must be her birthday. This is an incorrect conclusion given the initial statements because it is possible that she is having a party for some other reason. This type of error is known as affirming the consequent. Another common error is called denying the antecedent. Suppose that you learn that it is not Joan's birthday. Many people would conclude incorrectly that she is not having a party, whereas it is possible that she could be having a party for some other reason. A summary of these four conditions with examples of each is shown in Table 1.

Inductive Reasoning

When we reason deductively, we can be certain about the validity of a conclusion. By contrast, when people reason inductively, they are making judgements about the probable certainty of a conclusion. Much of our knowledge about the world was learned through the process of inductive reasoning. Consider the following example:

Evan gets a rash on the days he eats strawberries.
Evan does not get a rash on the days he doesn't eat strawberries.

Therefore, Evan is allergic to strawberries.

While there may be other explanations for Evan's rash (perhaps he eats strawberries on the same days he eats watermelon or perhaps he wears a wool sweater on those days), the premisses make it likely that Evan is allergic to strawberries. Evan could test this conclusion by carefully eliminating other variables and verifying the observation that he gets a rash from strawberries.

In everyday contexts, we need to be able to figure out the causes for a host of events in order to reduce uncertainty (Holland, Holyoak, Nisbett, & Thagard, 1986). People usually have to generate their own list of possible premisses and then evaluate how well the premisses explain the conclusion. These are the same processes that scientists use to generate and test hypotheses about the physical world.

Of the two processes involved, hypothesis generation and evaluation, the generation phase is probably the more difficult. Consider the problem of Evan's allergy. In order to consider strawberries as a likely cause for the rash, he has to detect events that vary in regular manner along with the rash. (In statistics, this is called detecting covariation.) In arriving at the possibility that strawberries are responsible, he may have considered and eliminated many other variables such as a type of soap or the insulation used in the building at his school. This is a high-level cognitive task that requires consideration of many factors that might cause the appearance of a rash. When children learn that their parents will appear when they cry or that they will succeed if they work hard, they are engaging in an inductive reasoning task.

Just as there are errors that commonly occur in deductive reasoning tasks, there are common errors in inductive reasoning. Perhaps the most prevalent error is an insensitivity to sample size when deciding if the conclusion is supported by the premisses. Consider once again the example given above about strawberries as the cause of a rash. If Evan broke out in a rash after he ate strawberries one time, this would be weaker support for identifying strawberries as the cause of the rash than if he had eaten strawberries ten times and got a rash each time.

Another common error in inductive reasoning tasks is the ubiquitious tendency to seek confirming rather than disconfirming evidence. In order to test the hypothesized relationship between a

food and an allergy, for example, most people would look for evidence that the two are related. It is even more critical to look for instances when the two events are unrelated. Are there rashes at other times or with other foods and are there times when strawberries were eaten and there was no rash?

A third common error is an insensitivity to variability. All people and all events are unique. Consider the relationship between smoking and lung cancer. Although there is a strong relationship between these two variables, some people are heavy smokers and will not develop lung cancer. Newspapers and magazines like to feature stories about old people who claim to have smoked two packs of cogarettes a day since a young age and are now approaching 100 years of age in good health. It would be wrong to conclude that there is no relationship because it is not found in every instance.

Everyday Reasoning

We are constantly bombarded with statements that provide support for conclusions. Informally, the premisses or statements that provide support are called reasons. Reasons are provided to persuade a reader or listener that the conclusion is true. The reasons are the "why" part of an argument; the conclusion is the "what" part of an argument. Most of our conversations and virtually all advertising follows this simple principle:

Buy Bif's Beer because it has fewer calories than the other brands.

In this example, the conclusion is presented first: "Buy Bif's Beer." The reason that is offered to support this conclusion is: "because it has fewer calories than the other brands."

All arguments or attempts to persuade a listener or reader to do something or believe something will contain at least one conclusion and at least one premiss or reason to support the conclusion. Sometimes the premisses or the conclusion is implied, and the listener has to supply the missing part of the argument. Real-world arguments may also contain counterarguments or reasons that suggest that the conclusion is not true and assumptions that modify the meaning of the premisses. When people reason (i.e. use reasons to determine the truth or probable truth of a conclusion), they need to consider three criteria for sound arguments:

1. The premisses have to be consistent and acceptable. A premiss

is acceptable if it is probably true or if it is reasonable to believe that it is true. Premisses are consistent if they do not contradict each other.

2. The premisses must support the conclusion. If you read that dogs make good pets because oil has been discovered in Mexico, you would surely wonder about the connection between the premiss and the conclusion. This is an example of a poor argument because the premiss does not provide any reasons for believing that the conclusion is true.

3. Premisses have to be strong enough to support the conclusion. This criterion is sometimes called adequate grounds. Suppose you read that two rock stars are getting divorced. The reason for this conclusion is an observation that they did not have dinner together one night last week. When considered alone, the premiss or evidence for the conclusion is too weak to support the notion that they are getting a divorce.

Thus, in everyday reasoning, the quality of the reasons and the nature of the support they provide for a conclusion are considered in assessing the acceptability of a conclusion. Reasoning is the cognitive process that underlies all rational decision making.

Recall

A characteristic of recall is that it involves the reproduction of information from memory. It is often contrasted with Recognition Memory, which involves not reproduction but instead the identification of a match between information in memory and information that is newly available.

The importance of reproduction in recall poses a problem in studying the recall of material (e.g. pictures) for which reproduction may be poor even in the absence of a memory burden. This factor has probably tended to propel the study of recall toward using readily reproducible linguistic symbols as material to be recalled: letters, numbers, words, sentences, stories. Certainly, verbal materials have predominated in studies of recall.

Two principal methods of experimenting on recall can be distinguished: cued recall and free recall. In cued recall, the person is provided at recall with a cue (or cues) – information that is related in some way to the information to be reproduced. In free recall, no such cue is provided. For example, a person in a cued call experiment

might be shown ten pairs of words, with one member of each pair subsequently being provided as a cue for the other. A person in a free recall experiment, on the other hand, might be shown ten single words and subsequently simply attempt to reproduce all ten.

By varying the number and the nature of the cues that are present in a cued-recall procedure, it is readily shown that the probability of successful recall depends upon the outcome of the various retrieval processes that they set in train. With a free-recall procedure, the role of retrieval processes is not made explicit in this way (one can of course point nevertheless to implicit cues specifying, say, the most recently presented set of words). As a consequence, the use of free-recall procedures has often been associated with a focus upon processes in memory hypothesized to occur prior to retrieval. Chief among these processes have been those concerned with Forgetting owing to limitations of storage.

A number of proposals have been made as to how forgetting occurs owing to limitations of storage. Information stored in memory may in general decay over time. Information in a particular memory store may be lost by displacement when the demands on that store exceed its capacity. Or the preservation of stored information may depend upon the way in which it is initially processed. These three possibilities – the decay approach, the memory-stores approach, and the Levels Of Processing approach – have generally been regarded as exclusive alternatives, although it is not clear that they are in fact necessarily exclusive. Of the three proposals, the memory-stores approach has over the last 30 years proved the most influential. The limited-capacity memory store that have been proposed include Echoic Store, Iconic Store and components of Working Memory such as the articulatory loop.

Before turning to cued recall, we may note a particular variety of free recall. This is serial recall, in which a person attempts to reproduce a set of events in the same other as that in which they occurred previously (rather than reproducing them in an unspecified order, as is usually the case in free recall). Serial recall is the procedure adopted in measuring memory span. A person's memory span is usually defined as the size of a list of items which that person can serially recall without error on 50 percent of occasions. A figure of around seven items indicates the usual order of magnitude of memory

span although, as might be expected, the size of a person's span depends not only upon the person but also upon the nature of the material being recalled.

As implied earlier in this essay, studies using cued recall make it clear that forgetting occurs not only as a consequence of limitations upon storage but also as a consequence of the absence of appropriate retrieval cues (it may be noted incidentally that this kind of forgetting, unlike forgetting due to limitations of storage, may be overcome when the right cue becomes available). Indeed, it has been suggested that in principle all forgetting could be the consequence of retrieval problems. This view can be related to two other approaches to understanding the origins of errors in recall. The first of these is the idea that forgetting may be the consequence of a process of repression (e.g. Freud, 1901/1960). The second is interference theory, which holds that forgetting results from contamination by other information stored either earlier in time (Proactive Interference) or later in time (Retroactive Interference).

When target information is successfully reproduced in cued recall, two possibilities may be distinguished concerning the way in which it is retrieved from memory. The first possibility, which has been termed the Encoding Specificity Principle, asserts that reproduction is successful only if the cue directly matches the original stored information. The second possibility, which has been termed the generation-recognition hypothesis (Bahrick, 1970), asserts that for reproduction to be successful the cue need not directly match the original stored information, but rather need only be related to it in some way. In this latter case, two separate stages of recall can be distinguished. The first stage is one of generation, in which information in memory that is related to the cue is activated. The second stage is one of recogntion, in which it is assessed whether segments of this activated information do in fact comprise the target information. The different retrieval possibilities identified by the encoding specificity principle and the generation-recognition hypothesis have generally been regarded as exclusive alternatives, in a manner similar to that described earlier in the case of forgetting due to limitations of storage. But again it is not clear that exclusivity is appropriate, and in this case evidence has in fact been provided that retrieval can proceed along both the direct-access route envisaged by the encoding specificity principle

and the indirect route envisaged by the generation-recognition hypothesis.

Recency Effect

The term denotes enhanced memory for the last items in a series. When people attempt to remember a series of events, the chance of their correctly remembering a particular event is usually a function of its position within the series. In its commonest form, the shape of this function (usually termed a serial position curve) resembles a letter U: the level of performance declines over the first handful of items (the primacy effect), stays roughly flat over the central part of the list, and then climbs again over the last handful of items (the recency effect).

From the preceding description, it is apparent that the recency effect may alternatively be described by saying that the items in the list which occurred most recently are remembered particularly well. Given that few are likely to find this result surprising, why has the recency effect been for some decades the focus of much theoretical interest in the study of human memory? The answer is that under different experimental conditions the recency effect has been found to wax and wane, to disappear altogether, or even to become negative in value. As more and more is discovered about it, the recency effect bids fair to become the drosophila of human memory.

Since the 1960s, the most commonly used tool for investigating the recency effect has been the method of free Recall. A person is presented with a list which contains, say, 20 words and subsequently attempts to recall these in any order. The level of recall of the last half dozen or so items increases steadily above the level of earlier items, corresponding in all to about three extra items recalled at these last positions. This recency effect may be abolished by interposing a short period of distracting activity between presentation and recall. A very influential interpretation of the recency effect has therefore been that it represents the recall of a small number of items held in a Short-Term Memory store from which they are readily displaced by director activity (e.g. Atkinson & Shiffrin, 1971). In contrast, remaining items are assumed to be held not in short-term (or primary) memory but rather in long-term (or secondary) memory. Consistent with this, it is found that some other factors, such as rate of presentation, appear to differentially affect prerecency but no recency items.

The short-term memory explanation of the recency effect has been extended in a number of directions. As one example, it is found that the effect is somewhat greater over the last one or two items when presentation is auditory than when it is visual, and this has been attributed to auditory items having available to them a further small modality-specific memory store. As another example, the explanation is consistent with the observation of negative recency. If people are given a final-free recall test for all the items from a series of lists, then under these circumstances the level or recall of the last few items of a list may dip below the level for prerecency items. This is what would be expected if those items had earlier had less opportunity to be transferred from short-term to long-term memory, owing to the end of their presentation periods being abruptly reached.

The main problem for the short-term memory account of the recency effect comes from the fact that rather similar effects are observed in a number of other situations that appear to exclude short-term memory involvement, for example when items are learnt slowly over repeated presentations or when a distractor task is performed after the presentation of each item. In these cases it has been suggested that it is the greater distinctinveness of items close to the extremities of a list which leads to their advantage in recall. But if this is so, then a distinctiveness type of explanation could presumably apply also to the results described earlier.

Recognition Memory

The identification of a match between newly available information and previously acquired information held in memory is referred to as recognition memory. It may be contrasted with Recall, in which the previously acquired information is reproduced from memory. Recognition is a more flexible method of interrogating memory than is recall, in that it is readily used even in cases when information (such as that concerning faces) is difficult to reproduce. Recognition differs also from recall in that it is often considered from the standpoint of perception rather than that of memory. Thus, in the study of word recognition, for example, interest often resides primarily in the classificatory processes whereby a written or spoken item comes to be identified as corresponding to an entry of the mental lexicon, rather

than interest focusing on the memorized lexicon itself.

Any act of recognition presumably involves adopting a particular degree of confidence to be satisfied concerning the degree of match between new and old information. Thus, in considering the determinants of recognition performance we may make a distinction between those that relate to the match with memory and those that relate to the translation from match to action. The latter kind of factor is exemplified by the consequences of manipulating recognition procedures in experiments. For example, the fewer the number of alternatives from among which a target item is to be recognized, the higher is generally the level of recognized, the higher is generally the level of recognition. In principle, this relation could arise because guesses are more likely to be correct when there are fewer alternatives. Suppose a person has an underlying probability of matching new and old information, m, but achieves a higher apparent probability of recognition, r, by guessing randomly among n alternatives (one of which is the correct one) when a match has not been detected. Then r is of course an inverse function of n and indeed m could be estimated following a simple correction for guessing as

$$[r- (1/n)]/[1- (1/n)].$$

The preceding classical analysis can be viewed as a form of threshold model, in that a discontinuity between matched and non-matched states is assumed. An attractive alternative is to replace this assumption in a manner similar to that which occurred with the advent of signal-detection theory within psychophysics. This type of analysis can be used not only for the n-alternative forced-choice procedure described earlier but also when the candidates for recognition are considered one by one, and the match/nonmatch decision is conveyed either by a confidence judgement or else by a simple yes/no. There have been some attempts to choose between different theoretical models within the signal-detection theory framework, but empirical discrimination between models has proven quite difficult to achieve. Instead, signal-detection theory analysis of recognition is generally pursued not as an end itself in but only as a means to an end. A particularly tractable model the equal-variance normal-distribution model is employed in what is hoped to be a theoretically neutral role to transform raw recognition data into two types of parameter value. One of these, d' ("d-prime"), relates to the quality of memory evidence

and the other, B ("beta"), relates to the evaluation of that evidence.

In addition to studying the accuracy of recognition, we may also study the speed with which it proceeds. For this, the simple yes/no method of assessing recognition is generally used. For a given recognition test, latency is likely to be a function of accuracy – the greater the accuracy, the greater the latency – and thus a complete specification of recognition performance would involve mapping an extensive speed-error trade-off function. In actuality, this formidable undertaking is rarely attended and instead mapping only of one of its boundaries is attempted (more accurately, a second boundary, if we count as first boundary the level of accuracy observed when recognition performance is not speeded at all). In this latency mapping, the aim is to measure speeds when they are kept at as high a level as is consistent with the maintenance of error-free performance (in practice, the error rate that is observed is usually low, rather than zero).

Some of the most extensive work on recognition latencies has been carried out using a method in which a person is presented with a short list of up to about six items. This is followed by a further item, and the person makes a speeded judgement as to whether or not that item was in the preceding list. Mean latency is generally found to be a linear function of list length, consistent with recognition involving involving the attempted matching of the target item with each item of the memorized list in turn. However, it has been shown that these results are consistent not only with serial scanning but also with a number of alternatives, including the possibility that attempted matching of all memorized items proceeds in parallel, but that the greater their number the longer it takes for them all to be completed.

Finally, one empirical phenomenon that promises to throw light on the underlying relation between recognition and recall is that of recognition failure – or, more fully, recognition failure of recallable information. When the same memories are probed, recognition performance is usually better than recall performance. However, recognition superiority may on occasion be reversed, and is in any case generally incomplete. A proportion of items is recalled but not recognized and, further, the size of this proportion appears to be a parabolic function of the overall level of recognition. A number of explanations of this result have been suggested. It has, for example,

been shown that a function of this type would arise if Recall could occur in two different ways, one route (namely, generation-recognition) being dependent upon recognition and the other (namely, direct-access) being independent of recognition.

Reductionism

Reductionism is the process of trying to explain events occurring at one level in terms of events at a more basic level. Explaining a car accident in terms of the laws of physics is an example of reductionism. In cognitive psychology, the search for causal explanations and attempts to specify the mechanisms that underlie cognition tend to generate reductionist explanations. However, reductionist explanations are not always causal. They may take the form of Analogies whereby our understanding of cognitive processes is illuminated by the fact that they are analogous to simpler, more basic mechanisms.

In physiological reductionism cognitive processes or aspects of behavior are explained in terms of physiological structures and functions, including both biochemistry and neurophysiology. Thus, for example, biochemcial reductionism provides a biochemical explanation of depression, whereby the overt behavioral symptoms are caused by a noradrenaline deficiency. Similarly, biochemical abnormalities are offered as explanations for alcoholism and other forms of addiction. Critics object that biochemical reductionsim ignores other causal factors, such as social and environmental ones, that are implicated in those conditions, so the reductionist explanation is incomplete. However, a same-level explanation, couched solely in terms of social factors, is equally incomplete. It is generally agreed that a complete account of depression or alcoholism must include causes at several different levels.

Neurophysiological reductionism allows cognitive processes of perception, memory, and language to be modeled as neural networks. Malfuctions are explained in terms of underlying neural pathology, and many cognitive processes such as Priming, interference, or reaction latencies are explained in terms of neural properties such as thresholds, levels of activation, and inhibition. One objection to this kind of reduction is that the mapping of the cognitive event to the neural event lacks precision.

Information-processing reductionism provides an analogy rather

than a causal explanation. Cognitive processes such as selective attention, recoding, storage, and retrieval are represented as a sequence of stages in the familiar box and arrow models. Clarification is achieved at the expense of drastic simplification. Some salient aspects of the cognitive processes are abstracted out and represented, but others, such as individual strategies, intentions, and emotions, are neglected.

Machine reductionism, whereby cognition is explained in terms of computational processes operating on symbolic representations, achieves a precise and detailed mapping of cognitive processes on to computational ones, but is also open to the objection that many features of human performance are missed out of the equation.

Animal reductionism has been applied to human social behavior with ethological observations of the simpler social interactions of animals supplying interpretations for analogous behavior in humans. The advantage of this approach is that experimental manipulation is less ethically constrained, but species differences may make extrapolation invalid.

Non-reductionist, same-level explanations may be phenomenological, or humanist, explaining behavior in terms of feelings, beliefs, or desires. However, this approach is also incomplete. It may identify correlations between same-level events but fails to specify causal mechanisms.

Rehearsal

When attempting to remember verbal information, people show a natural tendency to repeat the information either aloud or silently to themselves; psychologists call this kind of behavior "rehearsal" and there have been many experiments investigating its importance in memory function. Rundus (1971) examined the relationship between rehearsal and free recall of a word list. He found that, overall, the extent to which people recalled words correlated with how often they were heard to rehearse those words during the learning process. However, rehearsal rate bore no relation to how well the last few words in the list were recalled, thus indicating that rehearsal is an unimportant factor in determining recall of very recent information.

There are occasions when an experimenter may wish to prevent a subject from rehearsing during a memory experiment. One example is experiments which attempt to measure the rate at which people

forget information. Clearly, if rehearsal is allowed, there can be no clear time point from which forgetting can be measured. This problem can be overcome with the use of a distractor task in which, following presentation of the information, the subject is required to engage in some irrelevant verbal activity such as counting backwards in 3s prior to recall. This appears to be an effective means of preventing rehearsal but one must note that the degree of distractor task difficulty does influence the extent to which memory is affected. In addition, distractor task performance must be monitored so as to prevent concurrent rehearsal.

Rehearsal is thought to improve memory by increasing the likelihood of forming associations between the information being learned. This form of rehearsal is often termed elaborative and it can be distinguished from maintenance rehearsal where repeating items is not thought to promote retention but merely retain information in consciouness. There are doubts, however, about whether this form or rehearsal can occur (Nelson 1977).

Many theorists have identified rehearsal as a crucial process in the transfer of new information from temporary to permanent storage. Unfortunately there are difficulties with this view. One problem is that much of what we remember cannot be easily rehearsed, either because it is too complex, such as the content of a speech, or because the information cannot be described verbally, as is the case with smells. Furthermore, the concept of rehearsal suggests that learning is an intentional process in that what we remember is determined by what we decide to rehearse. Intentionality does play some role in learning but evidence from incidental learning studies indicates that much of what we learn occurs without any intention on our part to do so.

Rehearsal is best regarded as a specific learning strategy that can promote the retention of verbal information but it does not constitute a general principle of memory function.

Retroactive Interference (RI)

A term derived from interference theory, describing the tendency for new learning to interfere with memory for information learned earlier. RI is sometimes known as retroactive inhibition. The amount of RI observed will increase as the number of learning trials interpolated between learning and test increases. The amount of RI observed will also be influenced by the degree of similarity between the information being tested and the information presented during the interpolated learning sequence: the greater the similarity the greater the amount of RI.

The mechanism or mechanisms responsible for RI have not as yet been discovered. However, it remains a useful descriptive term when examining the various factors that can contribute to forgetting.

10

Schemata

Schemata consist of structured groups of concepts which constitute the generic knowledge about events, scenarios, actions, or objects that has been acquired from past experience.

Bartlett (1932) made use of schemata to explain why it is that, in understanding and remembering stories, people tend to reconstruct the story to fit in with expectations based on their prior knowledge and past experience. The original story undergoes processes of rationalization, deletion, elaboration, and distortion which, according to Bartlett's view, are shaped and guided by pre-existing schemata. Bartlett's ideas were ahead of his time. In the prevailing behaviorist ethos of that period, such mentalistic notions were rejected as being too vague, subjective, and unobservable. However, in the cognitive, psychology of the 1970s, with the growing interest in mental representations, schemata have been reinstated and modern versions play a central role in current theories of memory.

A schema can represent any kind of knowledge, from simple knowledge about, for example, the shape of the letter A, to more complex knowledge about topics like political history or physics. Schemata can be linked together into related sets, so that the schema for "table" would be linked to other schemata for "furniture," "rooms," and "houses." These groups of schemata may be organized hierarchically with high-level general schemata (like "shopping") subsuming more specific, lower order schemata (like "buying new shoes"). A schema consists of relations together with slots, or variables,

which can be filled with optional values. So the "buying shoes" schema would include the relation "buy" and variable slots which could be filled with values like "walking shoes" or "trainers." Schemata also supply default values which are the most probable or typical values and these are assumed if a slot is unfilled. In a schema for "table," for example, the slot for "material" would take the default value "wood." In this case, if somebody tells you about a table you will tend to assume it is made of wood unless you are specifically told it is made of plastic.

According to schema theory, schemata influence the way that new information is processed in a number of ways. The schema that is currently activated guides the selection of what is encoded and stored in memory, so that information relevant to that schema is more likely to be remembered than irrelevant information. The schema provides a framework within which the information can be stored and which can be used at retrieval to guide search processes. The effects of schema driven processes are that information becomes more abstract and general: specific details drop out and the features that are common to other similar experiences are retained. Memories are normalized to fit the expectations derived from pre-existing schemata, so bizarre or unexpected features are omitted or transformed. Missing information is added to the memory representation as the schema supplies default values.

Schank and Abelson (1977) extended the idea of schemata to explain how knowledge of complex event sequences is represented. These knowledge structures, called "scripts," represent the elements common to repeated experiences of events like eating in restaurants. A script consists of a sequence of goal-directed actions which are causally and temporally ordered and includes the actors, objects, and locations that are typically invoived. Like schemata, scripts can be broken up into hierarchically organized scenes and subscripts. Script elements function as default values and allow us to supply missing elements and to infer what is not explicitly stated. A narrative about a restaurant experience need not state that the diners sit down or that they pay for their dinner, since these actions are already present in the pre-existing script.

Experimental evidence for the psychological reality of schemata and scripts has accumulated. Bower, Black, and Turner (1979) asked

subjects to generate scripts for events like visiting the dentist or attending a lecture, and found good agreement about the component actions and sequencing. In tests of memory for scenes or events, people often falsely remember having experienced stereotypical elements that are part of the script or schema, but which were not part of the actual experience. The strongest evidence comes from findings that memories can be dramatically manipulated by supplying inappropriate schemata or no schemata at all (Bransford & Johnson, 1973) or by changing schemata (Anderson & Pichert, 1978).

In spite of this supporting evidence there are some problems with schema theory. The theory does not explain how schemata are acquired in the first place or how any sense can be made of unique or novel events for which no schema exists. A poorly specified process of abstraction or induction is assumed to underly schema acquisition. Another difficulty is that whereas schema theory predicts that schema-consistent, script-relevant information should be remembered best, it is often the unexpected, unusual, or deviant aspects of experiences that seem to be particularly memorable. To accommodate this finding the so-called schema-plug-tag model (Graesser & Nakamura, 1982) proposes that a specific memory trace consists of a pointer to the relevant generic schema plus a tag, or tags, encoding any deviant or novel aspects of the experience. In this way the specific details of a particular experience can be recovered. A number of current models of memory conform more or less closely to this pattern, incorporating a generic memory representation (the schema) linked with a specific memory representation.

Self-focus and Self-attention

The concept of self is one of the oldest and most basic in all of psychology. The word pertains simultaneously to the sense of personal continuity that characterizes every individual's personality; to the organized body of knowledge that everyone has about who one is, what one feels and believes, and who one wishes to be; and to the subjective sense of being at the center of the experiences in which one is involved. The self also exhibits a curious and unique property that has been termed reflexivity. That is, the self has the capability of somehow turning backward on itself, of taking aspects of its own content and its own functioning as the object of its awareness. When

a person is making use of this reflexive capability, that person is in a state of self-focused attention.

Background

The self's reflexive capability has captured interest at least since the time of William James. Though many eminent theorists made reference over the years to the potential importance of self-directed attention, the idea led to virtually no empirical work until, in 1971, Wicklund and Duval first published experimental research on the possibility that self-directed attention might exert an influence on people's behavior and judgements. The success of these studies led to an expanded program of research and theoretical work on their part (Duval & Wicklund, 1972). These efforts sparked the interest of others, and a good deal of further work ensued.

Relationship to cognitive psychology. The work described here was done by social and personality psychologists rather than cognitive psychologists. Unlike most researchers in "social cognition," those studying self-focus have only rarely adopted the research techniques of cognitive psychology and have examined aspects of behavior that are relatively molar rather than molecular. With few exceptions, the work has tended to focus on regulation of people's overt actions and on subjective judgements such as emotional states, attributions, and the like.

The theorists working in this area have also been less precise in their treatment of the concepts that underlie the construct of self-focus than most cognitive psychologists might prefer. That is, they have not fully articulated either theories of the self or theories of attentional processes. What is of interest is the consequence of directing more versus less attention to the self. No stand is usually taken on the question of whether attentional resources are divided or are instead timeshared. Indeed, even the usual assumption of a relatively fixed attentional capacity often seems pro forma.

Operationalizations. The typical procedure for increasing self-focus is simply to place subjects before a stimulus that serves to remined them of themselves: e.g. mirrors, TV cameras, or audiences. The introduction of the manipulation usually is timed so as to coincide with the point in the experiment at which a brief state of self-focus should in theory be most relevant to the behavior of interest. Although

unorthodox, these manipulations do produce the psychological state they are intended to produce.

Another technique was also developed to vary self-focus in research. In 1975, Fenigstein, Schier, and Buss published the Self-Consciousness Scale, a measure that captures individual differences in the tendency to spend time thinking about the self. The individual differences measured by this scale are indicative of differences in the ease or the frequency with which people slip naturally into states of self-focus.

Researchers who use both experimental and personality techniques have adopted a terminological shorthand to distinguish between the two operationalizations. In general, the term "self-awareness" is used to reflect experimental manipulations; the term "self-consciousness" is used to reflect individual differences. Terms such as self-focus and self-attention are used more indiscriminately. We have adopted this convention throughout our discussion here.

Cognitive Consequences of Self-Focus

What are the consequences of attending inward to the self? Because the informational content that defines the self has many aspects, the phenomenology to self-directed attention can take several forms, and the consequences of self-focus can vary considerably. The effects of self-focus range, in fact, from very simple to very elaborate. We begin our discussion by addressing relatively simple effects, which we have characterized here as "cognitive" consequences of self-focus.

Salience and weighing effects. A very simple consequence of self-attention is to cause a self-aspect that is salient to seem phenomenogically more prominent. This tendency toward greater phenomenological prominence is reflected in several different ways, depending on what aspect of self is situationaly salient.

In some circumstances, the most salient aspect of self is some transitory internal state. For example, when people have been exposed to an affect-inducing stimulus, affect tends to be salient as an aspect of the self's experience. When attention is self-focused after a feeling state has been created, the feling is subjectively more intense. On the other hand, there are also occasions when people are led to expect internal states that fail to occur. In these situations, self-focus appears

to lead to greater awareness of the *absence* of the expected condition. Thus, self-focused people are less likely to be misled about what is happening inside them.

Sometimes what is salient is not so much an internal state as a more global sense of the self as an entity that plays a role in events in the day to day world. In some instances, this awareness means that the self becomes more salient as a potential causal force. Thus, when people make causal attributions for hypothetically experienced events, greater self-focus seems to produce a greater weighting of the self as a causal agent, and causes greater attribution of causality to the self (e.g. Duval & Wicklund, 1972).

If self-focus enhances the salience of the self's involvement in events, self-focus should also increase the extent to which people view themselves as being the target of others' actions. There is substantial evidence that this is so (Fenigstein, 1984).

Self-focus, self knowledge, and articulation of self-knowledge. Does the scrutiny of self-focus provide a clearer or more accurate reading of stored information about the self? The implicit assumption underlying this possibility is that enhanced self-attention prompts a more through search of relevant memories. The result would be a more accurate self-report to one's prior actions, or of one's more general action tendencies. This hypothesis has been supported in a number of studies, using self reports of such qualities as hostility dominance, and sociability to predict actual behavior in subsequent sessions (reviewed in Carner & Scheier, 1981).

An additional derivation from this line of thought is that attending to the self over a long period should promote greater articulation of self-knowledge (Nasby, 1985). Indeed, it is arguable that this is why self-reports of very self-conscious people are more valid than those of less self-conscious people. That is, having spent so much time thinking themselves, people high in self-consciousness have well developed and highly articulated schematic representations of themselves.

The idea that self-consciousness leads to better developed representations of the self leads to several predictions, two of which have been tested in research paradigms adopted from cognitive psychology. In one paradigm, subjects in an initial phase rate the self-descriptiveness of a large number of adjectives. In phase two, they

view a subset of the adjectives, and indicate (yes or no) whether each is self-descriptive. After a distractor task, another set of adjectives is presented in phase three, some of which had been presented in phase two (old) and some of which had not (new). The task in phase three is to identify which items are old and which are new.

The tendency to misjudge new words as being old is termed the "false alarm" effect. It is most common among words rated initially as highly self-descriptive. This tendency is strongest among subjects high in self-consciousness. Among words previously rated as *not* self-descriptive, in contrast, highly self-conscious subjects were more likely to correctly judge new words as new. Taken together, these findings are consistent with the idea that highly self-conscious persons have a more articulated representation of themselves than do less self-conscious persons (Nasby, 1985).

A similar conclusion emerges from an incidental learning (depth of processing) paradigm, in which subjects make a decision about each of a series of words, decisions ranging from superficial (involving only shallow processing of the word) to more complex (involving deeper processing). After a distractor task, subjects are given a surprise recall test for the words in the first task. Incidental recall (reflecting deeper processing) is particularly high for words applied to the self in the first task (via the question "Does this word describe you?"), a finding that is commonly termed the "self-reference" effect. Hull and Levy (1979) found that the self-reference effect is stronger among subjects high in self-consciousness than among those lower in self-consciousness. Again, this finding is consistent with the idea that highly self-conscious people have highly articulated self-schemas.

Self-Regulation Consequences of Self-Focus

A more complex set of consequences follows from self-focus when self-attention is invoked under other circumstances. Because these effects appear to have an impact on the guidance of behavior, we will characterize them here as self-regulatory in nature.

Attention to standards and adjustments in self-regulation. A person who is engaged in goal directed activity has adopted some value, standard, or point of reference as a guideline for his or her action. In these circumstances, self-focus appears to have two interrelated effects: it enhances the tendency to compare one's present

state or action with the salient point of reference, and it promotes an adjustment in one's behavior so that the former comes into closer conformity with the latter.

Evidence that self-focus causes an increased tendency to compare one's current behavior with salient reference values is difficult to obtain. The difficulty stems from the fact that the postulated comparison is an internal mental check, a process of monitoring that takes place at an abstract level and is not directly observable. Indirect access to this event can be obtained, however, in situations where the abstract comparison requires concrete information that is available to the subject but must be sought out. A series of studies based on this reasoning has yielded evidence that seeking of comparison-relevant information is increased by self-focus–before starting a task, while engaged in a task, and even after the task is finished (Scheier & Carver, 1983).

The comparison between behavior and standard is presumed to lead to behavioral adjustments (when possible) that reduce discrepancies between the two values. This discrepancy reduction effect was one of the first consequences of self-directed attention to be shown experimentally (Wicklund & Duval, 1971) and it has since been obtained in numerous conceptual replications. Subjects follow instructions more closely, adhere more closely to salient norms, and adjust their actions more in response to systematic situational pressures, when they are more self-focused at the time than when they are less so.

Why does self-focus lead to discrepancy reduction? Two explanations have been offered, which differ substantially in their metatheoretical underpinnings. According to Duval and Wicklund's (1972) view, any discrepancy between one's current state and a salient standard creates an aversive drive state. The larger the discrepancy, the more aversive is self-focus and the stronger is the drive state. One option for reducing that aversive state is to reduce the discrepancy.

Carver and Scheier (1981) have proposed a different view (a view more compatible with the thrust of contemporary cognitive psychology): that the discrepancy reduction effect is a straightforward illustration of the process of feedback control. In a feedback loop, a sensor perceives a current quality. The perception is compared against a reference value. If a discrepancy between the two is discerned, an

adjustment is made in a behavioral output. If the adjustment is appropriate, the subsequent perception of present state either will be closer to the reference value or will not deviate from it all.

This, according to Carver and Scheier, is what happens when attention is self-focused in a setting where a reference value for behavior has previously become salient. People monitor the fit between what they see themselves doing and what they intend to be doing, and make adjustments as needed to diminish discrepancies between the two. In this view, the discrepancy reduction effect is a natural aspect of the feedback-based self-regulation of action.

Self-focus and Disengagement. Though self-directed attention promotes behavioral conformity to salient standards in many circumstances, sometimes the opposite effect occurs. This is most likely to occur when movement toward the goal seems beyond one's capabilities or is precluded by some aspect of the situation one is in. When a giving-up response is evoked, self-focus enhances the extent to which it is carried out. As a result, self-focus sometimes results in diminished persistence or even in withdrawal from the behavioral situation entirely (Carver & Scheier, 1981, 1986).

Carver and Scheier's interpretation of this sort of effect is that it involves a second critical element besides self-attention. In this view people continue to exert effort (and thus are facilitated by self-focus) until they become convinced that they are unable to attain the goal they are pursuing. Only when they pass this psychological watershed of loss of confidence does the impetus to disengage begin to emerge and self-focus yield a decrement in behavior. Thus, the important co-determinant of this effect of self-focus is an unfavorable outcome expectancy (Carver & Scheier, 1986).

Sometimes the disengagement impulse that follows from these unfavorable expectancies is expressed indirectly, rather than overtly. In some behavioral settings, overt withdrawal is prevented by implicit social sanctions. In other settings physical constraints prevent behavioral disengagement. The Carver and Scheier position is that in these settings the disengagement impulse is often expressed covertly, via activities such as daydreaming or off-task thinking (Fig. 1). If the person is supposed to be engaged in a task during this period, the off-task mentation can result in performance impairment.

The model displayed in Fig. 1 appears to be consistent with the

results of research that examined the consequences of self-focus under circumstances of difficulty or adversity. That is, even under conditions of adversity, if subjects have confidence of being able to overcome the adversity, self-focus enhances efforts and outcomes. If subjects are doubtful enough about attaining their goals, however, self-focus is associated with reduced efforts and impaired outcomes.

We should make one further point here, which deviates slightly from the theme of self-focus per se, but which also has implications for that theme. Disengagement (behavioral or mental) often is a viable response to impediments. Life would be impossible if people did not give up some of the goals they took up. On the other hand, it is difficult to disengage from a goal which is proving difficult or impossible to attain but to which one is very committed. In this situation, when attention is self-directed a cycle occurs. The person focuses on the comparison with the desired goal, assesses expectancies, finds them unfavorable, experiences the impulse to disengage, cannot execute it fully because of the importance of the goal, and is precluded from moving toward the goal by the disengagement impulse. Because the goal is important, it is repeatedly reconfronted and focused on. This cycling produces a phenomenology of self-deprecatory rumination and enhanced distress.

As a concrete example, consider the situation of a test-anxious student taking an exam, whose most fervent wish is to escape from the test room. Though he or she does not do that, the student lapses into off-task thinking as a way of disengaging mentally. This withdrawal of mental resources makes failure more likely. Moreover, even this mental disengagement cannot be sustained for long, as attention repeatedly returns to the test and thence to the unfavorable expectancies that induced the disengagement impulse in the first place.

It has long been recognized that test-anxious persons experience the resulting phenomenology of self-deprecatory rumination while in test settings. (Wine, 1971). Though this phenomenology is often termed self-focus, we believe that such a label for that phenomenon is too simple and is potentially misleading (Carver & Scheier, 1986), because it implies that self-focus is antithetical to task focus. The fact is that self-focus impairs performance under some circumstances, but *enhances* performance under other circumstances – and does so even among the test anxious.

One final point, which stems from considering disengagement and discrepancy reduction effects together. We have characterized the difference between these two classes of phenomena here in terms of differences in people's expectancies for the outcomes of their goal directed efforts. It is important to recognize that this characterization places these effects of self-attention squarely in the tradition of expectancy-based analyses of motivation. From this viewpoint, self-focus effects are not a curious and somewhat esoteric side topic. Rather, they are embedded in a major perspective on behavior.

Public and Private Aspects of Self. There is one more important issue to address before closing our discussion of self-directed attention. This issue stems from the multifaceted nature of the self. That the self is multifaceted should already be apparent, given the range of topics we have touched on thus far. However, the issue has also been treated in a more systematic manner than is apparent from our discussion to this point. More specifically, a distinction has been made between public self-aspects and private self-aspects (Fenigstein et al., 1975).

Private aspects of the self are self qualities that are covert, hidden, and bear on personal desires and preferences. Public aspects of self are self qualities that are overt, more explicitly open to others' scrutiny, and bear on self-presentational desires or communal values. This simple dichotomy appears to capture a fundamental and important division within the self, although it will be obvious that each of these self-facets can also be further subdivided.

Given this dichotomy among self-aspects, a dichotomy can also be imposed among instances of self-focus. At any given moment, it is possible to direct attention selectively to one or the other of these facets of the self. The consequences of self-attention can be expected to depend on which aspect of self is taken as the object of one's inward focus. This statement is simple, and the processes to which it refers are also simple, but the idea has far-reaching implications.

To go further, we must first consider how the distinction has been operationalized in research. The individual difference measure of self-consciousness incorporates the distinction in its design. Separate scales of the Self-Consciousness.Scale measure tendencies to attend to public and private aspects of the self, respectively. The existence of separate scales implies that the two tendencies are not assumed to

oppose each other. Indeed, the scales correlate positively, implying that people who tend to think about one aspect of self also tend to think about the other aspect of self. On the other hand, the correlation is usually not so high as to make it difficult to study the effects of each tendency separately.

Many researchers have now concluded (based partly on convergence between effects of manipulations and effects of dispositions) that experimental manipulations of self-focus also embody the public-private distinction. Small mirrors, placed in such a manner as to make it clear that there is no hidden observer behind them, appear to make people selectively attentive to private aspects of the self. TV cameras, the sound of the subject's own voice, and the presence of observers (particularly evaluative observers) make people selectively attentive to public aspects of the self.

Behavioral consequences of the public-private distinction have been examined in a substantial number of studies. Of greatest interest is a set of studies that focused on situations in which the private self and the public self can be expected to exert *opposite* influences on behavior. Due to the juxtaposition of competing values in these situations, it is very important which self-aspect the person takes as focal. The results of these studies are strikingly uniform. When placed in situations that engender this sort of conflict, people act in accord with their private beliefs and perceptions to the extent that they are focused on their private selves. They act in accord with consensual preferences and self-presentational considerations to the extent that they are focused on their public selves.

Applications to the Conceptualization of Pathology

An outline of principles relating to self-attention should include at least some mention of two additional literatures. One of them deals with the effects of alcohol, the other with depression. Neither case is a simple application of the ideas already outlined – in each case the theorists have developed statements of their own – but both cases fit reasonably well with the ideas already outlined.

Alcohol and the reduction of self-awareness. The first of these literatures grows from the hypothesis that an important consequence of alcohol ingestion is a reduction in self-awareness. This hypothesis, developed by Hull (1987), accounts for a variety of observed effects

of drinking. For example, if self-focus promotes self-regulation (as we argued earlier), a reduction in self-focus via alcohol should cause behavior to become more poorly regulated. Consistent with this, it is well known that people's behavior when intoxicated is more responsive to cues of the moment, and less responsive to personal values and principles, than is the behavior of the same people when sober.

More important, Hull and his colleagues have gone on to investigate directly the effects of alcohol consumption on various manifestations of self-awareness in a series of studies. In our view, the most interesting of the findings concerns the idea that people use alcohol as a vehicle for what Hull calls "the strategic avoidance of self-awareness." We said earlier that if people feel unable to attain desired goals, they feel an impulse to disengage from those goals. Diminishing self-focus via alcohol would produce an effect similar to successful disengagement. Moreover, this tactic can be used even when actual disengagement is precluded. This strategic use of alcohol tends to put certain people (those who have a strong motive to disengage) at an elevated risk for sustained drinking problems. The considerable support that Hull (1987) has obtained for his predictions make this theory an important candidate for further exploration.

Depression and Self-focus. The second literature to be addressed here bears on depression. It has frequently been noted that depression is correlated with self-consciousness. Pyszczynski and Greenberg (1987) have proposed a mechanism by which that association comes to exist. In their view, depression develops when a person loses some important source of self-worth (thus having a large discrepancy between perceived self and desired self) and is unable to move toward regaining it. The person becomes immersed in self-deprecatory rumination (and concomitant distress) in the manner we outlined earlier.

As this pattern of negative self-focused thought begins to solidify and stabilize, the person's self-image becomes more negative, and the person tends to focus selectively on negative self-attributes. The result is a "depressive self-focusing style," in which the person is more likely to seek out self-awareness after failure than after success (Pyszczynski & Greenberg, 1987), a pattern that is opposite to that found among non-depressed persons. This selective focusing on the negative helps to maintain and even exacerbate the depression. Though the precise reason for this selective focus on the negative remains to

be determined, this approach to understanding aspects of the phenomenology of depression seems a promising one.

Semantic Memory

Perhaps the first formal definition of semantic memory was provided by Tulving (1972, p. 386) who described it as:

> The memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts, and relations.... Semantic memory does not register perceptible properties of inputs, but rather cognitive referents of input signals.

The term semantic memory has been used very broadly in modern research memory. At a descriptive level the term has much in common with other attempts to distinguish between the representation of knowledge about the world, and the events that gave rise to that knowledge (cf. the philosopher A. J. Ayer's distinction between habit and event memory).

Semantic memory is one dimension of Tulving's (1985) tripartite division of longterm store into episodic, semantic, and procedural memory.

At an experiential level Tulving has attempted to distinguish each of the forms of memory in terms of the nature of conscious experience that accompanies retrieval. Episodic memory is automatic in that recollection of an event reflects some awareness of a contiguity between the even and other features of the temporal record. Semantic memory is noetic – conscious awareness is involved but retrieval of information is not associated with any temporal contingencies. Retrieval from semantic memory does not therefore require access to the learning episodes that provided the original basis for that knowledge. In contrast, episodic memory requires, ipso facto, access to the individual's record of personal experience. Finally, procedural memory is anoetic because access to this store of knowledge proceeds without conscious involvement.

At the experimental level it seems relatively easy to think of tasks which unambiguously tap semantic memory. Defining the meaning of words, answering quiz questions, and solving

mathematical problems are all acts of memory which do not appear to require the retrieval of specific personal experiences before we can come up with an answer. However, when experimenters have attempted to distinguish semantic memory from episodic memory in normal human subjects difficulties have arisen. The primary problem is that episodic and semantic memory, as conceived by Tulving, are likely to be highly interactive. Consider the simple semantic memory task of lexical decision which requires the subject to decide whether a string of letters is an English word or not. Experiments have shown, for example, that repeats of the same letter string produce faster responses – a finding that can, in part at least, be attributed to subjects remembering their previous (episodic) encounter with that simulus. For reasons such as this the separable existence of semantic and episodic memory has been difficult to establish through experiments on normal people.

These difficulties have led proponents of semantic memory to emphasize clinical data in support of their arguments. The human amnesic syndrome can be characterized by a gross impairment in the ability to remember both new information and loss of memories acquired before the illness or accident that gave rise to the memory impairment. However, finer analysis of the deficit reveals that some abilities that would be classified as semantic memory, namely language and general intelligence, are largely unaffected. In contrast, memories that appear episodic in nature are grossly impaired. These findings have led a number of researchers, notably Tulving (e.g. 1985) to propose that semantic memory is selectively spared in amnesia.

Unfortunately there are a number of flaws in this argument. First, the demonstration of normal language and intelligence is usually based on performance levels attained on the Wechsler Adult Intelligence Scale (WAIS). However, WAIS primarily tests information that has been acquired early in life. Thus, rather than suggest intact semantic memory, normal WAIS scores may simply represent the sparing of early overlearned information.

The separable existence of semantic memory, at least in Tulving's terms, is further undermined by more detailed analysis of amnesia. If semantic memory is spared in amnesics, one would expect retrograde amnesia (loss of memories acquired before the brain trauma causing amnesia) to be restricted to event-based knowledge. There are several

lines of evidence that show that this is not the case. Amnesics do, for example, have difficulty defining vocabulary that has come into use in the decade or so preceding the brain injury that caused their memory loss. Also they have great difficulty recognizing once-familiar people – a response that could not have conceivably depended on episodic memory when the patient's memory was intact. Finally, attempts to show that semantic memory can continue to function in the absence of episodic memory have been completely unsuccessful. The famous patient HM became amnesic in 1953 and is still alive at the time of writing. Despite constant exposure to the evolving English Language HM has learned only three new words and attempts to teach him new vocabulary have met with complete failure (Parkin, 1987).

Despite difficulties in asserting the independent existence of semantic memory, the term is still accepted, relatively unambiguously, as referring to that aspect of long-term store concerned with the representation of language and other aspects of knowledge. It also provides a useful umbrella term for a wide range of research enterprises concerned with understanding the representation and organization of knowledge.

Semantic memory was an important research topic in the 1970s and may prove so again with the advent of neural network models. The principal aim of semantic memory research was to examine how knowledge about the meaning of concepts was organized and how this knowledge was utilized. Ideas about the nature of semantic memory have been heavily influenced by developments in the parallel field of Artificial Intellegence. A primary feature of a computer is that it can store and retrieve information. Given that this is something that humams also do it seemed reasonable to suppose that the representational system used by a computer might provide insights into the system used by the brain itself.

Collins and Quillian (1960) developed a computer-based hierarchical network model of semantic memory called the Teachable Language Comprehender (TLC). In TLC, information is represented as various combinations of three types of structure – units, properties, and pointers. A unit corresponds to anything that would be represented by a noun or noun phrase (e.g. a concept, an object). A property is something which can qualify the nature of a unit (e.g. "is blue" can qualify "sea"). Pointers provide the links between units and properties

(e.g. "sea" "is blue"). Within TLC the meaning of a unit is therefore expressed in terms of the properties associated with that unit.

It is not difficult to imagine how vast this network would have to be in order to represent the typical semantic memory of an individual (although, to be fair, this is a problem for all theories of knowledge representation). To simplify matters Collins and Quilljan considered how their hierarchical network might represent knowledge about one category – animals. They proposed that networks would be organized hierarchically. Thus, the units "bird" and "fish" would point to the common node "animal." In order to be economical it was assumed that properties common to all animals would only be associated with the animal node. At subordinate units only those properties unique to that unit would be represented – a principle which was termed cognitive economy. Thus, in order to establish that a fish has skin it would be necessary to move from the "fish" unit to the "animal" unit – this follows because "has skin" is a property found in all animals. In contrast, knowing that a fish has fins, directly from the "fish" node because, barring a few prehistoric monsters, "has fins" is a unique feature of fish.

TLC made clear predictions about the time required to retrieve different items of knowledge. It asserted that verification time for a sentence such as "X has Y?" would be a positive function of the distance between the unit X and the unit at which property Y was stored. Taking our above example, "A fish has fins? would be verified more quickly than "A fish has skin?" because, in the former case, the property is stored at the same level as the unit.

Initial experiments measuring verification times suggested that TLC might be a valid model of semantic memory. It was found that subjects did take longer to confirm more "distant" relationships (e.g. "does a fish have skin?") than closer ones (e.g. "does a fish have fins?"). However various pieces of evidence failed to support TLC. First, subjects did not always respond as if their semantic memory was organized hierarchically. Consider the two sentences "a bear is a mammal?", "a bear is an animal?" TLC predicts that the former should be verified more quickly because the distance between bear and mammal is shorter than that between bear and animal. In fact the reverse was found. However, one might argue that people are not natural zoologists and that the concept "mammal" may not be

represented very clearly. Many people, for example, think of birds as separate entities from other animals.

There were a number of objections to the approach of Collins and Quillian. At a theoretical level it was pointed out that many concepts do not have a clear set of defining features and could not be fitted easily into a hierarchical structures. Another problem was the "typicality effect" in which people took longer to confirm that atypical exemplars belong to a category than typical exemplars, e.g. "ostrich" is more difficult to define as a bird than "robin." Results such as this forced Collins and Loftus (1975) to abandon the TLC model and substitute it with the "spreading activation" theory of semantic memory. They rejected the hierarchical structure and instead proposed that semantic relatedness was the primary basis of organization. Within the model the meaning of a given unit or "node" is expressed by the nature and strength of its relationships with other units and properties. Thus, a typical instance of category (robin) is confirmed as such more rapidly because it shares many of the defining features (e.g. "flies") of a bird. Atypical exemplars (ostrich) have less in common and therefore result in a slower decision process.

The notion of spreading activation refers to the process by which activation of a given node also activates other nodes to which it is linked. The extent of spreading activation of other nodes was assumed to be directly proportional to the strength of the links between those nodes and the initially activated nodes and, furthermore, the weaker the link the longer the time required for activation. Unfortunately, only evidence supporting the second prediction has been found.

Despite its limitations, the spreading activation model does provide a reasonable account of subjects' performance on sentence verification tasks. However, the theory is rather complicated and this, along with other considerations, has resulted in an alternative "feature" theory of semantic memory. Smith, Shoben, and Rips (1974) suggest that the meaning of a word (e.g. a bird) is represented by a set of features that define its various attributes (e.g. has wings, flies, lays eggs). The feature theory provides a good account of typicality effects by arguing that typical exemplars are identified more quickly because they possess more overlapping attributes than atypical exemplars. However, the theory is somewhat constrained in that it has evolved principally to explain how subjects verify that items are or are not members of a

given category. Clearly, semantic memory allows us to achieve much more than that but feature theory offers little account of this. It cannot, for example, explain how we verify a sentence such as "the woman has a watch," in which the subject and the object do not have any overlapping features. Furthermore, the theory suffers because defining features for a given concept cannot always be specified.

An alternative view of representation in semantic memory is provided by prototype theory. Within this framework a concept is represented not by defining features, but by a set of characteristics which are considered, in differing degrees of importance, to be representative of that concept. In the case of "bird," for example, "feathers" and "wings" are highly representative features whereas "seven feet tall" and "flightless" are not. As a result "parrot" is close to the prototype of bird whereas "ostrich" is not.

The prototype theory accounts nicely for typicality effects in sentence verification as well as a number of other findings. When asked to generate items from a category people are far more likely to produce exemplars that are close to the prototype than exemplars that are not. Research on color categories (e.g. Rosch, 1975) has shown that, despite immense variation in the number and nature of color boundaries in different cultures, people are reasonably consistent in what they consider to be the best example of a focal color (e.g. red). Prototype theory is also supported by various linguistic conventions. People say, for example, that "technically speaking" a tomato is a fruit but that most people think of it as a vegetable. This suggests that people have some awareness of prototypical knowledge. The existence of prototypes can also be demonstrated in the learning of novel material. One study, for example, has shown that people falsely identify previously unexposed schematic faces that are prototypes of faces that were in the learning sequence.

However, like other theories of semantic memory, prototype theory also has its inadequacies. Although the notion of a prototype is easily understood with simple categories such as "bird", prototypes cannot be so easily identified when other sorts of concept (e.g. a belief) are considered. Prototype theory has also some difficulty explaining why people attach greater salience to some features than to others when forming a new concept. Despite these and other objections the available evidence indicates that prototypical information

is a salient aspect of representation in semantic memory.

Psychologists have also attempted to understand how knowledge more complex than simple concepts is represented in semantic memory. Relational concepts (e.g. hit) have received some investigation and the evidence suggests that, like simple object concepts, their representation is prototypical rather than one based on defining features. However, most research into more complex knowledge has concentrated on the role of schemata(singular schema) and the derivative ideas of frames and scripts.

The existence of schemata was first proposed by the neurophysiologist Head but introduced to psychology by Bartlett (Sir Frederic). Bartlett argued that, through experience, people build up a set of expectancies about what should and should not happen in any particular circumstances and it was this that he termed a schema. Thus, when remembering or interpreting information people use schema as a means of filling in missing facts in a way that is consistent with their view of the world.

The idea of internal schema was revived by a number of psychologists. Schank and Abelson (1977) proposed that people have scripts which we use in our comprehension of everyday events such as going to a restaurant. Studies have given plausibility to the idea of scripts. When describing the typical events in a restaurant, for example, people show a high degree of consistency in the component acts they describe. Script theory is also supported by experiments showing that information consistent with the script for a sequence of events is remembered differently to events that are not normally associated with that script (Mandler, 1984).

It has also been argued that people impose sets of expectancies during perceptual processing and it is these that are most often called frames (Yekovich & Thorndyke, 1981). The basic idea is that, by expecting certain features to be present in a particular scene, the amount of processing effort for this information can be reduced and thus more applied to the un-expected and more difficult aspects of the percept. The existence of frames has been supported by several lines of evidence, including the finding that people are far more likely to notice changes in components of a visual scene that are unexpected compared with components which are consistent with the frame.

Schema theory has a lot of intuitive plausibility but there are a

number of problems. First, there seems to be no experimentally proven theory as to how schemata arise. Also, it much of perceptual analysis involves the use of frames derived from expectancies we ought to experience more difficulty than we do in comprehending wholly unexpected events. If, without warning, an elephant were to crash through the ceiling during a seminar we would have no difficulty realizing what had happened. This rather extreme example illustrates that although scripts and frames may facilitate our mental activity, we must have other more fundamental mechanisms which enable us to deal with experiences independently of our expectancies. For this reason the explanatory power of scripts and frames may be far less than some have claimed.

Sentence Processing

A great deal of research in psycholinguistics has been devoted to uncovering the very complex processes that are involved when we understand and produce sentences. One area of particular study has been the unit of analysis involved in sentence comprehension and production. The question at issue is this: do we process language in sentential units or is the unit of analysis smaller than a sentence?

There are good reasons for supposing that, for spoken language, the sentence is a rather artificial unit. Indeed, it has been argued that sentences only exist in written text where their boundaries are delineated by capital letters and full stops. Spoken language typically consists of incomplete sentences as any attempt to provide a written transcription of spontaneous speech will readily show. For this reason, it is probably more appropriate to think of something more like a clause as the main unit in both language production and comprehension.

Research on language production has shown that there are many differences between written and spoken forms. Most probably, spoken language usually contains frequent pauses and fillers (like "ah" and "um") as well as various kinds of error. The pattern of pauses produced by speakers provides important insights into the processes that are involved in language production. Beattie (1983) analyzed pauses occurring in spontaneous speech and found that over half appeared immediately before the start of a clause. This finding may be taken as evidence that the clause is an important unit in spontaneous speech. However, Beatties's data also suggest that when people are talking,

they do not plan one clause at a time but instead plan several clauses together.

Beattie argues that spontaneous speech has a cyclical character in which hesitant and fluent phases follow one after the other. During a hesitant phase – which is characterized by a great deal of pausing and the use of short causes – the speaker is planning ahead, not only to the next clause, but also to the next sequence of clauses. Once this forward planning has been completed, the speaker then enters a fluent phase during which there is significantly less hesitation and the speaker tends to use longer clauses. Once the limits of forward planning have been reached, the speaker then enters another hesitant phase. In Beattie's data, a complete cycle – a hesitant phase followed by a fluent phase – lasted for about nine clauses.

Planning what to say involves several different activities including the choice of topic, the organization of ideas within that topic, and the selection of appropriate words and syntactic structures to convey those ideas. The listener's task is, in some sense, the reverse of the speaker's because the listener has to derive intended meaning from what a speaker says. An important part of this task is concerned with integrating information contained in successive phrases and clauses and making necessary inferences. But before such higher level processes of comprehension can come into operation, essential processes of syntactic and semantic analysis are necessary.

It has been proposed – in the "clausal hypothesis" (Fodor, Bever, & Garrat, 1974) – that the clause is the main unit of analysis for the earlier processes of comprehension. In its strongest form, the clausal hypothesis makes two claims. The first is that syntactic and semantic processing does not begin until a whole clause has been heard or read. The second is that, immediately a clause has been processed, it is expunged from working memory and only its meaning is retained in long-term memory; details of syntactic structure and the particular words used by the speaker are lost.

Several experiments have shown that, in general, people do rapidly forget information about linguistic form. For example, subjects in a study by Sachs (1967) listened to a passage and were then presented with a single sentence (the test sentence) and asked whether it had occurred in the passage. Sometimes the test sentence was identical to one that had occurred in the passage (the target sentence),

but on other trials, the test sentence was slightly different, either in meaning or in form, from a sentence that subjects had actually heard.

Sachs systematically varied the position of the target sentence in the passage and she found that when the target sentence was the most recent one that subjects had heard, they could accurately recognize changes in either meaning or form. However, when several sentences had intervened between the target sentence and the test sentence, subjects could reliably recognize only changes in meaning.

Sachs's experiment strongly suggests that, with the exception of the most recently heard sentence, only the meaning – and not the form – of sentences is retained in memory. However, while the general conclusion of Sachs's experiment is undoubtedly correct, it is by no means clear that forgetting occurs on a simple sentence-by-sentence or, even, on a simple clause-by-clause basis. Indeed, an elegant experiment by Jarvella (1971) shows that the extent to which lexical and syntactic details of a clause are lost depends on its structural relationship to other clauses.

Subjects in Jarvella's experiment listened to prose passages. At various points their listening was interrupted and they were asked to recall what they had just heard. There were two versions of each passage and the main focus of interest was subjects' ability to recall the three clauses that were presented immediately before recall. In one condition the most recently heard clause was structurally independent of the immediately preceding clause. In the other condition the same clause was structurally related to the immediately preceding clause. An example of Jarvella's materials is shown below (clause 3 was the one heard immediately before recall):

Structurally independent target clause	Structurally dependent target clause
1. The document also blamed him	1. The tone of the document was
2. for having failed to disprove the charges.	2. Having failed to disprove the charges
3. *Taylor was later fired by the president.*	3. *Taylor was later fired by the president.*

Jarvella found that, as expected, subjects in both conditions

were very accurate at recalling the clause they had just heard. However, recall of the immediately preceding clause (clause 2) was significantly different in the two conditions. Where clause 2 was structurally dependent on the most recently heard clause, subjects could recall about 80 percent of clause 2, whereas when clause 2 was structurally independent of the last heard clause, recall dropped to 50 percent.

These data suggest that, in the case where a preceding clause is structurally related to the clause that follows it, the preceding clause is remembered almost verbatim until the following clause has been fully analyzed. It is easy to see why this is the case. Imagine that you came across the clauses that Jarvella used. If you heard "Having failed to disprove the charges," you would realize that there was more information to come even though you could carry out the majority of the syntactic analysis of that clause. In particular, you would know that "someone" had failed to disprove some charges but you would not know that person's identity. It would only be on hearing the next clause that this information would become available, and it would only be on determining this that you could fully complete the analysis of the earlier clause. Once the analysis of clause 2 was complete – and only then – would it be appropriate to expunge verbatim information about it from Working Memory.

There is another way in which some modification is required of the claim that verbatim information about clauses is immediately lost once they have been processed. This arises from findings of Kintsch and Bates about long-term recall of speech in real life (Kintsch & Bates, 1977, Bates, Kintsch, Fletcher, & Giuliani, 1980). Kintsch and Bates tested students on their memory for lectures. The students had not been warned in advance that they were to be tested but, somewhat surprisingly, 48 hours after the lecture students were able to remember quite a lot of information about the form, as well as the meaning, of what they had heard. On a recognition memory test, they were able to make reliable distinctions between sentences that had been in the lecture and paraphrases that meant the same.

One reason why the performance of the students in the Kintsch and Bates study was so much better than that of other experiments that have tested memory for language form – such as Sachs (1967) and Garrod and Trabasso (1973) – lies in the complex relationship between the form and meaning of language. The way in which we

choose to say something is usually not arbitrary. Important pragmatic factors influence the selection of a particular linguistic form to convey a particular piece of information. For example, the choice of a passive rather than an active form – that is, the choice of "Several candidates were chosen by the committee" rather than "The committee chose several candidates" – is strongly influenced, among other things, by the viewpoint of a particular discourse.

Such pragmatic factors were undoubtedly at work in the natural discourse that Kintsch and Bates studied, and these may explain why the form of some parts of the lectures that the students heard was particularly memorable. However, pragmatic factors influencing the choice of linguistic forms are usually ignored in the construction of stimuli for experiments that look at memory for language and hence there is no reason why the form of sentences in such experiments should be remembered.

Another aspect of the clausal hypothesis has also proved to be open to doubt. This concerns its first claim, that speech processing does not begin until a listener has heard an entire clause. A series of experiments by Marslen-Wilson and his co-workers (Marslen-Wilson & Tyler, 1980; Marslen-Wilson, 1987) has demonstrated that processing of syntactic, semantic, and pragmatic information is carried out as soon as the first word in a clause has been heard. Subjects in Marslen-Wilson's experiments performed a variety of tasks in which they had to identify a particular word in a sentence to which they were listening. Target words occurred in different serial positions across the test sentences, varying from the second to the tenth position. The syntactic, semantic, and pragmatic plausibility of target words was also varied by using three different types of prose. The first of these, normal prose, consisted of sentences such as "Some thieves stole most of the lead off the roof," which were pragmatically, schematically, and syntactically appropriate, whereas the second condition, anomalous prose, consisted of sentences such as "No buns puzzle some in the lead off the text," which obeyed the rules of English syntax but were meaningless. The third condition, scrambled prose, consisted of strings of randomly ordered words that obeyed neither the rules of syntax nor semactic-pragmatic constraints.

If processing of the syntax, semantics, and pragmatics of language does not begin until an entire clause has been heard, then the

plausibility of a word should not have any influence on the speed with which that word can be recognized unless that word occurs right at the end of a clause. That is, the time taken to identify a word should not be affected either by its position in a sentence or by the prose condition in which it occurs. (The argument here is that syntactic, semantic, and pragmatic plausibility within a sentence can only operate once information about these various linguistics aspects has been processed).

Marslen-Wilson's data clearly demonstrate that, in the normal and anomalous prose conditions, the time taken to recognize a word decreased with its position in the sentence – that is, the later in a sentence a target word appeared, the faster subjects were to identify it. Words occurring right at the end of a sentence were identified faster than words occurring at an earlier position, but there was a steady decrease in recognition time *throughout the entire sentence.* However, for the scrambled prose condition – where words occurred in random order – recognition time was the same for all word positions.

This result is only explicable if we assume that subjects were processing the sentence as they heard each word and were thus able to exploit syntactic, semantic, and pragmatic constraints that built up throughout the sentence. Subjects did not delay processing until they had heard an entire clause.

Recognition time for words in the normal prose condition was also influenced by whether there was a context sentence preceding the one containing the target word. Words occurring early in a sentence were recognized faster if there was a preceding context sentence. This is because information from a previous sentence can also influence the plausibility of a word and so make it easier to recognize.

The fact that recognition of a word is influenced by information from previous sentences, as well as that contained in earlier parts of the same sentence, suggests that processing is not only carried out as we hear each new word, but also that information from new sentences is rapidly integrated with information that has been extracted from earlier ones.

Short-Term Memory

Short-term memory refers to memory for events which occurred within the last few seconds; theoretically, it has been argued that

there is an important distinction between short-term memory on the one hand and Long-Term Memory on the other. The essence of the theoretical distinction was embodied in James's terms "primary memory" and "secondary memory". According to James (1890), primary memory is concerned with information that remains in consciousness after it has been processed, and thus forms part of the psychological present. In contrast, secondary memory is concerned with information about events that have left consciousness, and as a consequence form part of the psychological past.

One of the major differences between the short-term memory store (or primary memory) and the long-term memory store (or secondary memory) is in terms of their capacity. There are no known limits on the capacity of long-term memory, but it was argued by Miller (1956) in a very influential article that no more than approximately seven chunks (i.e. familiar units of information) can be stored in short-term memory at any given time.

The notion that there are separate short-term and long-term memory stores was incorporated into a number of major theories of memory (e.g. Atkinson & Shiffrin, 1968). Some of the strongest support for this notion came from research into amnesic patients. In essence, it was discovered that most amnesic patients have severely impaired long-term memory but largely intact short-term memory. However, there are a few amnesic patients who exhibit the opposite pattern of deficient short-term memory but essentially intact long-term memory. It is difficult to account for these findings without assuming that there is a valid distinction between short-term and long-term memory systems.

In spite of the empirical support for the distinction between short-term and long-term memory stores, it has increasingly been realized that neither the short-term store nor the long-term store is unitary. For example, Baddeley and Hitch (1974) and Baddeley (1986) argued that the concept of a unitary short-term store postulated by Atkinson and Shiffrin (1986) should be replaced by that of a working memory system consisting of three components. This conceptualization has the advantage of providing a more adequate account of the complexities of short-term memory. It also has the advantage that the working-memory system, combining as it does active processing and transient storage of information, provides an

explanation of performance on a wide range of cognitive tasks which are not explicitly concerned with short-term memory.

Short-Term Memory, Retrieval from

Information is Shor-Term Memory or Working Memory is information which is currently being processed. As James (William) (1890) pointed out, information in short-term memory forms part of the psychological present, whereas information in long-term memory forms part of the psychological past. In light of these considerations, there has been little interest in assessing the probability of retrieval from some short-term memory store, since it has been assumed that all information actually in that store can be retrieved. As a consequence, research and theory on retrieval from short-term memory have focused on the speed with which information can be retrieved from short-term memory rather than on the accuracy of retrieval.

It is a matter of some theoretical importance to establish as precisely as possible the characteristics of short-term memory. The short-term memory or working memory system is obviously involved in many memory tasks, but its significance extends far beyond that. It appears to be the case that this system is used in the performance of most cognitively demanding tasks, so that understanding the detailed functioning of the short-term memory system may increase our knowledge of the processes involved in comprehension, Problem Solving, mental arithmetic, and so on.

The basic paradigm for studying speed of retrieval from short-term memory was introduced into psychology by Sternberg (1960). In this paradigm, between one and six items are presented to the subject, and these items are generally referred to as the "memory set." A probe item is presented very shortly after the memory set, and the subject's task is to decide as rapidly as possible whether or not it matches one of the items in the memory set.

Since relatively few errors are usually made on this task, interest centres on the speed of positive and negative responses as a function of the number of items in the memory set. While the pattern of results is affected by various factors, the most common finding is that the positive and negative functions are both linear and parallel. According to Sternberg (1969), the results are consistent with a serial exhaustive model in which the probe is compared against each of the items in the

memory set at a rate of 25-30 items per second. It might seem more likely that there would be a self-terminating scan, in which the search would stop as soon as a match for the problem was located. With a self-terminating scan, half of the memory-set items would be searched on positive trials, whereas all of the items would be searched on negative trials. The natural prediction on a self-terminating scan view is that increasing the number of items in the memory set would increase the response time to the probe at approximately twice the rate on negative trails as on positive trials, but this does not happen.

Sternberg's (1960) complete theory consists of four separate processing stages, which together account for the time taken to respond to the probe. We have so far considered the serial comparison stage, which is the second of these four stages. The first stage is that of probe encoding, the third stage is that of decision (i.e. whether to respond "yes" or "no"), and the fourth stage is that of translation and response organization. According to Sternberg, manipulations of memory-set size affect only the serial comparison stage and have no effects on the remaining stages.

Although Sternberg's model provided an accurate and elTegant account of the basic findings from his paradigm, it became clear that there were various findings which it could not readily accommodate. For example, since the search through the memory set is allegedly exhaustive, it should not make any difference to the response speed whether the probe matches with the first item, the last item, or one of the middle items of the memory set. In practice, however, there is frequently a recency effect, in which the probe is responded to fastest if it matches with the last item or two in the memory set.

An alternative view of performance on the Sternberg paradigm is that the speed of performance is determined almost entirely by the strength or familiarity of the probe. According to a strength theory, the reason why subjects take longer to respond when there are more items in the memory set is that under such circumstances the average familiarity of the memory-set items declines. The recency effect has been accounted for by assuming that the last item in the memory set is in some sense stronger than the other items.

Atkinson and Juola (1974) proposed an interesting theory which combined elements of the Sternberg and strength theory approaches. According to them, subjects respond rapidly to the probe on the basis

of its strength or familiarity if its strength is very high or very low. If its strength is intermediate, then there is a slower search through the memory-set items. This search may or may not be the serial exhaustive scan proposed by Sternberg.

One of the advantages of the Atkinson and Juola model is that it provides a reasonable explanation of speed-accuracy trade-off (i.e. subjects can choose to respond rapidly but at the expense of making many errors, or slowly but accurately). If subjects attempt to respond rapidly, then this can be done by basing nearly all of their decisions on familiarity or strength alone. If they want to be accurate, then they can nearly always make use of the search process. More detailed predictions from the model have been confirmed in various studies (e.g. Banks & Atkinson, 1974).

11

Social Cognition

The study of social knowledge (its structure and content) and cognitive processes (including acquisition, representation, and retrieval of information) provides a key to understanding social behavior and its mediating factors. Social cognition leans heavily on the theory and methodology of cognitive psychology to provide precise and detailed models of social information processing. In particular, memory processes (encoding, storage, and retrieval) are seen to play a significant role in mediating a range of social judgements.

Although the popularity of the cognitive approach to social psychology has grown in the last ten years, it is not new. Indeed, it owes a great deal to the input of Gestalt ideas into American social psychology from Europeans such as Kurt Lewin and Fritz Heider. The vast majority of social-psychological data are about thoughts – as judgements, opinions, attitudes, or attributions. Furthermore, cognition pervades social psychology at three levels – the level at which the problem is formulated, the level of methodology, and the level of theorizing. Arguably, social cognition is more complex than cognition in general, as evidenced by several general features. First, social perceivers typically go beyond the information given. Second, the objects of social cognition can be changed by being its focus, because it has consequences for them (thus, social psychologists talk of self-monitoring, objective self-awareness, and evaluation apprehension). Third, nearly all perception and cognition is evaluative; a consequence of this affective involvement is increased encoding

idiosyncrasy, the between-subjects variance that social psychologists seek to explain.

Three fundamental questions underlie the study of social cognition: first, what type of social information is stored and how is it organized in memory?; second, how does social information stored in memory affect subsequent information processing, decision making, and behavior?; and third, how and when is stored information changed, both by new information and by cognitive processes? (Sherman, Judd, & Park, 1980).

Conceptually, the metatheory driving the social cognition approach is that of the "cognitive miser." Its central idea is that people are seen as capacity-limited information processors, who can deal with only a small amount of information at any time. Given these limitations, people use short cuts and strategies (heuristics) to simplify complex problems of judgement, decision, and attribution. These heuristics produce fast and quiet adequate, rather than slow, normatively correct, solutions, as illustrated by three heuristics governing intuitive prediction and judgement. According to *the representativeness heuristic,* an object is assigned to a conceptual category by virtue of the extent to which its main features represent or resemble one category more than another. *The anchoring/adjustment heuristic* refers to people's general failure to adjust their initial judgements in the light of subsequent evidence. *The availability huristic* refers to the general tendency to judge events as frequent, probable, or causally efficacious to the extent that they are readily available in memory. This approach provides a general understanding of human judgement, although it remains difficult to evaluate the different heuristics competitively and rigorously.

Social Categorization

The process of social categorization is another example of an information-processing short-cut. It involves matching a target person to an existing social category, and it is suggested that the process is set in motion as soon as information sufficient to activate a relevant social category has been encountered. The information necessary to cue an appropriate category can take several forms. For example it may take the form of an observable feature (e.g. skin color or clothing). Alternatively, and more common in psychological research, it may take the form of a written category label or a cluster of category-

consistent attributes. Rather than try to process all the stimulus information available, the social perceiver simplifies matters, by assigning individuals to social categories on the basis of their perceived similarities.

To the extent that social cetegories capture and reflect real differences and similarities between people, then social categorization can be considered a valuable cognitive tool. A growing number of studies have, however, demonstrated biases which stem from the very process of social categorization itself. For example, once a person has been classified into a particular social group, it is assumed that he or she possesses all the characteristics that define the group as a whole, and shares these characteristies with the other group members. That is, the person is deemed to possess a range of category-relevant attributes, despite the fact that none of these attributes may have served as the basis for the initial categorization. Furthermore, once we have classified two people into different categories, we tend to ignore their similarities and exaggerate their differences. Conversely, once we have assigned two people to the same category, we exaggerate their similarities and ignore their differences. These effects are accentuated if the social perceiver belongs to one of the groups in question (i.e. ingroup or outgroup). In general, the perceiver shows greater differentiation between ingroup members, with representations of ingroup members generally more complex than those of outgroup members. Consequently, ingroups and outgroups reflect an asymmetrical relationship, with perceived ingroup heterogeneity being contrasted with outgroup homogeneity (Tajfel, 1981; Fiske & Taylor, 1984).

Schemata

The process of social categorization is central to both theorizing and experimentation within social cognition. This is because once a social category has been activated it plays an important part in subsequent information processing. The activated category affects the encoding, representation, and retrieval of social information. A central tenet of social cognition is that people assimilate what they observe to pre-existing cognitive structures or Schemata. These abstract knowledge structures, stored in memory, specify the defining features and relevant attributes of a stimulus domain, and the

interrelations among those attributes. Five main types of schemata have been identified (person, self, role, event, and procedural), each of which can influence three main types of social information processing: perception of new information; memory for old information; and, especially, inferences that go beyond both.

The impact of schemata on social information processing can be interpreted in terms of both functions and liabilities. They tend to aid memory for schema-consistent information, increase confidence in schema-consistent recognition, enhance memory for schema-consistent information that was, in fact, never presented, and guide inferences and predictions. Further, it appears that under certain conditions people will tend to remember different types of information. For example, when forming an impression of someone subjects tend to show preferential recall of schema-inconsistent information. This is consistent with a range so studies that suggest that the presentation of relatively infrequent, novel, or salient behaviors will result in increased attention during impression formation and, as a consequence, be more accessible in memory. Schema-inconsistent information also tends to be preferentially recalled when stimulus targets are perceived as being members of groups. In contrast, consistent information is more readily recalled when targets are perceived as individuals, and when established beliefs or impressions are being tested.

The social cognition perspective has been usefully applied to the study of stereotypes, attitudes, and the self, via the concept of schemata. Stereotypes, as social categories, may be formed and represented in memory much like any other object categories. Thus, they may be organized around a prototypical representation, with exemplar or specific instance information stored as well. Social cognition research is currently evaluating prototypic and exemplar-based representational models in this domain. Stereotypes have been equated with schemata and, as such, have been shown to influence all stages of social information processing. This approach has been particularly useful in clarifying how stereotypes are formed, how they are maintained, and also how they might be changed.

A range of cognitive biases are implicated in the formation of stereotypical conceptions. These include differential attention to salient stimuli, and illusory correlation. The causes of social salience all depend upon contextual factors. Fiske and Taylor (1984) state that a person

can be salient relative to the perceiver's: (a) immediate context; (b) prior knowledge or expectations; and (c) attentional tasks. Regardless of the cause of social salience, its effects are robust. Salience or prominence accentuates a range of stimulus-based judgements. Essentially, evaluations and judgements are exaggerated in whichever direction they initially tend. For example, if a person is viewed negatively, being socially salient will increase this negative evaluation. Similarly, an initially positive evaluation will be accentuated if the person becomes salient. Illusory correlation concerns the imposition of a relationship between two variables when none actually exists. The main factor claimed to be responsible for the phenomenon is paired distinctiveness. This basis of illusory correlation has been used to account for the negative stereotyping of minority group members. Majority group members rarely interact with minority group members, and in comparison with positive behaviors, negative behaviors are relatively infrequent. It is therefore suggested that majority group members perceive an illusory correlation between the two and overestimate the frequency of occurrence of negative behaviors by minority group members.

At the stereotype-maintenance stage, researchers have demonstrated persistent biases in patterns of causal attribution and in memory retrieval which allow the perceiver's erroneous stereotypical beliefs to persevere in the face of potentially disconfirming information. Specifically, perceivers tend to attribute stereotype-confirming behavior to internal stable causes, and to attribute disconfirming behavior to external, unstable causes. Such an attributional tendency allows the perceiver to explain away stereotype-discrepant behavior. Similarly, the preferential recall of stereotype-confirming information, together with the tendency to select only confirming information when testing stereotypical beliefs, further perpetuates social stereotypes. Finally, schematic models of how stereotypes might change in response to the amount and distribution of disconfirming information have been tested. Research has shown that it is better to disperse disconfirming information across several outgroup members (each of whom displays some goodness of fit with the category prototype), rather than to concentrate disconfirming information in a few, highly atypical members. Interest here centers on when subcategories may be formed as a result of encountering new group exemplars that are divergent

from the category prototype or central tendency. Thus, a social cognition approach to "intergroup contact" emphasizes that successful contact (i.e. positive contact that generalizes to other outgroup members) depends upon the prototypicality of the outgroup member (category exemplar) with whom contact takes place.

Attitudes also appear to function like schemata, as illustrated by selective attention to, and encoding, retention, and retrieval of, attitude-relevant information. For example, attitudes facilitate recall of attitude statements that were strongly agreed or disagreed with, compared with statements that elicit more moderate responses on the agreement scale. Thus, the social cognition perspective has led to a major shift in the way attitudes are defined and assessed. In many situations evaluations seem to be formed spontaneously and stored in long-term memory for future retrieval and use. Other information stored with the evaluation may include information about the attributes of an attitude object, affective responses elicited by the object, past behavior toward the object, and the evaluations of significant others. The cognitive approach has also stimulated interest in the manipulation and measurement of cognitive responses to persuasive communications (Petty & Cacioppo, 1986). Persuasion via a central route requires active processing of the message, which itself requires cognitive effort (as when we are highly involved in an issue). Persuasion via a peripheral route involves minimal cognitive elaboration of the message, and is due to non-message factors (e.g. high credibility of the source). The amount and quality of cognitive responding can be measured by a written thought-listing technique, which has also been used to specify more precisely the cognitive processes involved in social influence by majorities and minorities.

There is now theoretical and empirical support for a multifaceted and complex concept of the self as a cognitive structure. The self is conceived as a special kind of schema, or a system of schemata, owing to its size, complexity, and affect. Each schema is a generalization about what the self is like and contains trait information, behavioral information, and inferences. The self is involved in attention to, and encoding and intepretation of, new information, but a person may be "schematic" or "aschematic" with respect to certain roles or traits. There is a general finding, however, that material is better recalled when encoded with

reference to the self (the self-reference effect).

Methodology

As social psychologies have studied the metaphor of information processing, so they have used a wider array of cognitive measures: visual attention, subject self-reports, requests for information, recall (quantity, errors, sequence, clustering), recognition, and chronometric analyses. These measures all attempt to sidestep a major methodological problem – the fact that we can never tap directly what is going on in the heads of our research subjects. Examples of problems studied include encoding vs. retrieval biases in memory for stereotypical information about group members, and negative moodpriming in relation to antisocial behavior.

More ambitious still are attempts to use these different measures to build process models. A process model is, simply, the description of everything that goes on in the subject's head, from start to finish of an experimental task (cf. think-aloud verbalizations – Introspection). It is a statement of the presumed stages through which information is processed, such as encoding, storage, retrieval, recall, and attribution. The aim of process analysis is to provide methodological precision, to specify the stages in social information processing and at what stage a given effect occurs. In the area of causal attribution, process models have been constructed to investigate several issues. Salience-based attribution, for example, was examined by using structural modeling techniques to relate measures of visual salience, attention, recall, and attribution. Process models have also been used to explore whether causal attribution involves Automatic Processing. Different response times for measures of causal judgement, trait judgements about an actor, and the actor's intent indicated that causal judgement was not an automatic process, and indeed that, under certain circumstances, it was mediated by more basic triat judgements and intention judgements.

Theories about process are potentially more general than theories about content, because the same procedure may operate over a wide range of stimuli, and they can provide methodological precision. By definition, however, the attempt to measure cognitive processes interferes with normal thinking and may not be informative about normal, extra-laboratory social information processing.

Affect and Motivation

Because social cognition is rarely dispassionate, or emotionally neutral, it is important to integrate social cognition with affect and motivation. Feeling states, negative and positive, influence what people think about and the judgements they make, and thus have important effects on social behavior. Social cognition is still central, however, because the impact of feeling states can be best understood as a function of cognitive processes such as perception memory, and inference. Thus, people in a positive feeling state are more likely to retrieve positive material from memory than people who are not (Isen, 1987). The affective state seems to function as a cue to prime and organize related cognitive material, via both automatic and controlled processes.

It has also been suggested that extremity in mood and evaluation is tied to the complexity of knowledge structures. More simple thinking about a domain (e.g. an outgroup about whom one has little knowledge) is associated with more extreme affective reactions within that domain (e.g. liking or disliking); greater complexity is associated with more moderate reactions. Research has also suggested that affect is stored with schemata, and is available immediately upon categorization (fitting an instance to a schema).

Social interaction and information processing about other people is often motivated by factors such as current concerns, values, and beliefs. In particular, social cognition is often goal directed, whether in terms of personality analysis, self-presentation, or information seeking. This motivational side of social cognition has been rather overlooked. The goals individuals bring to social situations may influence their perception of and memory for others' behavior; thus, subjects are more likely to organize information around individuals in memory, if they anticipate future interaction.

Conclusions

Social cognition represents an approach or set of assumptions guiding research in social psychology, rather than a separate theory or domain of enquiry within the discipline. The fruits of this approach to a variety of traditional substantive domains can be seen in relation to causal attribution, attitudes, and attitude change, intergroup relations (especially stereotyping), judgement and decision making, and the self.

Speech Perception

Research on speech perception attempts to understand how the brain can recognize in speech discrete linguistic units, such as words or phonemes. Traditionally the field has concentrated on the sound of speech, but recent research has indicated the contribution of vision. The classical work in the field pointed out the complex relationship between acoustic cues and phonetic categories. Current work compares the contributions of auditory and specialized phonetic mechanisms to phonetic perception, explores the development of speech perception in infants, and tackles the problem of extracting words from the speech stream.

The Natural of the Speech Signal

The unusually complex nature of speech sounds was revealed in the 1940s by the recently invented spectrograph. The spectrograph revealed two basic properties of the speech signal: first, that speech was a continuously changing pattern of sound, rather than a sequence of discrete elements; and second, that the pattern in the region of a particular linguistic element (such as a phoneme) varied with context. These two properties reflect two aspects of the way in which even careful speech is produced. First, the articulators must move from one target position to another, so the sound that emerges also makes transitions. Second, the specification of consonants generally involves only a subset of the articulators (e.g. closing the lips for a bilabial stop), leaving the other articulators free to anticipate subsequent segments (e.g. the tongue body is free in the bilabial stop to take up the position appropriate for the following vowel). This phenomenon is known as co-articulation.

Normal, rapid, casual speech introduces further complications since the target positions of segments may be undershot or the segments themselves may disappear as a result of the problems of moving the articulators rapidly enough (e.g. "What do you want?" becomes "Wotchawan?") Speakers adjust the clarity of their speech depending on the listening conditions and the context available to help the listener (Lieberman, 1963).

Yet more variability in the relationship between the linguistic segments and the sound wave is introduced by the fact that different speakers have different-sized heads and speak different dialects and

idiolects (Ladefoged & Broadbent, 1957). Finally, the environment through which sound passes from speaker to listener adulterates sound, and adds in sound from other speakers and sound sources.

As an antidote to this emphasis on variability, Stevens (1972) has pointed out that the articulation of some consonants uses places of articulation that give a stable acoustic output despite small perturbatons in the place of articulation. His emphasis on invarient acoustic cues to phonetic categories has been taken up by Blumstein and her colleagues (e.g. Blumstein, Issacs, & Mertus, 1982).

Fundamental research on the acoustic cues to phonetic categories has been conducted (mainly at Haskins Laboratories) since the early 1950s. It has emphasized the complex relationship between acoustic cues and phonetic categories and pointed out a number of perceptual phenomena (categorical perception, left-hemisphere advantage) which appear to be confined to speech perception. Mattingly and Liberman (1988) later proposed that speech is perceived by a specialized module that is responsible for phonetic, "speech-mode" percepts. Perceiving sound as speech requires knowledge of the particular constraints that apply to sounds produce by the human vocal tract. These relationships do not apply to the general sounds of the environment. Although few workers would disagree that speech perception must require some separate mechanisms from general auditory perception, the need for a module, and the particular perceptual phenomena which demonstrate special processes are both contentious.

Our ability rapidly to label the continuous flow of speech sound with discrete linguistic categories lies at the heart of speech perception. Research has concentrated on two aspects of this labelling: where listeners put the boundary between one phoneme category and another, and our apparent inability to distinguish sounds within a particular category (categorical perception).

Phoneme Boundaries and Trading Relations

There has been considerable debate over the question of whether the position of phoneme boundaries is determined by general auditory constraints, or by the operation of special speech perceptual mechanisms. For the voice-onest time (VOT) continuum which cues the voicing distinction between say, "ba" and "pa," there is clear evidence that chinchillas trained to respond differently to the two

ends of the continuum show a behavioral boundary at the same point as do human listeners (Kuhl & Miller, 1978). The boundary occurs at around 30 millisec VOT. There is a good reason why the boundary cannot occur at a shorter value than this, namely that the (chinchilla) auditory system in unable to represent accurately the difference between VOTs from this continuum shorter than 30 millisec. Such purely auditory limitations obviously constrain the position of phonetic boundaries: different phonetic categories must be both easily heard. But other phenomena concerning the position of phoneme boundaries cannot be attributed to general auditory mechanisms. One such class of phenomena involves "trading relations" between the acoustic cues to phonetic categories.

In general phonetic categories can be cued by a variety of different acoustic events. Thus, for instance, "say" can be turned into "stay" by introducing silence after the/s/ and lowering the subsequent starting frequency of the first formant transition (both of these cues are consequences of closing the vocal tract for the /t/). Listeners show both a sharp and consistent boundary between "say" and "stay." They are also sensitive to changes in both of these cues, trading off a change in one cue against a change in the other. Such trading relations can be used to demonstrate perception in the speech mode. The demonstration uses an intriguing stimulus called "sine-wave" speech, in which sine waves track the formant frequencies. Naive listeners can hear these sounds as either nonspeech whistles or as distorted speech, depending on their perceptual set. Listeners who hear sine-wave analogs of "say" and "stay" which differ both in silence duration and first formant onset frequency treat the two cues quite differently than listeners who hear the same sounds as speech. The particular trading relation between the two cues found for more natural sounds is only apparent for listeners that hear the sounds as speech. It is impossible to attribute this difference to an underlying psychoacoustic or physiological cause.

Categorical Perception

For sounds along simple physical continua (frequency, intensity, etc.), we are better at discriminating pairs of sounds than we are at labeling them as distinct categories. This principle appeared to be dramatically violated by the "categorical perception" of certain speech

sounds. Here, for some complex auditory continua that cue simple phonetic distinctions (e.g. voicing – between "ba" and "pa," or place of articulation – between "ba," and "da," and "ga"), listeners can discriminate much better between pairs of sounds that fall into different phonetic categories than they can between pairs that do not. In extreme cases, the ability to discriminate along the continuum is predictable from the ability to assign different phonetic labels, hence perception is "categorical." This striking phenomenon provided much of the impetus for special theories of speech perception, and continues to attract considerable attention.

For some continua (e.g. VOT) the peak in discrimination at the phoneme boundary does appear to be due to auditory limitations. Kuhl's (1981) chinclillas show peaks in similar places to human listeners. Moreover, the peaks persist when discrimination is tested using a fixed rather than a roving standard, which is a technique that reduces the contribution that perceptual categories make to discrimination performance. Discrimination measured with a roving standard (the normal ABX method) can show peaks near category boundaries which disappear when the more stringent fixed standard method is used. So some of the peaks in discrimination functions reported in the literature are attributable to basic auditory mechanisms; other are a consequence of phonetic labeling.

The traditional ABX method of testing speech sound discrimination nevertheless remains a useful indication of a native listener's perceptual categories. Adult listeners of a language that does not make a particular distinction (e.g. between [l] and [r] in Japanese) show very poor ability to discriminate sounds along the appropriate continuum when tested with the conventional ABX procedure. The contrast between this poor ability of adults of making within-category discriminations, and the infant's remarkable abilities to discriminate speech, has led to some of the most interesting recent work on speech perception.

Infant Speech Perception

Eimas, Siqueland, and Jusczyk (1971) introduced the first viable technique for detecting an infant's ability to discriminate sounds. Infants will suck a pacifier to hear a sound but soon habituate if the same sound is repeated. Sucking increases again if the sound is changed

and gives an index of the perceptibility of the change to the infant. Eimas's experiment showed that one- and four-month-old infants more readily discriminated sounds which straddled the /ba/-/pa/ boundary along the VOT continuum than those that did not, a result which with hindsight probably reflects general auditory mechanisms rather than a special speech perception one. Subsequent work has shown that young infants are able to distinguish not only the phonemi contrasts used in the language to which they are exposed, but also many that are not. We know that adult listeners can be insensitive to contrasts that their language does not use (at least when tested by some discrimination procedures), so does the infant gradually lose the ability to hear within category distinctions? Careful and imaginative work by Janet Werkert and her colleagues has shown that around 10 to 12 months infants lose the ability to hear differences between sounds that are not distinguished by their native language. But the same distinctions are preserved in infants and adults exposed to the appropriate languages (Hindi or Nthlakapmx). Werker emphasizes that the loss of the ability to make these discriminations is not a basic psycho-acoustic loss, since adults who study the language can learn to (re-) hear the foreign distinctions and because more sensitive discrimination paradigms also reveal some appropriate ability. Mann (1986), for instance, has shown that although Japanese listeners are not able explicitly to distinguish [1] from [r], they are nevertheless sensitive to the consequences of their different listeners, like English-speaking articulation. Mann showed that Japanese listeners, required different cues after [r] than after [1] to hear /d/ (versus/g/). The Japanese listeners were thus sensitive to the co-articulatory consequences of [1] or [r], even though they could not explicitly hear the difference between the two sounds.

Vision In Speech Perception

Most people are not aware how much speech perception is helped by visual information about articulation, despite the common observation that "I can hear you better with my glasses on." The McGurk effect (McGurk & MacDonald, 1976) encapsulates the bimodal nature of speech perception with a striking illusion. A dubbed video that shows a face executing the movements of/ga/together with the sound of/ba/is "heard" as/da/. The listener (be he or she a trained

phonetician or naive) is not aware of any conflict between sight and sound. The visual information about articulation is smoothly incorporated into the auditory percept by the brain.

Even very young infants appear to be sensitive to the relationship between seen and heard speech articulation since they will look selectively at the face that is articulating a heard vowel rather than a visually different one. Blind infants are slower than sighted ones at developing phonological categories that have visual features.

The important of seen as well as heard articulation for speech perception has emphasized the links between perception and the dynamics of action.

Auditory Word Recognition

The last 15 years have seen an increasing interest by psycholinguists in the perception of spoken (rather than the technically more tractable written) language. Marslen-Wilson's work (e.g. Marslen-Wilson & Welsh, 1978) stimulated research on the temporal course of lexical access from the speech wave, drawing attention both to the interaction with higher constraints and to the fact that under appropriate conditions lexical decisions could be made before the physical end of the relevant word. The role of different processing units in lexical access remains contentious with spectral templates, phonemes, and syllables all having their advocates; equally the mechanism whereby context influences the perception of words continues to be debated. These issues are reviewed by Frauenfelder and Tyler.

Under good listening conditions, a word becomes perceptible at its uniqueness point – the point in the word at which the word first becomes distinct from other words in the language. However, speech is still intelligible under conditions where word onsets may be indistinct, so it is likely that listeners are forced to use subsequent as well as preceding context and to rely on the most accoustically distinct parts of words for their identification. Stressed syllables are generally the most intense and the most carefully articulated, and for English at least may play an important role in the difficult problem of segmenting the speech stream into words.

Speech Production

Approaches to speech production should explain speech planning

and execution, and the explanation for planning should account for novelty in speech. Speech exhibits novelty ("generativity") in two fundamental ways: speakers produce novel, yet informative sentences, and they coin words. They can utter novel, understandable sentences, because sentences are composed of familiar parts – words – ordered and marked according to also-familiar syntactic rules. Speakers coin words by combining familiar morphemes (stems or stems and affixes) or phonemes (consonants and vowels) in new, ruleful ways.

A comprehensive theory of speech production should explain both kinds of novelty, and it should also explain other aspects of planning and execution. Such a theory has yet to be developed, but there are theories of each part.

Planning

Speech utterances must be planned, in part because the sequential order of words in a sentence is linguistic; it is not inherent in aspects of an idea or event being described. Also, the syntax allows creation of nonlocal dependencies (e.g. between underlined subject and predicate in "*the pen* I bought yesterday *has run out of ink*"); production of the first dependent part anticipates production of the second, however separate the parts.

Spontaneous speech errors ("slips of the tongue") provide important evidence concerning speech planning. Dell (1986), Garrett (1980), and Shattuck-Hunagel (1987) offer theories based on slips. Slips of the tongue are fluently produced departures from an intended utterance.

The errors are not, literally, slips of the tongue, but, rather, planning mistakes. A very unlikely series of motor slips would cause "salt" to substitute for "pepper," for example.

Components of utterances that move individually in errors are just those that recombine generatively in language: words, morphemes, and phonemes. Apparently, the freedom to create new orderings of familiar elements also enables serial ordering errors.

Errors reveal that planning takes place in stages, each one sensitive to different linguistic attributes. When words exchange over long distances (e.g. "I don't know that I'd knew one if I hear it" for "I don't know that I'd know ons if I heard it"), they are almost always the same parts of speech. When affixes are stranded in the

exchange, a word moving next to an affix integrates with it, resulting, almost always, in a real word. But when words or stems exchange over short distances (within a phrase), they are generally different parts of speech, and, integrations of stems and affixes may create nonwords (e.g. "slicely thinned"). Phonemes slip over short distances, between words of different grammatical classes, and the errors often create nonwords. Constraints on phoneme slips are phonological . Apparently, successive planning stages focus on progressively narrower sentential domains and on progressively lower-level linguistic elements and attributes.

Finally, errors reveal that the different levels of planning are largely analogous; all involve inserting elements into frames. Exchange errors show this most clearly. In exchanges at all planning levels, element B replaces A in A's intended location, and A replaces B in B's. By implication, when B replaces A, it leaves an empty slot for the left-over A to occupy. Thus, planning involves successively selecting elements ar three linguistic levels and inserting them into planning frame slots.

Execution

Speakers control the respiratory system, larynx, nasal cavity, and articulators of the oral cavity. During speech, they approximate the vocal folds of the larynx (for all but unvoiced sounds, such as/p/ or/s/) so that exhaled air causes the folds to cycle open and closed. We hear the cycling rate as voice pitch. Consonants are produced by constrictions in the oral cavity, for example at the lips (/p/, /b/), or the hard palate using the tongue tip (/s/, /d/). In addition, during some consonantal constrictions (e.g./m/), the soft palate lowers allowing air to pass through the nose. For vowels, the oral cavity is unobstructed, and the tongue body creates different cavity shapes. The acoustic speech signal reflects sound frequencies of the laryngeal source filtered by the changing oral cavity shape.

In contrast to the neat correspondence between units of language and of speech planning as revealed by speech errors, the correspondence between those units and vocaltract movements is complex and poorly understood. Whereas planned elements are discrete and sequentially ordered, movements of the articulators are continuous and overlapped. Planned phonemes "coarsticulate"; that is, movements

of different articulators for different phonemes overlap temporally, and influences of different phonemes converge on common articulators.

Coarticulation is necessary. For example, anticipatory production of vowels during consonantal constrictions prevents production of unintended vocalic sounds as consonantal occlusions are released. However, its effects are to obscure articulatory evidence of the discrete, ordered, context-free phonemes of the speech plan.

Listeners can recover phonemes from speech, however, and so phonemes must be present in some as yet unknown guise. Appreciation of the role of "synergies" in speech production (e.g. Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984) provides a promising lead. Synergies are organized relationships among articulators that achieve a goal. A synergy between the jaw and lips allows closure at the lips for /p/, /b/, and /m/ despite variable contributions to closure by the jaw caused by co-articulating vowels. More generally, synergies allow phonemic goals to be achieved invariantly despite the context-conditioned variability of movements contributing to goal achievement. Successive, partially overlapping implementation of consonantal and vocalic synergies may allow ordered, sufficiently context-free realization of phonemes in vocal tract activity.

Spelling

Writing is only 5000 or so years old, and until recently the skill or writing was possessed by only a small minority of individuals, even in literate societies. Only within the last hundred years has the expectation arisen that all members of an advanced society will be able to both read and write. When considering the problems children now experience in learning to spell, we should bear in mind the fact that the English spelling system was not designed with ease of learning for the general population in mind. Given that we spend perhaps a dozen years learning to write, the 60 practicing the skill, it could be argued that the design features of writing systems should still be biased toward the expert rather than the novice writer.

Writing shares with speech production the property of not being easy to study experimentally. Much of what we know about spelling and writing comes from one of two sources. The first is the analysis of the involuntary slips of the pen that even skilled writers make from

time to time; the second is investigations of the various patterns of spelling and writing impairment that may occur in previously literate individuals as a consequence of brain injury.

The latter approach, based in Cognitive Nueropsychology, provides the bulk of the evidence to support the idea of there being two cognitive "procedures" for spelling available to the skilled speller that are at least semi-independent one from the other. These may be termed the "lexical" and "sublexical" procedures, and may be likened to the two procedures that are being invoked to account for aspects of skilled word recognition in reading.

The "sublexical" procedure is the one that would be invoked by a skilled speller who was asked to construct a plausible spelling for an unfamiliar word or invented nonword. It requires that the sound pattern of a word be broken down into its constituent syllables and phonemes (distinctive speech sounds) and that a string of letters be assembled on the basis of the speller's knowledge of English sound-spelling correspondences.

The use of this procedure is highly error prone in English because of the number of "irregular" words whose spelling does not match their pronunciation (words like women, yacht, and biscuit). Attempts to spell such words using the sublexical procedure result in "phonetic" misspellings such as wimmen, yott, and biskitt. The spellings of irregular words *must* be stored in memory and retrieved as whole units. In fact, the evidence suggests that skilled writers store and retrieve the spellings of *all* familiar words in this way, even the spellings of highly regular words which they could, in principle, assemble *de novo* on each occasion using the sublexical procedure. One line of evidence pointing to this conclusion comes from studies of patients with acquired "phonological dysgraphia" whose brain injury has deprived them of the capacity to spell using the sublexical procedure but who remain able to spell a high proportion of familiar words, both regular and irregular.

Phonetic misspellings look childish, the reason for that being that children, whose store of memorized spellings is still limited, use this procedure regularly, especially in spontaneous writing where a child may not want to break off to seek the correct spelling of a word. A similar pattern can be seen in patients with acquired "surface" dysgraphia. These patients are in many ways the converse of the

patients with phonological dysgraphia. Surface dysgraphics can no longer remember the conventional spellings of many once-familiar words but retain a capacity for spelling by the sublexical procedure. Because that procedure will often generate the correct spelling of a regular word, surface dysgraphics show a regularity effect (spelling regular words more accurately than irregular words), and the errors they make are predominantly phonetic.

The processes by which the conventional spellings of familiar words are stored and retrieved may be termed the "lexical" procedure). Little is known about the storage and retrieval mechanisms involved, though they are generally thought to involve a distinct long-term memory system. Skilled writers will occasionally make involuntary slips of the pen in which they produce a real word different from the one intended. The error word in such circumstances is often either a homophone of the intended word (e.g. writing their for there, or seen for scene) or a word similar in sound (e.g. writing surge for search). Sometimes, though, the error is similar in meaning to the target word (e.g. writing speaking for reading, or last for next). It has been suggested that such errors arise because spellings are retrieved from memory in response to specifications of both the meaning and the sound-form were used it would be impossible to differentiate between homophones with the same pronunciations. The combination of a meaning and a sound-form were used it would be impossible to differentiate between homophones with the same pronunciations. The combination of a meaning and a sound-form specifies each word in a speller's vocabulary uniquely and may be the most efficient combination of cues to use for the retrieval of spellings. The price of this combination is occasional similar-sound and similar-meaning slips of the pen.

The lexical and sublexical procedures, though partially separable in neuropsychological patients, may not be totally insulated one from another in the cognitive systems of normal writers. The fact that the spellings of words that a subject has recently written in a dictation task can influence the way that subject then spells an invented nonword shows that there is some interaction between the lexical and sublexical procedures. Both the lexical and sublexical procedures must result in the creation of a representation of a string of letters that is relatively abstract, and has been termed a "graphemic" representation. The

abstractness of this representation is shown by the fact that it can be used to "drive" output in a range of different forms, including cursive handwriting, print, typing, and oral spelling. Patients have been reported who can still spell aloud but are no longer able to write, suggesting a locus of impairment between the abstract graphemes and the execution of pen movements. Certain slips of the pen (e.g. the tripling of double letters as when a writer produces "butter" for "butter") may have a similarly peripheral locus of origin.

State-Dependent Memory

Though it has long been known that events experienced in a certain physical or psychological state are best remembered in that state – Winslow (1860), for example, related the story of an Irish porter who, having lost a package while drunk, got drunk again and remembered where he had left it – only within the last 25 years has state-dependent memory become a subject of sustained scientific enquiry. In this essay, we chronicle the principle research developments of this period, beginning with mid 1960s demonstrations of drug-dependent memory in animals, and progressing through to late 1980s studies of mood-dependent memory in humans.

The first contemporary researcher to study state dependence in a systematic manner was Donald Overton. In a series of experiments published in 1964, Overton trained undrugged rats to turn right in a T-maze, in order to avoid an electric shock to their paws. All of the rats learned to make the correct directional response within six training trials. On trials 7 through 15, the rats were drugged with a high dose of the barbiturate pentobarbital, and trained to turn left in the maze, again to escape shock. Next, a set of 28 test trials was given, alternately with and without drug. On these trials the rats could escape shock by turning either to the right or to the left. Overton observed that the rats consistently turned right when undrugged and left when drugged, indicating that the animals' memory for a particular directional response could be evoked simply by reinstating the particular drug condition under which the response had been learned.

Today, 25 years after Overton's seminal studies were reported, a large literature exists on the state-dependent effects of drugs in animals, from which several replicable findings have emerged. Chief among these are that drug-dependent memory is (a) associated with

several classes of centrally acting agents, including anesthetics, anxiolytics, and antimuscarinics (and to a lesser extent with hallucinogenics, narcotic analgesics, and tricyclic antidepressants); (b) positively correlated with drug dosage (i.e. performance deficits produced by a shift from, say, drug-present training to drug-absent testing are magnified as the dosage is raised); (c) negatively correlated with response training (i.e. shifts in drug state typically do not impair the performance of well-practiced, overlearned responses); (d) demonstrable in the erformance of either aversively oer appetitively motivated responses by a variety of infrahuman species (e.g. goldfish, mice, monkeys, and, as already mentioned, rats); and (e) often, though not always, asymmetric in form (i.e. a shift from drug-present training to drug-absent testing results in poorer performance than does a shift in the reverse direction). Later in this essay we will discuss asymmetric drug-dependent memory in more detail.

Though drugs are the most common means of eliciting state dependent effects in animals, a number of nonpharmacological methods are also available. As an example, responses learned by rats within minutes after the administration of electroconvulsive shock (ECS) are more apt to be performed if ECS is delivered rather than withheld prior to retention testing. This pattern of results has been obtained in studies involving whole-brain ECS, as well as in those employing either kindled convulsions of the amygdala or subseizure stimulation of the caudate. Other effective nonpharmacological manipulations include movitational drives such as hunger or thirst, changes in body temperature, and even changes in time of day.

Like their animal-learning counterparts, students of human memory initially sought to secure evidence of state dependence through research involving drugs. Representative of this type of research is a study by Goodwin, Powell, Bremer, Hoine, and Stern (1969). Their study entailed two sessions, which we will refer to as "encoding" and "retrieval." During the encoding session, medical students completed a collection of cognitive tasks, such as generating association to words and memorizing a series of pictures. The students completed these tasks after they had consumed either a soft during or a potent cocktail containing an average of 10 ounces of 80-proof vodka. During the retrieval session, held one day later, the students performed a new battery of tasks, which included recall of the previously generated

associations and recognition of the previously memorized pictures. Subjects performed this second set of tasks either in the same drug state – sobriety or intoxication – that they had experienced the day before or in the contrasting pharmacological context.

Results of the test of association recall revealed reliable evidence of state-dependent memory: on average, subjects committed 45 percent fewer errors of recall when their encoding and retrieval states matched than when they mismatched. There was, however, no significant sign of state dependence in recognition memory for the pictures: only 9 percent fewer errors were made under matched as opposed to mismatched conditions.

In the decade following publication of Goodwin et al.'s (1969) research, many other studies of human drug-dependent memory were conducted. Most of these studies employed moderate doses of commonly used – and frequently abused – drugs such as alcohol, amphetamine, and marijuana, and most – like Goodwin et al.'s – produced mixed results, with drug-dependent effects materializing in some situations but not in others. Indeed, by the late 1970s, roughly half of the published studies could be counted as successful demonstrations, and half as failures. Consequently, human drug-dependent memory came to be commonly regarded as a capricious phenomenon of little practical or theoretical significance.

The unpredictability of human drug-dependent memory proved to be more apparent that real. Reviewing the results of 27 studies comprising 57 separately identifiable experimental conditions or cases, Eich (1980) observed that when retrieval of the to-be-remembered or target events occurred in the presence of specific, experimenter-provided reminders or cues (as in tests of cued recall or recognition), evidence of drug-dependent memory rarely emerged (3 of 31 cases, or 10 percent). Conversely, when retrieval occurred in the absence of such cues (as in free recall), the odds of demonstrating a reliable drug dependent effects improved sharply (23 of 26 cases, or 88 percent).

Why was this the case? According to one account (Eich, 1980), if what is stored in memory about a target even includes information about the state – drugged or undrugged – in which the event was encoded, then that state can act as a cue for the event's retrieval. Hence, restoration of the state in which the event was encoded may be seen to provide the remembered with an "invisible" cue for its

retrieval. However, if other powerful "observable" cues are available, then the rememberer will rely on them, which will overshadow the more subtle state cues. When observable cues are not provided, then the invisible state cues come into play. Consequently, drug-dependent effects are more apt to obtain when memory is assessed in the absence of potent observable cues than in their presence.

It is of interest to note that the presence or absence of observable cues is a critical factor not only in the occurrence of *drug*-dependent memory, but of *mood*-dependent memory as well. That is, a shift in affective state from, say, elation at encoding to depression at retrieval has sometimes been shown to impair the free recall of target events, but to have no appreciable influence on cued recall or recognition performance.

Still, even under conditions of free recall, mood dependence seems to be a "now-you-see-it, now-you-don't" effect. Bower (1981), for instance, described a study that used post-hypnotic suggestions to induce happy and sad moods in highly susceptible subjects. Every participant learned one list of words while happy and a second list while sad. Later, the subjects were tested for recall of both lists in either a happy or a sad mood. Recall averaged 70 percent when learning and testing moods matched and 46 percent when they mismatched, signaling a strong mood-dependent effect. Nevertheless, a direct replication attempt by Bower and Mayer (1985) failed to find this pattern or results. Conflicting outcomes – some positive (e.g. Mecklen Brauker & Hager, 1984), some negative (e.g. Johson & Klinger, 1988) – have also been obtained in studies in which moods were modified by non-hypnotic means. It is no surprise, then, that the status of mood dependence as a bona fide phenomenon of human memory is currently a matter of controversy and concern (Bower & Mayer, 1980).

Ultimately, the matter will only be resolved through programatic research aimed at understanding which factors play pivotal roles in the expression of mood-dependent effects. One such factor may be the nature of the moods in which the encoding and retrieval of target events take place. Just as drug-dependent effects in animals are correlated with the amount of drug administered, so too might mood-dependent effects in humans be correlated with the intensity of the moods induced. Thus, for example, it is possible that a shift from an

extremely happy to an extremely sad state would impair memory more than would a shift from extreme happiness to a neutral mood, or from moderate happiness to moderate sadness. Data directly relevant to this possibility have not yet been collected, but it merits mentioning that some of the most robust mood-dependent effects ever reported were obtained in studies involving patients with either manic-depressive illness or multiple-personality disorder – conditions characterized by profound alterations in affect or mood.

A second potentially important factor has to do with how people perceive the relation between their present mood and the events that transpire within it. Recently, Bower (1987) has hypothesized that in order to establish effective associations between ongoing events and the mood in which they occur, contiguity alone between the events and the mood is not sufficient. Rather, people must perceive the events as causing their current mood, for only then will a change in mood cause the events to be forgotten. Though initial tests of Bower's hypothesis have yielded variable results (Bower & Mayer, 1989), additional research is clearly called for. As an aside, it should be noted that a similar idea has been advanced by Fernandez and Glenberg (1985) in connection with environmental context or place-dependent memory. On their account, events experienced in a particular physical environment (a classroom, a courtyard, or the like) may not become associated with that place unless subjects conceive of the environment as causing or enabling the events to happen. This too is an attractive idea that warrants further investigation.

Yet a third factor of potential importance is the source of the target events. Intuitively, it seems reasonable to suppose that events that are generated through internal mental operations such as reasoning, imagination, and thought may be more colored by or connected to one's current mood than are those that emanate from external sources. If so, then a shift in mood state, between the occasions of even encoding and even retrieval, should have a greater adverse impact on memory for internal than for external events.

To investigate this inference, Eich and Metcalfe (1989) performed four experiments that relied on a lengthy regimen of music plus guided imagery to induce rather intense and enduring levels of happiness or sadness. During the encoding session of each experiment, subjects either read a target item that was paired with a category name and a

related exemplar (e.g. *milkshake flavors:* chocolate - Vanilla), or they generated (with a very high probability) the same item when primed with its initial letter, in combination with the category name and exemplar cues (e.g. *milkshake flavors:* chocolate – V). In this manner, memory for one and the same target item could be assessed in relation to its source: either internal (the generated condition) or external (the read condition). Assessment of memory ensued two days later, during the retrieval session, and entailed tests of both free recall and old/new recognition.

Pooling the results from the four experiments, who key findings emerged. First, relative to subjects whose encoding and retrieval moods matched, those who experienced a shift in state recalled 32 percent fewer generated items but only 18 percent fewer read items. Thus, as anticipated, mood-dependent effects were more pronounced for internal than for external events. Second, there was no reliable difference between matched and mismatched moods in the recognition of either generated or read items, which is consistent with the claim, made earlier, that mood dependence is seldom seen when retrieval is tested in the presence of specific, observable cues. Taken together, these two findings imply that the more one must rely on internal resources, rather than on external aids, to generate both the target events at encoding and the cues required for their retrieval, the more likely is one's memory for these events to be mood dependent.

These, then, are three of the factors that may figure prominently in the manifestation of mood-dependent effects. Doubtless there are more, and there is also no doubt that their identification represents a difficult and demanding task. Yet there are other, broader, perhaps even tougher challenges in store. One is to determine whether drug-, mood-, and place-dependent memory are fundamentally different phenomena, or whether retrieval impairments produced by shifts in pharmacological state or physical environment are mediated by alterations in affect. Another is to explore the practical and clinical implications of state-depdendent memory, especially as they may apply to drug abuse, dissociative amnesias, and negative thinking in depression.

A third challenge pertains to the observation, alluded to earlier, that animals who shift from drug-present training to drug-absent testing perform more poorly than do those who experience the opposite

shift in state. This asymmetric pattern has also been found in several studies of drug-dependent memory in humans. For instance, in a study of the state-dependent effects of marijuana on the free recall of conceptually categorized words, the difference in mean percentage recall between the encode intoxicated/retrieve intoxicated and encode intoxicated/retrieve sober conditions (8 percent) was more than twice the difference between the sober/sober and sober/intoxicated conditions (3 percent). Similar results have been realized using other centrally acting depressants, such as alcohol, and also with the psychomotor stimulant, nicotine. What is more, asymmetry appears to be an attribute to mood-dependent as well as drug -dependent memory. To illustrate, one study (Bartlett & Santrock, 1979) showed that although reinduction of happy mood was crucial for the recall of words which children had learned while they were happy, a neutral mood at recall conferred little, if any, advantage on the recall of words memorized during a prior neutral state. Generally speaking, then, it seems that information transfers more poorly in the direction of an abnormal or special state of consciousness (such as drug intoxication or extreme happiness) to a normal or standard state (such as sobriety or a neutral mood) than it does in the reverse direction. How and why this happens is currently unclear; several plausible proposals have been offered, but none has yet been thoroughly articulated. Granted that asymmetric state dependence runs counter both to most people's intuitions and to most established theories of the phenomenon (especially those that incorporate the venerable concept of stimulus generalization), and given that asymmetry is demonstrated by humans as well as animals, its explanation would seem to be an essential step in understanding the mechanisms that subserve state-dependent memory.

Statistical Inference

Psychological interest in statistical inference lies in the study of judgements and inferences that people make about probability and uncertain events – not necessarily involving data normally described as "statistics." Such *intuitive* statistical inferences and judgements arise in many situations when, for example, we forecast the likelihood of a future event (e.g. the chance of rain on a summer weekend) or when we draw general conclusions from our experiences (as when

we form our opinion of the quality of a restaurant by our experience of eating a few meals).

Psychological interest in intuitive statistical inferences stems from the study of decision making. It is generally assumed that choice between alternative actions is made by considering the likely consequence of each decision and trying to assess which will lead to the most beneficial result in the long term. Thus, decision making involves forecasting possible states of the world that are conditional upon the decision made. However, the world is full of uncertainties and the great majority of decisions will require us to estimate probabilities of different outcomes of our choices. *Decision theory* has been developed to deal with the problem of how to make optimal choices under uncertainty, and a number of psychologists have been involved in helping decision makers to apply the theory to their real-world problems, a practice known as "decision analysis". However, such analysis depends upon the ability of people to generate useful estimates probability.

Much research has, then, focused upon the notion of whether or not humans are good intuitive statisticians. Research in the 1960s initially led to an optimistic conclusion (Peterson & Beach, 1967). However, subsequent research has led many, if not most, psychologists in this area to the view that probability judgements are biased by a large variety of factors, and the real decision making is not based upon the principles of formal decision theory. This view is, however, disputed by some contemporary psychologists, who insist on a more rational picture of human judgement (e.g. Beach, Christensen-Szalanski, & Barnes, 1987).

The major theoretical influence in recent years has arisen from the publications of Daniel Kahneman and Amos Tversky, who have advocated what is known as the "heuristics and biases" approach. The theory proposes that intuitive inferences and judgements are based not upon the laws of probability or decision theory, but on simple "rule of thumb" heuristics. One proposal, for example, is that people judge the likelihood of events by the ease with which they can draw examples to mind (Tversky & Kahneman, 1973). While often useful, this "availability" heuristic can lead to systematic errors and biases due to ways in which human memory and retrieval is organized. For example, a clinician might maintain a false theory of the diagnostic

value of a procedure, by selective encoding and retention of case-studies in which the expected association between test result and diagnosis was observed.

Stress, Adaptation to

There have been three different approaches to the study of stress and to the way people cope with stress: the stimulus based or engineering approach, the response based or medico-physiological approach, and the interactional approach exemplified by "appraisal" theories of stress. This essay briefly considers the nature of stress in the framework of appraisal theory, then explores how people attempt to cope with such experiences, and the role of individual differences in coping. This discussion of coping and adaptation to stress is necessarily limited to individual processes, and makes no sustained reference to social processes, such as social support and group coping, although they are important elements in the complete picture.

The engineering approach treats stress as a stimulus characteristic of the person's environment, usually cast in terms of the load or level of demand placed on the person or some aversive or noxious element of that environment. Stress, so defined, produces a strain reaction. In contrast, the medico-physiological approach considers stress as a "generalized and non-specific" response to aversive or noxious environmental stimuli. Stressors give rise, among other things, to a stress response. These approaches have been judged to be inadequate both in terms of their ability to account for the available data and in terms of their theoretical sophistication. Essentially, they fail to take account of individual differences and the perceptual cognitive processes which underpin such differences and which drive the experience of stress. The interactional approach to the study of stress treats it as a psychological state: the internalization of a particular transaction between the person and his or her environment. Such theories owe much to the work of Lazarus (1966) and his notion of "cognitive appraisal." It is interesting to note that not all avowedly *cognitive* theories of stress adopt this interactional perspective. Those that are rooted in the assumptions of information theory, and usually wedded to laboratory experimentation, tend to have more in common with the engineering approach in their treatment of the processes involved and in their essentially "local" nature. In contrast, those

concerned with "appraisal" offer more of a "framework" for understanding and draw on a variety of methods for developing and testing such a framework. Framework theories place an emphasis on "synthesis" and "consensus" between different models and include natural observation as well as experimentation as legitimate data sources.

Appraisal Theories of Stress

Appraisal models of stress make explicit its psychological definition. They treat stress as a psychological state which is the internal representation of a particular and problematic transaction between the person and his or her environment. Appraisal is the evaluative process that imbues such situtional transactions with meaning. Numerous authors have offered appraisal theories of stress and there is a high degree of consensus between them.

In many ways, Lazarus's ideas formed the initial platform on which most other appraisal theories have been constructed. According to Lazarus's approach the outcome of a stressful transaction is mediated by two components, appraisal and coping.

Appraisal is composed of primary and secondary processes. With regard of the former, individuals ask themselves the question: "Is this particular encounter relevant to well-being and in what way?" If the encounter is relevant and is defined as stressful, rather than as benign, three more specific appraisals are then made, those of loss, threat (of harm), and challenge. Primary appraisal is associated with the emotional content of stressful transactions. Secondary appraisal is concerned with the question: "What, if anything, can be done to overcome or prevent harm?"

Influenced by the work of Lazarus and his colleagues, and also by that of McGrath (1970), the theoretical contribution of Cox, and of Cox and Mackay, has been developed over the last 12 years. Originally their "transactional" model of occupational stress was described in terms of five stages (Cox, 1978). The first stage, it was argued, represents the sources of demand faced by the person and is part of his or her environment. The person's perception of these demands in relation to ability to cope represents the second stage: effectively primary appraisal. Stress was described as the psychological state which arose when there was a personally significant imbalance

or mismatch between the person's perceptions of the demands on him or her and his or her ability to cope with those demands. The psychological and physiological changes which are associated with the recognition of such a stress state, and which include coping, represent the third stage of the model, which leads into the fourth stage, which is concerned with the consequences of coping. The fifth stage is the general feedback (and feed forward) which occurs in relation to all other stages of the model.

This model has been further developed in three respects. First, the authors have attempted to describe the process of primary appraisal in more detail (Cox, 1985; Cox & Mackay, 1981). According to the authors cognitive appraisal appears to take into account:

1. the external and internal demands that people experience, *matched* against.
2. their personal coping resources and behaviorual style;
3. the constraints under which they have to cope; and
4. the support that they receive from others in coping.

Second, the stress process (including coping) has been set in the context of "problem solving" and a clear distinction has been made between primary appraisal (is there a problem?) and secondary appraisal (how and how well can I cope with it?) (Cox, 1987). Primary appraisal is seen as a continual process of monitoring while secondary appraisal is seen as a more discrete activity contingent on the outcome of primary appraisal. Third, there has been some discussion of the problem of measuring stress based on this approach (Cox, 1985, 1990) with the development of possible subjective measures of the experiential (mood) correlates of the stress state.

There is a conceptual overlap between the work at Nottingham of Cox and his colleagues, and that of French, Caplan, and van Harrison at Michigan. Those authors and their colleagues have been responsible for the development and popularity of the Person-Environment Fit (P-EF) model of occupational stress. This model distinguishes between objective and subjective variables, both of which can refer to either the environment (E) or the person (P). Furthermore, the authors refer to both *demands* on the person's abilities and *supplies* for the person's motives (such as income). Mental health is then defined in terms of four dimensions: (a) objective person-environment fit (P-EF); (b) subjective P-EF; (c) accessibility of self (fit between objective and

subjective P variables); and (d) contact with reality (fit between objective and subjective E variables). This model, and the hypotheses that it has been used to generate, have been tested by the authors through a series of survey-based field studies.

Useful additions to "appraisal" theory have been made by Pearlin and Schooler (e.g. Pearlin & Schooler, 1978). They distinguish between acute and chronic stress,, and suggest that together eventful experiences (acute events) and chronic strains (hassles) drive the perception of stress: possibly through the experience of life events causing people to be more aware of daily hassles. They also make a distinction between, on the one hand, coping resources (both personal and social) and, on the other, coping responses – the former being what is available to the individual and the latter what they actually do.

Appraisal theories generally offer "frame-works" for understanding, and sadly have not always been used to generate the detailed hypotheses necessary for experimental validation. Having said this, there has been a proliferation of quasi-experimental and survey based field studies which have contributed to the development of such theories.

Coping

There have been two approaches to the study of coping: that which considers coping as a problem solving *process,* and that which attempts to classify the different types of coping which exist and produce a comprehensive *taxonomy* of such strategies.

Coping as problem solving. Coping is increasingly being viewed as a problem solving strategy (e.g. Cox, 1987). Cox (1987), for example, has described a cycle of activities, beginning with recognition and diagnosis (analysis) following through actions and evaluation to re-analysis, which possibly represents the ideal problem-solving process. However, all this implies only the positive side of a double-edged sword and Schonpflug and Battmann (1988) have emphasized a more negative "problem generation" side: that is, by adopting the wrong actions, or by failing in coping, a person may create further problems and stress. At the same time, Meichenbaum (1983) argues that "catastrophizing" or over-reacting to such failure serves no adaptive purpose. It is often said that one of the few positive aspects to coping with stress is that the person learns from such experience.

However, Einhorn and Hogarth (1981) suggest that there are, at least, three problems with this proposition: first, one does not necessarily know that there is something to be learned; second, what is to be learned is not clear; and third, there is ambiguity in judging whether one has learned. Furthermore, the problem solver may be fully occupied and not have any spare cognitive capacity for learning, and the emotion associated with stress may interfere with the learning process.

Notwithstanding, detailed models of coping based on rational problem solving and decision making have been offered.

Coping taxonomies. In an attempt to develop a parsimonious understanding of the coping process, various authors have tried to identify and categorize the different coping strategies which exist.

Lazarus (1966) sees coping as having three main features. First, it is a process: it is what the person actually thinks and does in a stressful encounter. Second, it is context dependent: coping is influenced by the particular appraisal that initiates it and by the resources available to manage that encounter. Finally coping is defined "independent of outcome."

In dealing with a stressful situation, Lazarus (1966) argues, the person usually combines task–and emotion–focused coping strategies. The former attempt some form of action directly targeted on dealing with the source of stress: adaptation *of* the environment, while the latter attempt to attenuate the emotional experience associated with that stress: adaptation to the environment. The perceived success, or otherwise, of such strategies feeds back into the appraisal process to alter the person's perception of the situation. Lazarus and his colleagues also emphasize that the importance of the situation to the individual is critical in determining the intensity of his or her response.

Pearlin and associates (e.g. Pearlin & Schooler, 1978) further develop this view and distinguish between responses concerned with changing the situation, those concerned with changing its meaning, and those relating to the management of the symptoms of stress. In a different vein, Miller (1979) has distinguished between two informational styles which she terms "blunters" and "monitors": the former tend to use denial strategies and the latter information-seeking strategies in relation to stressful situations. However, these various coping strategies are not meant to be mutually exclusive and most authors emphasize that no one type is necessarily better than any

other in solving a problem. Most people use a mixture of strategies in most situations, although certain situations may tend to be associated with particular types of strategy.

Dewe (1987), for example, examined sources of stress and strategies used to cope with them in New Zealand ministers of religion. Using factor-analytical techniques, he identified five clusters of strategies: seeking social support, postponing action by relaxation and distracting attention, developing greater ability to deal with the problem, rationalizing the problem, and drawing on support through spiritual commitment. It was possible to classify 33 percent of the strategies which made up these clusters as task focused and 67 percent as emotion focused. The most frequent source of stress experienced by the ministers related to the emotional and time difficulties associated with crisis work, and the experience of such problems appeared to be associated with coping by seeking social support and rationalizing the problem.

Individual Differences and Coping. Coping does not occur in a vacuum, but is context dependent and partly reflects an interaction between personal and social resources, and situational demands and constraints. Attribution of control, hardiness (Kobasa, 1979) and type A behavior, informational styles, and several other factors, have been implicated in the coping process and in determining the effects of stress. For example, increases in perceived control have been shown to be important in reducing the effects of stress on performance and on health. Control is one of the defining traits in Kobasa's (1979) concept of hardiness: control, commitment, and challenge. Hardy individuals report less illness than non hardy individuals when exposed to high levels of stress. Hardiness appears to affect the person's appraisal of events: hardy individuals, for example, experience events in a quantitatively similar but qualitatively different way to type A individuals. Kobasa, Maddi, and Zola (1983) suggest a further difference between the two groups: hardy individuals may be intrinsically motivated while type A individuals appear extrinsically motivated.

Coping may be seen as functional in its attempts to manage demands, by either changing them, redefining them, or adapting to them. The styles and strategies used need to be relvant and applicable to the situation at hand. The choice and successful use of these

responses will be determined by the nature of the situation, by the personal and social resources available, and also by the type of causal reasoning adopted (appraisal).

Stroop Effect

Asked to name the ink color of a stimulus, people are much slower to say "green" for the incongruent word RED in green ink than for the control stimulus XXX in green ink. This is the Stroop effect, first reported by J. Ridley Stroop (1935) in his dissertation. Of particular interest is the fact that interference is asymmetrical: asked instead to read the word aloud, people are no slower to say "red" for the incongruent stimulus than for the control word Red is normal black ink. Later studies also demonstrated a small facilitation effect: ink color naming time is faster for congruent stimuli such as the word Red in red ink than for the XXX control.

There are several hundred Stroop-related studies in the literature. Numerous analogs now exist, such as the picture-word task, where a word is embedded in a line drawing of an object (e.g. a picture of a cat containing the word Dog). Here, interference and facilitation occur when naming the picture but not when reading the word. Because of its magnitude and reliability, the Stroop effect has long been treated as a benchmark measure of attention.

Two related explanations of the Stroop effect have dominated theoretical efforts: relative speed of processing and automaticity. The "relative speed of processing" account starts from the longstanding observation (Cattell, 1886) that people read a word faster than they name an ink color (or other perceptual quality of a stimulus). The idea is that word reading and ink color naming occur in parallel, but a response is produced faster by word reading. Assuming a limited capacity response buffer, if the response requested is the ink color name, the wrong response reaches the buffer first, causing interference.

The automaticity account, also due to Cattell, emphasizes the much greater practice accrued by word reading relative to ink color naming. This extensive practice has made reading automatic: it is beyond volitional control and does not require attention. In contrast, ink color naming is a controlled process requiring attention. An automatic process will interfere with a nonautomatic one, but not

vice versa, producing asymmetrical interference.

These two theories differ in that automaticity does not hinge on a direct mapping from processing speed to interference outcome. Thus, a slower but more automatic process could interfere with a faster but less automatic one. Recent studies suggest that the "relative speed of processing" view cannot accommodate all of the data, most notably when the onset asynchrony between the word and the ink color is systematically varied. Instead, a continuum of automaticity has been suggested as a better explanation. Direction and magnitude of interference are determined by the relative automaticity of the two stimulus dimensions (assuming a reasonable index of automaticity, such as extent of practice). This explanation directly links attention to learning and memory.

Student Learning

Although students in universities and colleges have been the subjects of experimental studies in cognitive psychology for many years, it is only recently that the ways they tackle their everyday academic learning tasks have been systematically examined and related to cognitive theory. A rapidly developing area of research has made use of existing theories from cognitive psychology, but it has also led to the identification of quite different concepts rooted in the experiences of the students themselves. This essay presents brief summaries of the research derived from existing cognitive theories which are well represented in other entries, with greater detail being provided here of the less familiar experiential concepts.

The dominant approach to research on student learning in North America has taken concepts and theories of learning and memory and investigated their application to student learning. In Europe, a quite different tradition has been developed over the last ten years, commonely by educational researchers. They have used both qualitative analyses of interviews and questionnaire surveys to develop concepts and theories derived from students' descriptions of their own learning. These different approaches make different assumptions both about the nature of learning and about data collection and analysis.

The first approach is familiar. The data are collected from carefully controlled laboratory experiments in which specially designed

learning materials avoid the effects of previous knowledge. Knowledge is generally viewed quantitatively and measured in terms of, say, the percentage accuracy of verbatim Recall. Concepts and theories then emerge from the statistical analysis of the data. These are used either to generate inventories in which students report their processes of learning and study activities within that predetermined theoretical framework, or to design interventions intended to improve students' learning strategies.

The second approach is a more recent development. It starts from the experience of the students themselves. Learning is generally seen as a process by which students alter their conceptions of the topic being studied. The outcomes of learning have, therefore, to be assessed qualitatively in terms of different levels of understanding or conceptions. Beside investigating the qualitatively different outcomes of learning, this research has also looked at the contrasting ways in which students tackle common academic tasks. The initial work was based on semi-structured interviews with transcripts being subjected to regorous qualitative analysis. This process involves the repeated reading of the transcripts to identify the most distinctive similarities and differences among the students, which are refined iteratively until the most powerful descriptive concepts and categories are identified. These concepts, rooted in the students' own descriptions, have then been used to construct self-report inventories.

In exploring these two approaches we shall be looking for convergence in the descriptions of student learning, which may also suggest more appropriate ways of helping students to learn more effectively.

Concepts from Cognitive Psychology

Perhaps the two most important lines of development deriving from cognitive psychology itself are those which have investigated learning strategies in experimental studies and those which have investigated the dimensionality of study strategies using inventory surveys. Both areas of research have subsequently led to training procedures designed to improve the efficiency of student learning.

Experiments on Learning Strategies. Although there is a long history of attempts to assist students with their learning strategies, until recently the evidence of their success was not strong. There

was little transfer from the training situation to everyday studying. There is now a set of related studies which show much greater benefits, and this is believed to be attributable to their firmer roots in cognitive psychology. Probably the best known system of strategy training is that developed by Dansereau and his colleagues using the acronym Murder. Students were encouraged to set a positive Mood for studying, concentrate while reading on Understanding, actively Recall what had been read, try to Detect errors in that recall, use mental imagery and previous knowledge to Elaborate what was being learned, and finally to Review what was to be remembered. This method has only rather indirect linkages to cognitive theory, and yet more recent methods, emerging more directly from work on metacognition, have pointed up similar components as being valuable and effective in strategy training, even in the earlier stages of education. In particular, training in elaboration and monitoring seems to lead to improvements in academic performance.

Surveys of Study Strategies. The idea of differing levels of processing in the memory has provided a productive theoretical framework for thinking about student learning. The idea that recall will be stronger where a deep level of processing, including elaborative processing, has been utilized has led to the development of inventories and related intervention strategies. The most direct application of these ideas can be seen in the work of Schmeck who developed a so-called Inventory of Learning Processes. This inventory contains four scales: deep processing, elaborative processing, fact retention, and methodical study. The first three of these scales contain items guided by cognitive psychology theory, while the final scale described study activities. Beside being rooted in this body of theory, these four dimensions have also repeatedly emerged from the factor analyses of the items.

Weinstein has developed the so-called Learning and Study Strategies Inventory, which contains scales derived from both cognitive psychology and the study skills literature. Its theoretical base includes ideas about both memory and metacognitive processes, while from the literature on study skills came, not just descriptions of organized studying, but also motivation and anxiety. A scale of information processing has been found to correlate fairly closely with Schmeck's elaborated processing (Schmeck, 1988). In parallel with construction of the inventory, Weinstein has also developed a course on individual

learning skills designed to help students form a more systematic and reflective approach to studying. The inventory is used to monitor progress and to adapt the course to student needs. Evaluations show substantial improvements both in the quality of work and grade-point average (Weinstein et al., 1988).

Educational Research

The work on student learning carried out by educational researchers has led to the definition of a series of concepts and categories derived from the qualitative analysis of interview transcripts and the parallel development of inventories to measure those concepts. The qualitative studies have been concerned with learning of subject matter as well as the processes of learning and studying.

Recalling Information and Developing Concepts. Effective learning of academic subject matter depends not only on the storage and recall of information, but also the formation and appropriate use of abstract concepts. The contrast between rote learning and meaningful learning is well established in the psychological literature, but a substantial amount of student learning in higher education and in schools is of an intermediate kind. It is of the form that Ausubel has described as meaningful reception learning in which information is conveyed in a lecture or a book and "directly" absorbed by incorporation into cognitive structure with little, if any, conscious effort. Of course, subsequent reorganization and systemization occurs, particularly during revision, but the initial process has neither the rehearsal properties of rote learning nor the active establishment of linkages which is the characteristic of more active forms of meaningful learning. This type of learning, although prevalent, has been rather ignored by researchers.

A good deal of effort has, however, been put into describing the differing conceptions held by students, particularly of scientific concepts. The starting point was in schools, where many pupils were found to have naive concepts of physics which interfered with their understanding of the scientific concepts they were supposed to be acquiring. Marton has since developed a research methodology – phenomenography – which describes in terms of a few distinctive categories the range of conceptions held by students. In this research, knowledge and understanding are seen as being reconstructed and adapted as new information is acquired or problems solved (Marton

& Entwistle, 1989), and it is interesting to see how recent ideas in cognitive psychology have suggested a similar, more holisitc, view of knowledge and the way it is stored.

Intellectual Development. Perry (1970) was one of the first researchers to use interviews to create a set of categories describing differences between students in their ways of studying. Analysis of the transcripts suggested that there was a distinct series of "positions" through which students progressed, representing a dimension of intellectual development from "dualistic" to "relativistic" thinking. Dualism indicates a belief in the existence of "right" and "wrong" answers, with successful students being those who are able to acquire a large number of "right" answers, and effective teachers being those who know those answers and can convey them in the clearest possible form.

The dualist position does not last long and thereafter students begin to see both the nature and the implications of relativism – that many answers are incomplete, have to be qualified, or are provisional. The degree of uncertainty perceived in relativism creates anxiety in some students and leads them to resist further progression. Those who persist, however, begin to recognize how evidence is used in the construction of sound argument, then begin to make a "commitment" to their own theoretical stance and to defend it effectively.

The implications of these ideas on the intellectual and ethical development of students have been explored in relation to the level at which first-year courses should be pitched, the need to emphasize skills as well as content in curriculum planning, and the counseling of students who find relativism threatening.

Approach to Learning. Perhaps the most influential concept describing aspects of student learning was introduced by Marton and his colleagues. They asked students to read an academic text and to be ready to answer questions afterwards. The questions explored not just what was understood, but also how the students went about reading the article. Using regorous qualitative analysis, they found that the concept which most clearly distinguished between the ways students tackled the article and their levels of understanding was their *approach to learning.* Students adopted either a deep approach or a surface approach to the task.

The crucial difference lies in the contrasting *intentions* shown

by students. A *deep approach* draws on a sophisticated conception of learning with an intention to reach a personal understanding of the material presented. To do this, the student has to interact critically with the content, relating it to previous knowledge and experience, as well as examining evidence and evaluating the logical steps by which conclusions have been reached. In contrast, a *surface approach* involves a simple conception of learning as memorization and an intention merely to satisfy task or course requirements, seen as external impositions largely remote from personal interests. The surface approach can still be active, but it relies on identifying the elements within the task most likely to be assessed, and then memorizing those details (in Marton, Hounsell, & Entwistle, 1984).

Marton's study was carried out as a naturalistic experiment, but in everyday studying the influence of assessment on approach has to be taken into account. Ramsden asked students about their ways of tackling their academic work. Again the deep and surface approaches represented important distinctions, but it was also necessary to add a third category – *strategic approach* – which related to studying as well as to learning. In this approach, the student adopts deep and surface approaches in a combination designed to achieve the highest possible grades. The approach involves using well organized study methods and careful time management, but above all there is an alertness to any cues given by tutors about what they are looking for in deciding grades or marks, or what questions they are going to set in the examinations (Entwistle & Ramsden, 1983). Students appear to have two distinct foci of attention – the content and the teacher's reward system. While lecturers expect the students to focus on the former, assessment demands shift attention to the latter.

Although the idea of approach to learning had its origins in a single naturalistic experiment, it has since been extended or modified to describe most tasks found in everyday studying (Marton et al., 1984), and has been introduced into many staff training workshops to help faculty understand the different ways students go about their academic work.

Study Orientations. Attempts to operationalize the concepts derived from the interviews led to the Approaches to Studying Inventory (Entwistle & Ramsden, 1983), which also contained scales describing motivation and study attitudes and methods. Factor analyses

of the items showed that each approach was associated with a different form of motivation – deep approach correlated closely with intrinsic motivation, surface was linked with fear of failure and instrumental motivation, while strategic was associated with need for achievement. These combinations of scales held together so consistently that they have been referred to as "study orientations" – meaning, reproducing, and achieving. A final "non-academic" orientation is characterized by low levels of motivation, negative attitudes to studying, and disorganized study methods.

A similar pattern linking motive to strategy has been reported independently by Biggs (1987) using his own theoretical framework and a different inventory. He has used this inventory, in ways similar to Weinstein, to suggest how students can be weaned from reliance on mechanical study skills to metacognitive awareness of their own goals and strategies.

The approaches to learning and studying are, however, not simply to be seen as reflecting consistent individual differences. They also reflect reactions to the students' perceptions of their learning environment. Approach is influenced not just by assessment procedures, but also by the perceived quality of the teaching, freedom in learning, and workload (Entwistle, & Ramsden, 1983).

One of the important growth points in this area of research seems to be the investigation of the parallelism, and the interactions, between study strategies and perceptions of the learning environment. It is becoming clear that approaches to learning have correlations not only with distinctive motivations, but also with contrasting perceptions of their learning environment (Entwistle in Richardson, Eysenck, & Warren-Piper, 1987; Ramsden in Schmeck, 1988; Janssen in Entwistle & Marton, 1990; Entwistle & Tait, 1990).

Reconciling The Contrasting Methodologies

There are rather few opportunities for cognitive psychologits and educational researchers to discuss their different approaches to describing student learning. When they do attend the same conferences, the differences in methodology and theoretical perspective are marked (Richardson et al., 1987) and mutual criticism strong. Cognitive psychologists attack both the methodology and the concepts used by educational researchers. For

example, the qualities analyses have been described as unconvincing, with the "approach to learning" being seen as a diffuse concept with too many component parts. In their turn, educational researchers have criticized the ecological invalidity of applying concepts derived from artificial experiments to the very different context of student learning, and also point out the limitations imposed by what they see as a mechanistic conception of learning and a reductionist research strategy (Entwistel & Marton, 1990).

It is clear that cognitive psychologists seek a different level of explanation from educational researchers. They focus on the common cognitive processes within the individual, paying little or no attention to individual differences or contextual variability. For the educational researcher, the internal cognitive processes are of only peripheral interest. What concerns them is how to explain, and influence, the study behavior of the student. The terminology used by the cognitive psychologist also creates a barrier of communication with their colleagues – hence the need for concepts which, while broad and somewhat diffuse, are essentially communicable and carry with them metaphorical meanings which encourage reflection on teaching and learning.

The research by cognitive psychologists has, however, led to interventions which have been shown to support student learning. The value of continuing to investigate common cognitive processes, and of exploring their applications, has thus been demonstrated. There is nevertheless a remaining area of concern about this approach. The focus on cognitive processes in these interventions leads to an overemphasis on the importance of strategy training. This has two disadvantages. First, much cognitive strategy training puts control of the learning in the hands of the instructor who decides what treatment to apply. Yet it has been found more beneficial to encourage students to take charge of their own learning and develop metacognitive awareness. Second the strong influence of assessment procedures and teaching indicates that the quality of student learning can only be improved by a re-evaluation of the whole learning environment (Newble & Clark in Richardson et al., 1987; and Eizenberge in Ramsden, 1988), and a recognition of how the various aspects of the environment are likely to interact with the strengths and weaknesses of individual students (Entwistle in Richardson et al., 1987).

Study Skills

Skills which lead to effective studying are those that make use of the processes that normally lead to good comprehension of what is being studied and, subsequently, to good recall. It is usually the case that the material being studied has a logical and meaningul content and structure, although identifying this may at first present a challenge to the student. Comprehension and memory are usually closely related. If the full sense of the text is not properly comprehended then it will be difficult to remember. Bransford and Johnson (1972), for example, showed that when subjects were given, in advance, the title "washing clothes" that clarified the theme of an otherwise obscure passage which they were given to read, subsequent recall of the passage was more than doubled.

Both comprehension and memory depend upon the use of the existing knowledge that the person possesses in the active construction of a new mental representation of the information that is being conveyed in the material under study. It is, therefore, not surprising that study techniques usually emphasize an active approach to the material to be learned, along with methods to ensure that the learner's existing knowledge of the area has been fully activated.

There are many further opportunities to improve upon the memorization of information beyond its initial comprehension. However, many students expect to be able to remember the details of a textbook after passively reading it through. They rarely check shortly after reading the passage to see how much they can remember. Such testing of one's memory is a very valuable component in effective study. Not only does it provide awareness of what has and has not been learned but it also makes the future recall of the information that can be remembered far easier.

Several other central factors in determining memory also need to be incorporated into effective study skills. One is that the more salient and interesting the material appears to the learner, the better it will be acquired. If the information falls within an area in which the person is already knowledgeable then the new information will be acquired more quickly (e.g. Morris, Tweedy, & Gruneberg, 1985). This is in addition to the need to use existing knowledge to comprehed the material being memorized. Recall depends heavily upon the availability of cues at the time of remembering that will

activate the memories. Where the information being learned has a framework or structure that can be used to organize both the learning and the retrieval, then memory is often considerably improved (e.g. Bower, Clark, Lesgold, & Winzenz, 1969). Study techniques need to reveal such organization in the material under study and, if necessary, impose a framework if one is not clearly discernible. Memory is much influenced by the similarities between items that have been encoded. Any activities by the learner which improve the distinctiveness of the information being memorized will increase its probability of later retrieval.

In the light of these general principles about comprehension and memory, various study systems have been developed. The oldest and most widely recommended of such schemes is the SQ_3R system (Robinson, 1946; Higbee, 1988). The technique involves the three stages: Survey, Question, Read, Recite, Review. This technique is principally intended to improve learning from textbooks. In the first, Survey, stage the student reads the chapter summary, skips through looking at headings, pictures and graphs, and so on and attempts on obtain an overview of the topics covered. In so doing they begin to identify the structure of the material and to cue from their memory knowledge that is related to the chapter or book under study. In the Question stage the learner again skims through the material, this time asking questions based upon the headings of the sections. These questions should relate both to the learner's own interests and to what they expect to find within the learning material. By so doing, they maximize their interest in the material and develop an active approach to their subsequent reading. In the next stage they read through the chapter. Then, in the Recite stage they reread the chapter, but do so by first attempting to anticipate what will come in each section and to answer the questions they yet earlier. In the final, Review, stage the student surveys his or her success in the Recite stage and examines the extent to which the original questions have been answered and locates such weaknesses and the need for further learning as will be required.

Study skills such as SQ_3R can improve the efficiency of study. However, students to whom the system is taught often complain that it involves too much mental effort and tend to fall back on more passive approaches to learning, despite the dangers for their future

performance.

In addition to study systems, there are other important aspects of memory that students should consider. One is the normal finding that spaced or distributed practice leads to better learning than does massed practice. The benefits of distributed practice have been demonstrated in language learning by Bloom and Shuell (1981) and in Learning to Type by Baddeley and Longman (1978). Thus, for example, more will be learned from three one-hour sessions on some given material than from one three-hour session.

When studying to remember information later, overlearning is beneficial. That is, rather than practicing upon and studying the material for just so long as one can report it from memory, the student should continue studying and recalling the material so that recall becomes progressively easier. Studies of over-learning have shown that it increases long-term retention. Material that is very considerably overlearned is particularly resistant to forgetting and may be retained for a lifetime while information that has been crammed for an examination will be rapidly forgotten (Bahrick, 1984).

Subliminal Perception

Subliminal perception is the phenomenon of responding to stimuli below the awareness threshold. It has been shown to occur for all sense modalities and in the context of several different research paradigms. The latter include presenting stimuli of low intensity or short duration, or at frequencies beyond the normal range for conscious perception. Signal to noise ratios for achieving "perception without awareness" may also be brought about by masking or by presenting stimuli in sensory channels that are not currently mediating conscious perception.

Subliminal perception implies that the processes responsible for conscious perceptual experience are not identical with those which mediate the transmission of information through the brain from receptors to effectors. Whereas the latter depends upon the classical sensory pathways linking peripheral receptors with their cortical projections, Consciousness of sensory inflow necessitates a coincident contribution of cortical excitation from the ascending fibres of the reticular system.

The theory that subliminal perception occurs when external stimulation is too weak or too brief to produce sufficient activation of the reticular system is supported by the research findings of Libet and his co-workers (1967). Recording from the somato-sensory cortex of fully conscious human subjects, they were able to detect electrical potentials initiated by tactile stimuli presented at subliminal intensities. When stimulation was increased to supraliminal intensities, subjects reported awareness of the stimulus at the same time as the wave form of the evoked potential manifested a contribution from the reticular activating system.

Yet other researches have shown that visual evoked responses, galvanic skin responses, verbal behavior, conscious perception of supraliminal stimulus arrays, and even dreams, may be significantly influenced by the meaning of verbal and/or pictorial stimuli presented below the awareness threshold. The data from such studies imply that subliminal stimuli may be subjected to extensive preconscious processing involving unconscious long-term memory and emotional classification of sensory inflow.

By far the most extensively researched manifestations of subliminal percetion are those of perceptual defence and vigilance. The data from a variety of experimental paradigms suggest that prior to awareness of a visually presented word or picture, which is gradually increasing in brightness or duration of exposure, a preconscious analysis of the latter's meaning may result in the raising or lowering of the threshold for conscious perception of the material in question.

The fact that stimulation below the awareness threshold is not subject to conscious appraisal has led to several clinical applications of subliminal perception. These include the investigation of unconscious psychopathology (Silverman, 1975), the treatment of various neurotic disorders, and the identification of accident prone candidates among applicants for flying duties in military airforces. In the Defence Mechanism Test the candidate for a flying career has to describe what he sees when flashed a complete picture in which the central figure is flanked by a "subliminal" threatening male face.

Prolonged validation studies have confirmed that those candidates whose perception of the central figure is distorted by the unconsciously

perceived "threat" will be particularly accident prone during military training. By using such a test in the selection of military fliers many countries have shown a considerable saving in lives and aircraft.

Subliminal stimulation has also been used in a therapeutic context. For example, the reduction of anxiety by subliminal presentation of reassuring messages has proved helpful in reducing compensatory overeating (Silverman et al., 1978) and improving performance at mathematics (Ariam, 1979).

Presumably because of the threat it poses to notions of free will there has been considerable resistance to accepting the reality of subliminal perception. A favorite criticism has been that so-called subliminal effects are only responses to consciously perceived fragments of the stimulus array. A recent study by Groeger (1984) which showed that semantic influences were *greater* for a subliminal stimulus than for one presented at the awareness threshold suggests that subliminal perception cannot be explained away by this "partial cue" hypothesis.

Tip-of-the-tongue Phenomenon

There are many times when people try to recall an answer, are unable to at that time, but still know something about the answer, i.e. the tip-of-the-tongue phenomenon. For instance, suppose someone is asked, "What is the capital of Denmark?" Just because the person does not immediately recall the answer, that does not necessarily mean the person does not know anything about the answer. Sometimes, a person who cannot recall the answer may nevertheless be able to recall something about the *attributes of the word* that is the correct answer. For instance, in response to the aforementioned question a person who cannot recall the answer "Copenhagen" may still recall that the answer begins with "C" or even "Cop" or that it contains four syllables. Other times, a person who cannot recall the answer may nevertheless be able to recall information about the *attributes of the referent* of the answer. For instance, in response to the aforementioned question a person who cannot recall the answer may still recall that the referent of the question is a city near the sea or that it is the city in which the Tivoli Gardens are located. Recall of either of these two kinds of attributes, in the absence of recall of the answer itself, is called the "tip-of-the-tongue phenomenton" (first investigated by Brown & MacNeill, 1966).

The phenomenon is conceptually different from-and is easily confused with) a related phenomenon called the "feeling of knowing." The latter refers to people's predictions about their subsequent performance on a nonrecalled item. The most common kind of prediction is about the person's subsequent likelihood of recognizing an answer that currently cannot be recalled (first investigated by Hart, 1965). For instance, for the question above, a person might not be able to recall the answer but might be quite certain that he or she could recognize it (e.g. pick it out of a list of Scandinavian cities).

Sometimes the attributes that are recalled during the tip-of-the tongue phenomenon may serve as a basis for the feeling of knowing (Koriat & Lieblich, 1974). For instance, if someone who could not recall the answer "Copenhagen" was nevertheless able to recall that the answer, then on that basis the person might predict that he or she could correctly recognize the answer if confronted with a list of Scandinavian cities (this and other related bases that might underlie the feeling of knowing are discussed by Nelson, Gerler, & Narens, 1984). Although the tip-of-the-tongue phenomenon and the feeling of knowing are not identical, both of them are indicants of underlying knowledge about an answer that cannot be recalled.

12

Visual Dominance

The term visual dominance refers to the fact that, under conditions of conflicting input from different sensory modalities, visual input tends to dominate. For instance, Gibson (1933) had subjects wear distorting prisms that made straight lines appear curved. Subjects then maintained that a straight edge felt curved to their touch when they saw and felt the edge at the same time. Similar effects occur in studies of prism adaptation, where subjects wear distorting lenses for a prolonged period until they adapt to behaving appropriately to the distored visual input (Kohler, 1972). Thus, walking up stairs is initially very difficulty, even though proprioceptive and kinesthetic input from our joints should be sufficient to enable the behavior to occur (since we can walk up stairs with our eyes closed).

Different accounts can be offered as to why visual dominance occurs. Posner, Nissen, and Klein (1976) proposed that it occurs because visual signals tend to be less alerting than other signals. An auditory tone works as a warning signal to visual and auditory stimuli even when the signal is unexpected and cannot be predicted. In contrast, a visual warning signal may sometimes be effective only when its occurrence is expected (Posner et al., 1976). Because visual signals need not alert us automatically we may need to pay them greater attention than other stimuli – so producing visual dominance.

However, visual dominance could occur for other reasons too. Some visual signals, particularly those produced by movement of observers or objects in environment, may link directly to motor

responses (such as balance reactions). Vision dominates some aspects of motor behavior because these direct links do not exist for other senses (such as audition). If an object looms toward an observer, the observer tends to sway backward (and there can be forward sway if an object looms away) – suggesting a direct link between balance reactions and visual cues for movement (Lee & Aaronson, 1974). The extent of visual dominance then depends on the relative strength of other signals to the motor system – with vision dominating when proprioceptive cues are reduced (as in young children).

Visual Illusion

An illusion is a perception which deviates systematically from physical fact. This definition is due to Gregory (1973) who nevertheless notes deep philosophical problems lurking in the notion of "physical fact." In practice, however, most psychological studies of illusions ignore these difficulties by taking as the question of interest: why should the perception of a given entity differ from measurements of the entity using instruments such as rulers, protractors, light meters, and so on? The latter measurements thus serve as operational definitions of "physical reality." The requirement in the definition of illusions for *systematic* deviations is to exclude the random fluctuations around some average response that are characteristic of all observations, whether by the senses or by human-made instruments. Numerous visual illusions exist: a useful reference work which reviews most of them and the classic psychological theories proposed to account for them is Robinson (1972).

Sharply differing views are held about the value of illusions for furthering understanding of perceptual mechanisms. At one extreme, J.J. Gibson held that illusions arise from impoverished stimulation that is quite uncharacteristic of the richly structured ambient optic array which normally confronts the observer. Gibson used this controversial assertion to conclude, equally controversially, that illusions provide a misleading basis on which to erect theories of perception.

Gregory (1970, 1973) has championed the opposite view. He holds that illusions can provide valuable clues about how perceptual processes work both when they produce correct outputs and when they produce illusions. He suggests that illusions might be generated in two basic ways: (a) malfunction of physiological mechanisms, and

(b) inappropriateness of the strategies built into those mechanisms. He suggests these two types of illusions might be called, in biological terms, "physiological" and "cognitive" illusions respectively; or in engineering terms, "mechanism" and "strategy" illusions.

Gregory believes that the "after-images" experienced following exposure to bright lights following a period of dark adaptation are clear examples of physiological/mechanism illusions. This is because their origins can be traced to physiological processes in the retina which are "over loaded" by bright lights following dark adaptation. He calls this state of affairs a "loss of calibration" (Gregory, 1973, pp. 58, 73).

Gregory offers the size-weight illusion as "probably the clearest and simplest case of a purely cognitive illusion" (1973, p. 60): if an observer lifts up two objects of differing sizes but with the same scale weight, then the smaller of the two objects is perceived to be heavier than the larger. Here the conclusion seems inescapable that the misperception is due to normal assumptions about the size/weight relations of objects being inappropriate for the atypical densities of the two objects generating the illusion. The illusion is a compelling one despite there being nothing unusually impoverished about the stimuli. If Gregory is right, an explanation of the effect must include an account of the strategies incorporated into the mechanisms subserving weight perception. Here in strategies are ones that use visual information to access stored information about the typical weights of objects in order to set the muscle effort likely to be necessary to lift things.

Gregory's view of cognitive illusions falls within the empiricist and constructivist traditions of perceptual theorizing. The famous dictum summarizing this approach is that of Helholtz: "perception is unconscious inference." This is explicated by Gregory (1973, p. 51) as follows:

> Perceptions [are] "conclusions" [or "hypotheses"], more or less likely to be true, depending on the sensory data available and the difficulty of the perceptual problem to be solved .... [they] are given by inference, from data given by the senses and [from data] stored in memory. On this view, any perception may be false, just as any argument may be false. It may be false because its assumptions are incorrect, or because the form of the argument is fallacious. On this view of perception, illusions take on the same importance that paradoxes and ambiguities have for philosophers concerned with the nature of argument; or how data can be used to discover a truth, or a fact.

Anyone who has programed a compute will find Gregory's basic distinction between mechanism and strategy an easy one to grasp in principle. Computer users frequently have to ponder whether some unwanted aspect of their machine's operation is a hardware fault or limitation, or some failing in the software. That question is not always an easy one to answer, and the same is true of visual illusions. Indeed, Gregory observers that the "Physiology or Cognition?" question is a particularly tricky one for geometric distortion visual illusions. These are phenomena to which he has devoted considerable attention because he thinks that at least some of them might be caused by the inappropriate triggering of size constancy mechanisms that work for images of normal scenes but which produce "errors" when applied to two-dimensional line drawings for which they were not designed.

But even if it is granted that the question "Physiology or Cognition?" is a useful one to bear in mind when considering these and other visual illusions, it needs to be remembered that when an illusion is judged to have its roots in a misapplied processing strategy, that does not render uninteresting the mechanisms that carry out that strategy. In other words, to fully understand visual systems requires understanding those systems at both the mechanism *and* strategy levels (and perhaps also other levels besides).

One can be in sympathy with Gregory's view that many illusions reflect the application of inappropriate assumptions and yet still doubt whether studying illusions is, in itself, an entirely satisfactory way of studying perception. The inconclusive nature of the controversy, after more than two decades of experiments, surrounding Gregory's interpretation of geometric distortion illusions as caused by inappropriately triggered size constancy scaling mechanisms is not an encouraging sign. The missing required ingredient seems to be computational experiments aimed at testing the putative value of processing strategies suggested by an illusion. Such experiments force attention to many important details which are otherwise all too easily neglected in psychological theories, and they thereby provide much needed clarification of the issues involved.

In computational theories of vision, a major requirement is identifying *constraints* which serve as assumptions (usually about the nature of the viewed world) upon which can be built useful methods for interpreting visual input data to satisfy some clearly stated goal. The mention of assumptions here clearly finds an echo in Gregory's views on perception: if the assumptions embedded in a computational theory are inappropriate for a given set of inputs then incorrect outputs ("illusions") will be generated. Numerous examples of this kind of reasoning can be found in the computational vision literature (Frisby, 1979).

Sometimes the constraints used in computational theories derive from knowledge of human vision, sometimes a computational study itself leads to the identification of a useful constraint. In the latter instance, this may in turn lead to experimental studies of biological vision systems aimed at testing whether they implement the constraint by seeing if they suffer from an illusion predicted by the computational theory when fed with certain sorts of inputs. Here the study of illusions plays a valuable role, but a secondary one, that of providing evidence as to whether a given computational theory is implemented in the visual system of interest. That is, the interpretation of illusions is much the same as in Gregory's cognitive account, but the research program has a different starting point: first, the computational problem and the identification of constraints, second, the search for an illusion to see if the constraints are used in a biological vision system. But whichever way around the process of enquiry proceeds, a key element for success seems to be computational experiments providing rigorous tests of whether the right questions are being asked. A recent collection of review papers on computational work relating to biological visual systems is provided by Landy and Movshon (1990).

13

Working Memory

The term generally refers to a system which is involved in both cognitive processing and in the transient storage of information that is being processed during the performance of a wide range of cognitive tasks. Cognitive tasks of any degree of complexity (e.g., problem solving; comprehension of text) involve a number of different processing stages, and working memory allows the updated "state of play" on the task to be stored and readily available. There is clearly some overlap between the notion of a "working memory" and that of a "short-term memory store" but the concepts actually differ in a number of important ways. The short-term memory store was typically regarded as unitary, i.e. it was assumed that it operated in a single, uniform fashion (e.g. Atkinson & Shiffrin, 1968). In contrast, the most developed theory of a working-memory system (Baddeley, 1986; Baddeley & Hitch, 1974) is based on the assumption that working memory consists of a number of different components.

The short-term memory store was used almost exclusively to account for data from cognitive tasks in which the subject's sole task was to remember various items of information. In contrast, working memory is, at least in principle, involved in numerous tasks where the primary focus is not on memory at all. For example, the focus in mental arithmetic is on the performance of accurate arithmetical calculations, but there is good evidence that working memory plays an important role on mental arithmetic tasks (Hitch, 1978).

The notion of working memory, or of something rather similar,

is to be found in various theories in cognitive psychology. However, the only systematic theory which is centered squarely on working memory is that originally proposed by Baddeley and Hitch (1974), and subsequently modified by Baddeley (1986). In its original version, Baddeley and Hitch (1974) put forward a working-memory system consisting of the following three components: a modality-free central executive closely resembling Attention; an articulatory loop which can hold a limited amount of phonological or speech-based information; and a visuo-spatial scratch pad, which is devoted to spatial and/or visual coding. This working-memory system is hierarchical. The central executive is at the apex of the hierarchy, and beneath it are its two "slave" systems, namely the articulatory loop and the visuo-spatial scratch or sketch-pad.

The three components of working memory will now be considered in more detail. One of the major characteristics of attention is that it possesses limited capacity. Baddeley and Hitch (1974) assumed that the same was also true of the central executive, but they did not investigate its functioning systematically. Baddeley (1986) subsequently argued that the central executive may be rather similar to the supervisory attentional system described by Shallice (1982). Among the purposes for which the supervisory attentioanl system is used are to perform tasks requiring planning or decision making, to handle situations where poorly mastered response sequences are involved, and generally to fulfill a trouble-shooting function when lower levels of the processing system are inadequate to the task in hand.

Baddeley (1986) also speculatively suggested that damage to the frontal lobes may impair the functioning of the central executive component of working memory. According to Rylander (1939, p. 20), the typical pattern following damage to the frontal lobes involves "disturbed attention, a difficulty in grasping the whole of a complicated state of affairs, well able to work along routine lines .... cannot learn to master new types of task, in new situations." As this description makes clear, those suffering from frontal lobe damage behave as if they no longer have a control system which permits the flexible and appropriate deploment of processing resources. This is exactly what one would expect if the central executive were located in the frontal lobes.

The articulatory loop has strictly limited capacity. The evidence suggests that its capacity is determined by temporal duration in similar fashion to a tape loop. Baddely, Thomson, and Buchanan (1975) proposed that the articulatory loop can hold approximately as much verbal information as can be spoken out loud in two seconds. They discovered that subjects could provide accurate immediate serial recall of more words that could be pronounced rapidly than of words that had a long pronunciation time, and they argued that the apparent time-based limitation on immediate serial recall reflected the limitations of the articulatory loop.

Subsequent neuropsychological evidence indicated that the original view of the articulatory loop was an oversimplified one. Several patients appear to have patterns of impairment which could not be accounted for by the theory proposed by Baddeley and Hitch (1974). For example, there was a patient, P.V., who was studied in some detail by Basso, Spinnler, Vallar, and Zanobio (1982). She seemed to have reasonably intact articulatory processes as revealed by her ability to recite the alphabet. In spite of this, she did not use this ability when asked to provide immediate serial recall of a series of visually or audiotorily presented letters. This was demonstrated by using an articulatory suppression task, in which the subject has to repeat something simple over and over again, at the same time as the letters for recall were being presented. Despite the fact that articulatory suppression largely prevents use of the articulatory loop for other purposes, it was found that P.V.'s performance on immediate serial recall was unaffected by the presence of the articulatory suppression task. The crucial additional finding was that her memory for auditorily presented letters was worse when these letters were phonologically similar (i.e. when they sounded alike) than when they were phonologically dissimilar. In other words, P.V. appeared to be processing phonologically (i.e. in a speech-based manner), but she was doing this without making any use of articulation.

In the revised version of the working-memory model (Baddeley, 1986), a distinction is drawn between an active articulatory process that is linked to speech production and a passive phonological store that is concerned with Speech Perception. Phonological information can enter the phonological store in a direct fashion through auditory presentation of verbal material, or it can enter indirectly either through

subvocal articulation or via phonological information stored in long-term memory. Within the framework of this theoretical framework, the impairments shown by the patient P.V. (and by other patients) can be accounted for by assuming that she has a deficient phonological store. Since auditory presentation of words always involves the phonological store, whereas visual presentation does not, it follows that P.V. should experience greater problems with auditory presentation than with visual presentation, and that is, indeed, the case.

An interesting issue relevant to this revised theory concerns the precise mechanisms involved in the subvocal articulation component of the articulatory loop. One obvious possibility is that the speech musculature is involved, but is also entirely possible that these articulatory processes occur at a rather more central level. It is relevant here to consider patients suffering from anarthria. This is a condition in which impairments to the system controlling the speech musculature mean that they are unable to speak, but in spite of this they possess essentially normal language abilities in other respects. It has been found that patients with anarthria can make use of subvocal articulation in spite of the inability to use their speech musculature (Baddeley & Wilson, 1985). As Baddeley (1986, p. 107) concluded.

> The loop and its rehearsal processes are operating at a much deeper level than might at first seem likely, apparently relying on central speech control codes which appear to be able to function in the absence of peripheral feedback.

The visuo-spatial scratch-pad (or sketch-pad) was defined by Baddeley (1986, p. 109) as "a system especially well adapted to the temporary storage of spatial information, much as a pad of paper might be used by someone trying for example to work out a geometric puzzle." In their investigation of the sketch-pad, Baddeley and Lieberman (1980) started from the assumption that there is an important difference between spatial and visual processing. For example, someone who is blind is by definition largely or entirely unable to use visual processes. However, a blind person may still possess rather accurate information about the spatial layout of objects in a room in spite of his or her visual impairment.

The experimental approach adopted by Baddeley and Lieberman (1980) made use of a task in which the location of digits within a matrix was described in an auditory message, with the subject being

required to reproduce the matrix afterwards. The message was constructed so as to be either easy or difficult to visualize. In an earlier study by Baddeley, Grant, Wight, and Thomson (1975), it was found that concurrent performance of the pursuit rotor (i.e. tracking a light along a circular track) produced a severe impairment in memory for easily visualized messages, but had no effect on memory for messages that were difficult to visualize.

Since the pursuit rotor involves both visual and spatial processing, it is not clear from these findings precisely why there was an interference effect on easily visualized messages. Baddeley and Lieberman (1980) devised a task which involved spatial but not visual processing. The subject, who was blindfolded, tried to point at a moving pendulum. When he or she was successful, there was an auditory signal. When this task was performed at the same time as memorizing an auditory message, it impaired recall of visualizable messages but not nonvisualizable ones. In contrast, a concurrent task that involves visual but not spatial processing impaired performance more on the nonvisualizable message than on the visualizable one. The implication of these findings is that the processing of visualizable messages depends primarily on spatial processing, and it is for this reason that a concurrent spatial task has such a disruptive effect.

In sum, there is reasonable evidence to support the notion of a three-component working memory system. However, far more is known about some components than others. The detailed functioning of the articulatory loop is reasonably clear, whereas it has proved very difficult to examine the central executive. The modular nature of the working-memory system is generally in line with current thinking. However, several cognitive neuropsychologists have postulated the existence of several modules, and it may well be that the working-memory model will need to be expanded beyond its current three components.

Writing

The first rendering of spoken language in pictograms and subsequent developments of the phonetic alphabet, the printing press, and possibly the word processor are landmarks in intellectual and technological history. For centuries the study of writing has been central to rhetoric and education. Curiously, writing was virtually

ignored by psychologists until the 1970s, but then drew the attention it deserves alongside other complex forms of thinking. Cognitive psychology diverged from the earlier literary approaches to writing by focusing on the process of composing rather than on the resulting text.

Flower and Hayes (1980) presented an influential model of writing processes. Planning included generating ideas, organizing ideas, and setting goals to be achieved in the structure of the text. Translating or sentence generation included the subprocesses of language production, such as lexical selection, that are also involved in speech. Reviewing included reading the evolving text and editing it for errors. These processes are controlled by an executive monitor that allocates limited attentional capacity. They occur within the context of a task environment, consisting of the writing assignment and the text produced thus far, and the writer's long-term memory.

Memory holds vast realms of diverse knowledge, methods of composing, and personality characteristics. The writer possesses extensive knowledge of language, topic, and audience. Strategies for retrieving relevant knowledge and for creating new concepts are an important aspect of the writer's method, as are the tools selected for composing and editing. Finally, personality characteristics such as anxiety and achievement motivation are needed to account for writer's block, on the one hand, and prolific productivity, on the other.

Empirical work by Flower and Hayes and other suggests that writers recursively attend to planning translating, and reviewing processes throughout prewriting, first draft, and subsequent draft phases of text development. For example, even when working on a final draft the writer interweaves planning, translating, and reviewing. This constant juggling of processes is highly effortful and frequently overloads the writer's limited attentional capacity.

Research on planning indicates important development differences. For example, in generating ideas immature writers rely heavily on an associative search of memory that Bereiter and Scardmalia (1987, pp. 7-8) have termed "a knowledge telling strategy." Identifiers of the topic, genre, and discourse type serve as cues that automatically retrieve related material from memory. As writers develop in skill, they continue to use knowledge telling, but add a strategy called knowledge transforming. This is a directed search of memory

that arises from the writer interactively reflecting on both content problems (e.g. How do I define this concept?) and rhetorical problems (e.g. How do I express this thought clearly?). The writer consciously searches for solutions to the difficulties at hand. Because of reflective planning, the writer's thoughts emerge through the act of composing itself; composing causes a transformation of knowledge, not merely a telling of it.

As another example, immature writers fail to organize information effectively, often presenting it in the order in which it was generated. More experienced writers attempt to order the text in a manner that is appropriate to communicating the subject matter to a particular audience using a particular type of discourse structure (Flower, 1979). For adult writers, creating an outline during prewriting results in superior quality texts. Kellogg (1988) found that outlining improved the quality of composition by alleviating attentional overload during the drafting of an essay; writers who outlined had less need to juggle planning, translating, and reviewing simultaneously.

Some research on translation overlaps with studies of Speech Perception and Speech Production. For example, whether an uautonomous syntactic component operates during sentence generation, as predicted by generative grammar models, is open to debate in both speech and writing. Other research focuses on the differences between speaking and writing (Faigley, Cherry, Jolliffe, & Skinner, 1985, p. 45). For example, oral discourse achieves cohesion in part through nonverbal means, such as posture and gestures, whereas written discourse elies exclusively on explicit lexical and syntactic ties.

Finally, investigations of reviewing indicate a progression of editing strategies as young writers mature (Graves, 1979). Initially, only changes in the forms of letters occur. Then spelling changes followed by word substitutions are noted. Later still comes rearrangement of sentences and paragraphs. Young writers review relatively little, wheres mature writers generally review extensively for long, difficult writing assignments. It should be noted, however, that highly skilled writers who are thoroughly comfortable with their writing assignments are capable of planning a document in such detail that revising is largely unnecessary (Faigley et al., 1985, p. 60).

Although the Flower and Hayes (1980) model has been influential,

major disagreements with it abound (e.g. Nystrand, 1982). One argument against it is that sentence generation and planning cannot be neatly separated. Another is that writing is inherently a social act, a point underemphasized by their model. Instead of focusing on the cognitive processes of an individual, the social approach studies the writer as an agent in a literate community of discourse.

14

Aging and Cognitive Science

The attempt to apply to assumptions, methodology, and models of cognitive psychology to the study of old age provides an important test of the extent to which the discipline can illuminate the human condition. A particularly provocative theoretical challenge has been whether and how the functional models for hypothetical "steady state" systems, found in the cognitive psychology of the 1960s and 1970s, but never in nature, can be developed to describe complex patterns of change. These static models were linked to a "modular" neuropsychology which assumed that changes in cognitive performance following brain damage were only of theoretical interest if they could be related to anatomically well defined lesions, putatively focal to particular hypothetical cortical mechanisms and implicitly to very tightly definable cognitive functions associated with these areas. In this context the study of the aging brain seemed theoretically sterile since it forces attention away from localized and complete extirpation of supposedly "modular" structures with putatively independent and discrete functions and toward the cognitive effects of global and diffuse biochemical, histological, and neurophysiological changes.

The study of cognitive aging also forces us to consider problems which contemporary neuropsychology still prefers to neglect – the extent to which "local" damage can have global effects on cortical function and the interrelationship between the extent to which cognitive skills have been practiced, or memory databases have been developed, and their relative vulnerability to insult or degeneration. In short, the

relationship between learning and the plasticity of the neurophysiological substrate on which cognitive processes depend.

Given that even very healthy people obviously do show gross changes in cognitive performance as they grow older, the main questions for cognitive gerontology are few and simple: At what age are the first changes detectable? How quickly do changes then proceed? Do all cognitive skills decline together – perhaps because they are all common outcomes of the same global neuropathological processes - or do different cognitive skills, and so perhaps the separate "modular" brain systems on which they depend, "age" at different rates? (Do we lose all our faculties together, or one at a time, and, if the latter, in what order?) Can we alter the rate, or the pattern, of cognitive change; for example can heroically persistent practice of cognitive skills prevent their decline? Finally there are the existential questions which psychology has alternately muddled and avoided for far too long: What is the *experience* of growing old? How far are individuals conscious of the cognitive changes they suffer, and can they adapt to, or circumvent, them? The present state of cognitive gerontology can be assessed in terms of the extent to which it can give sensible answers to these basic questions.

Cognitive changes in old age are driven by two distinct processes: gradual and, in our present state of knowledge, irreversible degenerative changes in the central nervous system (CNS) (e.g. Rockstein & Sussman, 1979) and the steady accretion, throughout life, of an increasing database of knowledge about the world and an armamentarium of cognitive skills. Considering this latter source of change it is not disingenuous to consider "cognitive aging" as a process which is continuous throughout life and which can only artificially be distinguished from the "developmental processes" of infancy, childhood, and adolescence. Without stretching a metaphor the same can be said of CNS aging. Most of our body cells have a limited lifetime, reproducing themselves by division before dying so that their population is entirely renewed every nine to twelve years. Neurons, in contrast, do not reproduce so that at any age the CNS units we have left are survivors of a much larger number with which we were born. These veterans may continue to grow new dendrites and form new connections with each other throughout their, and our, lifespans. If, as seems probable, changing networks of connections between

neurons constitute the biological basis of information storage in the CNS, the fact that our entire corpus of neurons, and so of established connective networks, does not change every decade is an essential design feature providing some stability for memories, information, and skills which we would otherwise have to continually refresh or lose.

The period of maximum neuron loss occurs within the first three years of life and, speculatively, may reflect a process of the establishment and preservation of networks of "useful" connections with loss of units which are not incorporated into such system. Autopsy data suggest that neuron loss is continuous throughout life so that by age 75 even healthy individuals may have lost up to 10 percent of their young adult brain mass, with temporal and frontal cortex more depleted than other areas. These gross changes are paralleled, perhaps preceded, by changes in individual cells whose bodies and processes begin to incorporate "neurofibrilliary tangles" of lipids, and which may progressively be denuded of their dendritic connections to other units. Surviving units may greatly increase the number of their dendritic processes, speculatively in compensatory attempts to preserve connectionvity. Increasing numbers of "senile plaques" mark miroscopic areas of neural degeneration. All these histological changes have been observed in young adult brains and, as far as we know, it is the increasing prevalence, rather than the precise nature, of these indicators that characterizes the aging of the brain. It is also likely that even in "normal" aging, specific changes in larger neural structures such as the cholinergic, dopaminergic, and serotinergic, nuerotransmitter, systems which are strongly marked in age-related neurological disorders such as Dementia of the Alzheimer type (DAT) or Parkinson's disease may also occur in "normal" aging. It is still a matter for debate how far these conditions (for example, DAT) reflect step functions of neuropathology, or acceleration of a continuous process toward an inevitable terminal state.

We should not bleakly conclude that intellectual decline may begin even before maturation is completed because it is likely that the gross redundancy of the CNS allows substantial losses of individual units to be tolerated without loss of efficiency, and it is certain that the highest levels of cognitive skills require extremely extended practice, making the passage of time a precondition, rather than a restraint, of

attainment. The common-sense idea that a steady decline in absolute capacity is offset by continuous improvements due to learning and practice fits the general pattern of data available. The largest bodies of data come from standardizations of tests of performance IQ and of creativity, and these show measurable declines in average scores between the ages of 25 and 35 years (e.g. Wechsler, 1956). Recent investigations of learning of complex, fast interactive video- games show impairment between 18 and 36 years (e.g. Rabbitt, Benerji, & Szemanski, 1989). Such results are typical of tasks in which novel information has to be processed as rapidly as possible, but strongly contrast with tasks which require unpaced deployment of acquired information, such as vocabulary tests, which may show no loss and even some increase throughout a lifetime of 80 years. Horn (1982) first formalized this contrast as a distinction between "fluid intelligence," associated with maximum information-processing speed, which steadily declines after a peak in the twenties, and "crystallized intelligence" or acquired information and intellectual skills, which may show little age-related loss.

This distinction reappears in statistical analyses of lifespan achievements pioneered by H.C. Lehman (e.g. 1957), whose analyses of biographical data suggest that in all professions a relatively early peak is followed by a long slow decline, but that peaks are later and declines less prolonged in disciplines such as history and literature which require the long acquisition and massive deployment of vast bodies of information than in those such as Mathematics which demand the ability to perceive new problems and find entirely original solutions to them. Lehman's conclusions have been criticized and qualified (e.g. by Dennis, 1966; Fox, 1983) but, in general, it seems that the lifetime course of cognitive attainment does represent a shifting balance between continuous loss of "raw" information-processing capacity and problem-solving ability and steady acquisition of information and useful techniques.

Thus the simple question "How fast do cognitive changes proceed once they have begun?" can only be discussed evasively in terms of a shifting balance between individuals' hypotheticals initial maximum potentials and their lifetime opportunities and motivation for self-development. It also raises severe methodological problems. Most data on age changes of any kind are "cross-sectional"; that is,

mean performance scores for different, samples of individuals in successive age groups. With the possible exception of vocabulary test scores and measures of "crystallized intelligence", all such data show a peak of performance attained in the twenties or early thirties followed by a continuous decline which sharply accelerates in the sixties, seventies, and eighties. A much more laborious alternative is to obtain longitudinal data on the course of aging within rather than between individuals by repeatedly testing them over many years. The logistic difficulty of such comparisons makes such studies rare and, usually, incomplete. However, some authors have found that particular individuals may show little or no change in cognitive function as they age through and beyond their seventies. This has led to an alternative to the "continuous decline" model of aging in which a long plateau of indefinite duration is succeeded by a "terminal drop" in performance, possibly associated with pathologist leading to immanent death; (e.g. Jarvik, 1983). On such assumptions "time backward from death" is a much more interesting index of current cognitive age than "time forwards from birth." It is important to note that it is not at all incompatible that "pleateau and drop" age-performance trajectories are often obtained for longitudinal data on individuals while "continuous decline" trajectories are invariably obtained for cross-sectional data on means of successive age samples. The terminal drop model assumes that while the trajectory for any individual shows a plateau, individuals tend to begin their terminal declines at different ages and that the ratio of "droppers" to "survivors" increases with group mean age. Thus it follows that there will be a continuous decline in mean performance levels for successive groups although the age-performance function for each individual is, in fact, rectangular.

Unfortunately, longitudinal studies do not offer a simple, if arduous, resolution to this problem because they also have their characteristic methodological difficulties. Retrospective reanalyses of published data by Schaie and Labouvie-Vief (1974), among others, have found that longitudinally assessed rates of change may differ between successive generations of individuals who apparently age at different rates. More disturbingly, recent and as yet unpublished very large-scale studies by Flynn and by Raven reveal further difficulties because scores on a wide range of IQ tests obtained from very large samples of young adults, in many different cultures, over the last 50

years seem to show a continuous increase with historical decade. It remains uncertain to what extent this may reflect general medical and social improvements such as in health, hygiene, nutrition, and general education or increasingly widespread specific practice on problems similar to those encountered in IQ tests. In any of these cases it is clear that the original standardization data for older IQ tests is no longer valid for contemporary population samples. This possibility that the performance indices we use as well as the people we test may "age" during the course of longitudinal investigations compounds difficulties of interpretation when large groups of people are, perforce, progressively practiced on the tasks on which they are repeatedly assessed. An important, optimistic conclusion is that the "continuous decline" age-performance functions currently observed may be contrasted with rectangular functions which represent the ideal, obtainable by medical and social advances, both for populations and for individuals.

The next question is whether all cognitive functions change at the same rate, possibly driven by a global change in CNS efficiency, or whether functionally distinct abilities decline at different rates, perhaps determined by idiosyncratic rates of change in the neurophysiological "modules" which support them.

It has been noted that cognitive skills which require the rapid handling of novel information show early and marked declines while those based on the retrieval of learned information or procedures show little change with age. Recent unpublished work by Brayn, Perfect, and Rabbitt suggests that even when information can be accurately retrieved, the speed with which it can be accessed markedly declines as age advances. It seems that age may affect the *accessibility* of information even when its *availability* is apparently unimpaired. Horn (1982) has shown that age changes in performance on tests of "fluid intelligence" show factor commonalities with tests of information-processing rate. Eysenck (1986) and Jensen (1985), among others, make a case for the idea that "g," a single factor or "general intelligence" on which most IQ tests appear to head highly, can be identified as information-processing rate, a performance index common to most laboratory tasks and so, plausibly, a sensitive index of general neural efficiency. Sathouse (1986) suggests that slowing of information-processing rate is the most general and sensitive index

of cognitive aging. Experiments by Waugh and Barr (1980) show that slowing of information-processing rate may underlie and determine age differences in learning efficiency, since learning rates of young and older individuals may be equated by allowing the elderly longer study times to compensate for their slower information uptake. These and many other observations seem to support. Salthouse's suggestion that a coherent theoryof cognitive aging may be predicated on decline in a single general factor of information-processing speed which affects all mental functions.

The attractive idea of a single common factor, which accounts for individual differences in fluid intelligence and for its decline in old age, together with all other aspects of cognitive performance, founders on points of logic rather than of empirical evidence or theoretical plausibility. When we attempt to quantitatively assess human performance, whether by laboratory tasks, intelligence tests, or evaluations of cognitive efficiency in everyday skills, we can only measure how fast or how accurately people do things. As Hick (1952) first pointed out, the index of "information-processing rate" necessarily represents a composte measure, or "equivocation" in which speed and accuracy are jointly taken into account. Thus, *information-processing rate is a performance index logically implicit in all our empirical measurements* rather than, as taken by Eysenck, Jensen, and Salthouse, a privileged index of some functional property of the CNS upon whose relative value theories of intelligence or of cognitive aging can be based. Because nearly all our measures of task performance, including scores on all time-limited intelligence tests, are directly translatable into measures of information-processing rate, it is statistically inevitable rather than functionally important that analyses of between-subject variance across a variety of tasks should yield a single, general factor common to all of them.

Thus, statements that information-processing rate can be considered a single, principal factor in cognitive aging, or in individual differences in intelligence, is bound to be true, but cannot be analytic; i.e. they do not address the question of whether discreate functional "modules," and the skills which they hypothetically support, age at equal or at different rates. Such questions can only usefuly be asked by identifying skills which, by studies of focal lesions, have been shown to be associated with particular brain systems and then by

carrying out very large-scale investigations to examine the joint regressions, across very large populations, of performance indices for these skills against each other, against chronological age, and against measures of information-processing rate so as to discover to what extent these sets of indices show common, and to what extent idiosyncratic, variance between individuals.

An interesting hypothesis which lends itself to such treatment is that the right cortical hemisphere "ages" faster than the left so that skills such as accurate perception and memory for patterns may decline more rapidly with age than "left hemisphere" skills such as logical computation or linguistic dexterity. While some interesting and well conducted studies seem to support this hypothesis (e.g. Albert, Duffy, & Naeser, 1987), as we have seen, the degree to which cognitive skills survive into old age seems to depend on the extent to which they are based on information and procedures learned over very long periods of time. Our models of the human cognitive system are not yet sufficiently developed to answer important questions about plasticity and its maintenance into old age, i.e. whether identifiable cognitive performances such as "pattern perception" or "language" have been rendered "domain specific" as much by the very extended practice which created them, and which may maintain them in old age, as by the architecture of the particular, anatomically localizable neuronal "modules" which support them. The further question as to whether, and how, practice of a skill may contribute to the continued health and longevity of the particular neural substrate that maintains it also remains obscure. In the present state of knowledge we can only be certain that any regimen which maintains bodily health must benefit the CNS also. As far as we know "jogging one's memory" is probably more effective if undertaken with the feet than by attempts at intellectual athleticism.

A final question is the extent to which individuals are aware of, and can adapt to or learn to circumvent, the cognitive changes which may occur as they grow old. This is of considerable interest in current cognitive theory as a particular instantiation of the general question of how far individuals consciously aware of their own cognitive processes and how accurately they can monitor and adapt them. The invariable methodology for such investigations is to ask individuals to rate their own ability, or their frequencies of lapses on self-report questionnaires.

Such studies of changes in self-report with advancing age give remarkably consistent results which highlight some logical problems with this technique. When large populations of older people are interrogated about their memory efficiency, their relative reported frequencies of cognitive lapses, or the frequency with they mislay their possessions, or their perceived memory deterioration over the last 30 years, the results are remarkably consistent. Indivisuals' self-reports bear no relation to their IQ test scores, and correlate *negatively* with their ages, so that individuals in their fifties complain of poorer performance than do individuals in their sixties and seventies.

It is, of course, difficult to assess individuals' subjective self-reports against their actual everyday competence, but recent studies have found precisely the same results in comparisons where performance can be objectively assessed – in the case of mild deafness or growing, mild visual handicap. Although, as we would expect, individuals in their fifties show significantly less impairment, their complaints are more frequent and severe. The clue to this paradox is that individuals have no way of making absolute judgements about their own cognitive or, it would seem, sensory abilities. They can only make relative judgements, assessing themselves against companions or colleagues or against the demands of their environments. Individuals in their fifties are usually very much engaged in life and may be unduly self-conscious of their aging in confrontation with taxing demands or able young colleagues. As people grow oder their lives and environments change, and may become steadily less demanding. Individuals who maintain homeostasis with their social and work environments receive little feedback which might draw their attention to changes in their capacities. It is pleasant to be able to conclude a commentary on cognitive changes in old age with the reflection that individuals who anxiously perceive deterioration in their cognitive efficiency are probably those who are, as yet, least affected, and have least cause for concern.

Aging and Memory

Most people believe that their ability to learn and remember declines with advancing age, and in general this belief is confirmed by experimental evidence. However, demonstrations of age-related decrements in ability must be treated cautiously for two main sets of

reasons: the first is that other factors besides age may be responsible for the observed differences, and the second is that the overall picture of decline masks a more complex pattern in which some types of memory are found to hold up well with age whereas other types decline substantially. In the first category, there is a greater incidence of depression in older people and some memory failures are clearly associated with this condition. Other difficulties of interpretation arise from the usual practice of comparing a young group in their teens and twenties with an older group in their sixties and seventies; the two groups almost always differ in terms of their early educational experiences, and may differ in health, motivation, recent educational experience, and other variables. However, the alternative research strategy of carrying out longitudinal rather than cross-sectional studies is somewhat impracticable, and a reasonable solution may be found in work demonstrating the modifying effects of health, verbal intelligence, activity levels, and the like on the relations between memory and aging.

The second complicating factor is the very different patterns of age-related effects observed on different memory tasks. For example, digit span (the longest string of random digits that a person can reproduce accurately) shows a very slight decline with age, as does memory for world knowledge and for vocabulary. On the other hand, the ability to recall a long set of unrelated words or objects decreases dramatically with increasing age, as does the ability to hold some material in mind while simultaneously carrying out manipulations on the material or while dealing with further incoming material. Various suggestions have been made to explain such differential patterns. One is that different memory stores exist (e.g. short-term and long-term stores) and that the processes of aging affect the stores differently. An analogous suggestion is that the proposed memory systems of episodic, semantic, and procedural memory are tapped differentially by various memory tasks, and that these systems age at different rates. These and other accounts will be assessed after a review of the evidence.

Short-Term Memory

The term "short-term memory" is used here to denote a type of task rather than any special store or system; the tasks in question are

those in which the person must hold a small amount of material in mind for a matter of 2-30 sec. As previously mentioned, performance on standard digit span declines only slightly from the twenties to the seventies, although *backward* digit span, and the "alpa span" test devised by Mary Gick and myself do show greater age-related deficits. In the alpha span test, a short list of words is presented and the subject's task is to rerrange the words mentally before reproducing them in alphabetical order. It seems that the requirement to manipulate materials held in mind is associated with poorer performance in older people.

Older traditional short-term memory tasks include the Brown-Peterson paradigm in which subjects are presented with three or four letters or unrelated words to be recalled after a filled retention interval of 0-30 sec. Age differences on this task are again slight, as they are also in the retrieval of the last few items in a list of words presented for free recall (Craik, 1977). Performance levels on the digit span test, the Brown-Peterson task, and in the last-mentioned task of retrieval all depend on the ability to hold a small number of items in mind and reproduce them in the same order. This ability has been described as depending on "primary memory" and shows little change as a function of aging.

It was suggested previously that the requirement to manipulate material held in mind does give rise to age-related decrements. This type of manipulation is inherent in working memory tasks, and indeed other such tasks appear to be vulnerable to the effects of aging. As one example, age differences were found in a task in which subjects were presented with a series of sentences whose truth had to be verified (e.g. "Sparrows build nests in the spring," "An elephant is smaller than a mouse"); then at the end of the series subjects recalled the final word from each sentence ("spring, mouse .... etc."). The effects of aging were especially marked when the sentences were grammatically complex (Gick, Craik, & Morris, 1988).

In summary, age difference are typically slight or non-existent in short-term memory tasks requiring verbatim reproduction of the material ("primary memory" tasks), whereas age differences can be substantial when the task requires manipulation and reorganization, or requires a division of attention between holding some items while dealing with further incoming material ("working memory" tasks).

However, rather than thinking of these two types of task as involving different stores or mechanisms, it seems preferable to think of a range of tasks requiring active manipulation to a greater or lesser degree.

Long-Term Memory

When memory proper is considered – that is, memory for events that took place minutes, hours, or years ago – the evidence again shows that age differences are much greater in some circumstances than in others. It is well established, for example, that performance on recognition tasks (in which the target item is re-presented along with new distractor items) declines only slightly with age, whereas age-related decrements are greater in cued recall tasks, and even greater in free recall tasks in which no explicit cues are provided (Craik, 1977; Salt-house, 1982; Craik & Rabinowitz, 1984). At first it might seem that the age difference is simply a function of task difficulty, with differences being slighter in the easier recognition task. This does not appear to be the case, however; in an experiment in which recognition was made more difficult than recall by manipulating list lengths and retention intervals, the age decrement remained greater on the recall task (Craik & McDowd, 1987). A straightforward account of this pattern of differences is that older people have particular problems with the *retrieval* of information, and that retrieval problems are greater in recall tasks than in recognition tasks (Burke & Light, 1981). An extension of this idea is that all memory processes, at both ecoding and retrieval, depend on a mixture of processes driven by external stimuli and processes initiated from within by the subject. Tasks like recognition memory receive a great deal of support from environmental stimuli, whereas recall tasks receive less environmental support and must therefore rely to a much greater extent on self-initiated mental operations. Craik (1983) suggested that age decrements are found in proportion to the amount of self-initiated processing that particular tasks demand; the point is similar to Hasher and Zacks's (1979) suggestion that age decrements increase as the task becomes more "effortful."

A number of recent studies yield results in line with the notion that age decrements in memory are least when the task receives good environmental support. As one example, Craik and Rabionwitz (1984) describe a study by Waddell and Rogoff in which the investigators

presented a set of 30 miniature items (animals, cars, people, etc.) either on model panorama consisting of roads, fields, and buildings, or in featureless cubicles, the task being to remember the spatial position of each miniature item. The subjects were middle-aged and older women. In the cubicle condition performance levels were 73 percent and 36 percent respectively, but in the more "supportive" panorama condition the levels were 75 percent and 72 percent respectively: that is, the greater amounts of environmental support in the panorama condition were differentially helpful to the older group.

Although the argument is still somewhat speculative it therefore seems possible to suggest that age-related differences in memory for specific events are the result of less efficient processing at both encoding and retrieval. At econding, older people do not form as deep, as distinctive, or as contextually specific representations of episodic events as do their younger counterparts, and at retrieval they are less efficient in reinstating the mental configuration that corresponds to remembering the original event (Craik, 1983). Both types of inefficiency can be overcome, however, by the provision of appropriate environmentally support.

Two other sets of observations also fit this general descriptive framework. The first set concerns "prospective memory" or remebering to do things at a future time – to phone a friend in an hour's time or to pick up milk on the way home from work, for example. Such prospective actions often have few environmental reminders associated with them and to that extent are liable to be forgotten. There is emerging evidence that older people are particularly vulnerable to this type of forgetting, although evidence that is both rigorous and "ecologically valid" is difficult to obtain. Anecdotally at least, people report that they are more likely to leave reminders for themselves – letters to mail placed by the door, notes by the telephone – as they get older. The second set of observations concerns so-called implicit memory tasks. Such tasks do not require the subject to recollect the original episode, but merely to carry out some operation like identifying a word flashed very briefly on a computer screen or completing a word presented with some of its letters missing (e.g. -SS-SS--). Performance on these tasks is enhanced if the target word (e.g. ASSASSIN) has been studied recently, and such priming effects do not appear to change with age. It is possible to argue that implicit

memory tasks require relatively small amounts of self-initiated or effortful processing – the solution is guided or driven by the stimulus array itself – and therefore slight age changes would be expected in light of the previously expressed theoretical framework.

Remote Memory

One of the stereotypes of aging and memory is that whereas memory for recent events becomes increasingly inefficient with advancing age, memory for the events of one's youth remain clear and easily recallable. This claim is often made by older people themselves, but it is open to a number of criticisms. One is that the early and recent events are rarely comparable in importance to the person in question. Thus, it does not make much sense to compare recent forgetting of what you had for breakfast yesterday with the memory of a salient event (a birthday party, death of a pet) from 70 years ago. A reacted difficulty is that early personal memories are almost always highly selective, and because they refer to interesting episodes in a person's life they are recounted quite often. Thus, the early memory is not really being retrieved after an interval of 70 years, but is probably a composite "memory" mixing details of the original with details added unconsciously over many retellings of the event. The psychologist Laird Cermak refers to these oft-told tales as "family folk memories"; their accuracy is usually impossibly to ascertain.

The results of more objective tests are somewhat mixed. One method is to contrast a questionnaire about public events stretching back several decades. The typical result of such studies is that memory for public events gets progressively poorer as the events recede in time from the present; there is no evidence for particularly good memory of those events that occurred in the person's youth. On the other hand, studies by Bahrick and his collaborators have shown very good retention of personally acquired knowledge, such as memory for the names and faces of high school colleagues or memory for the geographical layout of a person's college town, after 50 years and more. More research is required on this interesting but difficult problem.

Memory for Knowledge

So far this account of age changes in memory has dealt largely

with memory for episodic events. However, everyday memory is as often concerned with memory for factual knowledge, for Semantic Memory, and here the picture is more positive. In general, it seems that previously acquired knowledge is retained well by older people, although new knowledge may be more difficult to acquire. Several studies have demonstrated an increase in vocabulary and in general knowledge from youth to middle age, and only slight declines from middle age to older ages (Salthouse, 1982). One crucial factor may be the degree to which the information is *used* by the older person; practice at retrieving specific pieces of knowledge may serve to maintain its accessibility. Studies of expert knowledge in older chess and bridge players by Charness (1981) have shown that the older players retain their skills, although they may respond more slowly and have poorer memory for the specific details of a specific game. However, it would be difficult to argue that all types of semantic memory are immune to the effects of aging given the universal complaint of older people that they have increasing difficulty in remembering proper names.

A further aspect of general knowledge is a person's knowledge of his or her own memory and how to use it effectively. This information is referred to as "metamemory" – for example, which strategies to use, which events will be easy to remember, and which will be difficult. It seemed possible that some age changes in memory might be attributable to changes in metamemory, but so far the evidence is that changes in metamemorial knowledge are quite slight. Older people report more reliance on external cues and reminders, but this change in habits is a sensible response to an increase in everyday forgetfulness.

Underlying Causes

Given that there are substantial age-related decreases in the ability to remember at least some types of information, what gives rise to these deficits? It is likely that the final answer to this question will involve a complex mixture of biological and psychological factors (Poon, 1985). Several changes in the brain are known to accompany normal aging: there is neuronal loss, and neurons lose their connectivity; there is an increase in senile plaques and in neurofibrillary tangles; there are changes in neurotransmitter and neuroendocrine functions;

glucose utilization is less efficient. Any or all of these changes could be associated with cognitive changes, although at present the precise linkages are not well understood. At the psychological level, several researchers have suggested that a decline in "processing resources" accompanies the aging process and is related to a decline in cognitive functioning (Craik, 1983; Salthouse, 1982). These hypothesized processing resources may be linked to neuronal glucose metabolism. At a more practical level, it now seems clear that continued active involvement in social and intellectual pursuits plays a positive role in maintaining memory and other cognitive functions in the elderly.

Agnosia

Agnosia is a relatively rare clinical disorder, reflecting the failure by a patient (typically following a brain lesion) to recognize commonly occurring objects. It is a recognition rather than a naming disorder. Unlike anomic patients, who have problems in finding the appropriate names for objects, agnosic patients fail to show any recognition for objects they cannot name – for instance, they cannot describe or pantomime the object's use. Agnosic problems can be specific to objects presented in just one sensory modality – such as a failure to recognize only visually presented objects, or felt objects, or a failure only to recognize objects from their associated sounds. In each of these instances, it is important that the patients can be shown to have adequate sensory discrimination (e.g. visual acuity, tactile sensation, and hearing respectively) – so that agnosia reflects a recognition failure in the presence of intact sensation.

The first, and still influential, theoretical discussion of agnosia (specifically, visual agnosia) was provided by the German neurologist Lissauer in 1890. Lissauer made a distinction between an apperception process, concerned with generating a stable perceptual representation of a stimulus, and an association process, concerned with linking the perceptual representation with stored knowledge about the object's function and associations. According to this distinction, agnosia could result from an impairment of either the apperception or the association process – with apperceptive and associative agnosic patients being distinguished according to whether the patients show good performance on tests of perceptual processing, along with impaired recognition. For instance, in the visual modality, patients may be asked

to copy the objects they fail to recognize, to match the same object seen from two different views, or even to decide whether an object has a familiar structure (i.e. to judge whether they have seen the object before). Some patients can do well on all these tests – and yet still fail to recognize the object's function. This is consistent with the patient having a problem in associating perceptual information with stored knowledge of the object's function (e.g. Riddoch & Humphreys, 1987a). It further indicates that rich perceptual descriptions of stimuli can be derived without feedback based on functional and other forms of associative knowledge – since such patients fail to access functional knowledge and yet they can copy objects, match them across different views, and so on.

Other patients can fail on different aspects of the perceptual tests. Some have problems in copying (not linked to a motor disorder), others in matching objects seen in different viewpoints, and so forth. This suggests that we may need to distinguish different kinds of "apperceptive" problem in different patients (Humphreys & Riddoch, 1987a). In vision, a problem in early edge coding could result in poor copying, while a problem in integrating depth and surface information could produce problems in matching objects in different views. Yet other problems may be linked to the grouping processes that normally organize the world into coherent perceptual objects – coding parts of a scene as belonging to a single object if the parts are visually similar, and so forth (so that we may fail to identify an animal whose coat is similar in color or texture to the background under growth). Patients with impaired grouping procedures may tend to segment objects into separate parts, seeing each part as a distinct object. Distinctions between different types of apperceptive problem are also possible in the modalities, although the most detailed work has been carried out on visual agnosic patients. In general, studies of such patients provide a rich source of evidence for understanding the usually hidden processes underlying normal object recognition.

In some instances, patients can have especial problems recognizing particular types or categories of object – such as animals, fruits, and vegetables, or inanimate objects (e.g. Warrington & McCarthy, 1987; Silveri & Gainotti, 1988). For some patients the category-specific problems seem to reflect the close visual similarity between the objects within particular categories – with objects from

natural biological categories presenting the greatest problems, perhaps due to the similarity of the exemplars within many of these categories (Riddoch & Humphreys, 1987). However, patients with specific problems with other types of object (such as indoor objects) may have difficulties because of damage to the knowledge associated with the particular objects. These patients may be important for guiding theories concerned with the nature of the functional and associative knowledge mediating object recognition.

Amnesia

A general term covering any form of temporary or permanent memory loss. Amnesia can be divided broadly into psychogenic or functional states and organic states (Parkin, 1987). In psychogenic states, loss of memory has a psychiatric origin. In most cases the memory loss is related to an extremely negative life event such as bereavement. The most extreme and rare form is fugue, in which patients forget who they are and may even adopt a new identity. More common are selective hysterical amnesias in which people lose the ability to remember traumatic events; combat amnesia and amnesia for violent crime are the most commonly observed examples of this. Explanation of these disorders is difficult. It has been suggested that some hysterical amnesia is caused by emotional state dependency but in many cases an explanation based on a theory of repression seems the only course. In the case of amnesia for crime, the issue is further complicated because many violent crimes are committed under conditions of extreme intoxication resulting in consolidation failure, and there is the added problem of malingering.

Multiple personality is another and controversial form of psychogenic amnesia. Here the patient adopts a number of different personalities each of which denies conscious knowledge of the other. These personalities can often be used as an alternative way of relating negative life experiences without involving the normal self.

In organic amnesia the loss of memory is attributed to some form of brain dysfunction. Some forms of amnesia are transient in that the patient's memory returns to normal after a period of time. Post-traumatic amnesia following head injury and transient global amnesia (thought to be caused by temporary cerebro-vascular disruption) are the most common causes of temporary organic amnesia.

Epileptic seizures and electro-convulsive therepy also cause temporary amnesic states.

Permanent organic amnesia arises from a wide number of different causes including head injury, strokes, aneurysms, tumors, metabolic deficiencies, and malnutrition (Kapur, 1988). The irreversible nature of brain damage means that the patient's memory never returns to normal. This form of amnesia is often referred to as the amnesic syndrome because all patients suffering from it have certain unifying features, which are: intact immediate apprehension of information, normal or near normal intelligence, and intact language abilities. Amnesic patients are also known to have intact procedural memory. This term is rather imprecise but can be used to describe the memory processes underlying the acquisition of skills and other forms of knowledge that are not directly accessible to consciousness.

Patients with the amnesic syndrome have two major deficits: a severe anterograde amnesia which prevents them from acquiring any new knowledge, and retrograde amnesia, whereby the patient is unable to remember events and knowledge learned at a time prior to the brain injury that caused the amnesia. An interesting feature of retrograde amnesia is that it demonstrates a temporal gradient: memories formed early in life are more likely to survive than those formed during a later period.

Despite its wide variety of causes, organic amnesia, whether temporary or permanent, is caused by disruption to either one or two specific regions of the brain: a structure in the medial temporal lobe of the cortex known as the hippocampus and nuclei in the midline of the diencephalon, most notably the mamillary bodies. Diencephalic damage is most commonly found in Korsakoff's Syndrome – a result of brain damage due mainly to chronic alcoholism. Temporal lobe amnesia has more widespread origins including encephalitis, head injury, and stroke.

Patients with the amnesic syndrome have been studied intensively. The fact that some aspects of their memory remain intact while others are badly affected provides important information about the organization of memory. Preserved immediate apprehension supports the widely held distinction between short- and long-term storage processes. Intact procedural learning indicates the separate existence of a memory system concerned with the representation of

skills and related abilities although there is considerable doubt about whether there is a single procedural memory system.

Psychologists are less agreed on how to describe the deficit in amnesia. Some (e.g. Tulving, 1985) believe that amnesics have a selective loss of Episodic Memory with preserved Semantic Memory. However, others argue that this division is inconsistent with the evidence (e.g. Squire, 1987). In retrograde amnesia, for example, it has been shown that patients' general knowledge about the world is as disrupted as their ability to remember specific events. This has led Squire (1987) to propose that the amnesic deficit is best described as one of declarative memory. A third approach has been to define the amnesic deficit in terms of the characteristics of tasks on which they succeed or fail. Schacter (1987) has suggested that amnesics are poor on all tasks which require reference to a specific past experience – a function which he calls explicit memory. In contrast, amnesics perform quite well on tasks that do not require reference to a specific point in the past. This form of memory is known as implicit memory and it has become the focus of much interest in recent amnesia research.

There have been many attempts to explain human amnesia but as yet there is little progress or agreement among workers. However, one theory gaining reasonable support argues that amnesia represents a deficit in the ecoding of contextual information (Mayes, 1988). According to this theory amnetic patients are unable to encode those features of new events that are essential if those events are to be recalled at a subsequent point in time. In contrast, amnesic patients perform quite well on learning tasks which do not require the ecoding of contextual information. One problem, however, is that the evidence supporting this theory drives almost exclusively from patients with Korsakoff's Syndrome. There is evidence that this may represent only one form of amnesia and that amnesia arising from temporal lobe pathology may have a different origin (Parkin, 1987). This issue is in turn complicated by the discovery that certain features of amnesic learning performance arise from frontal lobe deficits that the unrelated to primary amnesia.

Absent-Mindedness

The term used to describe a lack of attentiveness to what is going on. Psychologists tend to study a particular form of absent-

mindedness known as "action slips" which involve the performance of actions that were not intended. As the term "abset-mindedness" suggests, attentional failures typically play a major role in producting action slips. However, a detailed analysis of action slips indicates that various other factors are involved.

Since action slips and other forms of absent-mindedness are difficult (though not impossible) to create under laboratory conditions, the usual method of investigating action slips is by asking several people to keep records of their own action slips. Norman (1981) and Reason and Mycielska (1982) carried out diary studies, with the latter investigators collecting a total of 625 action slips from 98 people. While it is possible to assign action slips to different categories, it would be unwarranted to attach much significance to the percentages of action slips falling into the various categories. One reason for this is that the figures are obviously based only on those action slips that were actually detected by those keeping diaries, and so we have no direct knowledge of those slips that were not noticed. Another reason is that the number of times a particular kind of action slip occurs is meaningful only in the context of the number of occasions on which that kind of action slip could potentially have curred.

It is possible to account for most action slips by drawing a distinction between two modes of control over motor performance (cf. Reason & Mycielska, 1982). A closed-loop or feedback mode of control is used during the early stages of motor learning. This involves a central processor or attentional system utilizing feedback of a visual and proprioceptive nature to provide moment-by-moment control of behavior. After prolonged practice, this closed-loop mode of control is more and more abandoned in favor of an open-loop mode of control in which behavior is controlled in a relatively automatic fashion by motor programs or by pre-arranged instruction sequences. In essence, action slips tend to occur when the open-loop mode of control is used at those points in behavior where use of the closed-loop mode of control is necessary for accurate performance.

One of the interesting characteristics of action slips is that they typically occur during the performance of activities which are highly practised and overlearned. This is somewhat surprising, because practice usually produces a substantial improvement in performance and a reduction in the tendency to make errors. Practice permits use

of the open-loop mode of control, which has the advantage of freeing attentional resources to engage in other processing activities. However, if the open-loop model of control is used too extensively, then errors will occur.

Analogies

Analogies are partial similarities, typically between rather different situations, that support further Inferences. More specifically, analogy has been described as a kind of similarity in which the same relational structure holds in different domains (Gentner, 1983). A good example is Rutherford's analogy between the atom and the solar system. We begin with the knowledge that the sun is *more massive than* the planet and *attracts* the planet, and that these two relations *cause* the plant to *revolve around* the sun. The analogy tells us that, given the correct object correspondences, a common relational structure will emerge. Indeed, if we map sun to nucleus and planet to electron and carry across the relational structure, we discover that the nucleus is *more massive than* the electron, that it *attracts* the electron, and that these *cause* the electron to *revolve around* the nucleus.

Psychologists study analogies for several reasons. First, analogies are important in learning. They are often used in explaining scientific concepts, such as the hydrogen atom, heat, or electricity. Once learned, they can serve as Menta Models for understanding the new domain (Gentner & Gentner, 1983). Second, analogy is important in Problem Solving. Gick and Holyoak gave subjects Duncker's radiation problem: how can one cure an inoperable tumor when enough radiation to kill the tumor would also kill the surrounding flesh? The solution is to converge on the tumor with several weak beams of radiation. Normally, only about 10 percent of the subjects discover this solution. If given a prior analogous story in which soldiers converged on a fort, however, three times as many subjects (about 30 percent) produced the correct answer, apparently transferring the convergence solution to the radiation problem. A final reason to study analogy and similarity is that they seem to underlie many other cognitive processes. Indeed, recent exemplar-based theories of conceptual structure in psychology and case-based reasoning models is artificial intelligence suggest that much of human reasoning and categorization is based on implicit or explicit analogies between the current situation and prior situations.

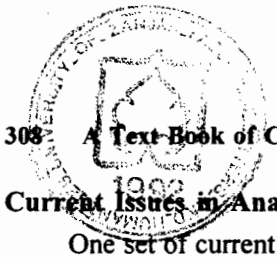
Psychological Approaches to Analogy

Although simple four-term analogies have long been used in intelligence testing, detailed modeling of the analogy process began fairly recently. Rumelhart and Abrahamson (1973), using a multidimensional space representation of knowledge, modeled analogy as a mapping from one subspace to another. They found that subjects given analogies like "Horse is to zebra as dog is to — ?" would choose the answer (e.g. fox) whose position relative to dog was the same as that of zebra relative to horse. Sternberg (1977) measured solution times to solve such four-term analogies as a way of studying component processes – encoding, inference, mapping, application, and response – and individual differences in their use.

Current approaches continue to treat analogy as a mapping. However, they generally use complex domain representations, such as Propostional Representations or Schemata. Gentner (1983) defines analogy as a "structure-mapping": a mapping of knowledge from one domain (the base) into another (the target), which conveys that a system of relations that holds among the base objects also holds among the target objects. Given such a match, any additional predicates that belong to the base system may be imported into the target as candidate inferences. Another prominent approach to analogy is Holyoak's (1985) pragmatic account. Holyoak defines analogy as similarity with respect to a goal, and suggests that, while structural mapping processes occur, they are oriented toward attainment of goal states.

Computational Approaches to Analogy

Artificial intelligence research on analogy has suggested computational principles applicable to human processing (e.g. Carbonell, 1981; Winston, 1982). Some current programs are explicitly intended as cognitive simulations of human analogical processing. Thus, for example, given two potential analogs, the Structure-Mapping Engine (SME) of Falkenhainer, Forbus, and Gentner (1986, 1989/90) first finds local matches, and then combines these into the maximal structurally consistent mapping and draws candidate inferences that should follow from the match. Connectionsit-style simulations have been developed, and M.H. Burstein's (1983) Carl simulates the use of multiple analogies to model a new domain.



Current Issues in Analogical Research

One set of current issues concerns the mapping process: How are the predicates that enter into an analogy selected, and what determines which matches are crucial and which can be ignored? How do contextual factors, such as the current goal state, interact with the intrinsic match in knowledge structures?

A second important question is how potential analogies are accessed in long-term memory. People often fail to access potentially useful analogs. In Gick and Holyoak's research, as discussed above, 30 percent of the subjects given the fortress convergence story spontaneously used it to solve the Duncker radiation problem. However, if they were told to use the story, between 70 and 80 percent could solve the problem. Ross (1984) demonstrated that, although people in a problem-solving task are often reminded of prior problems, these remindings are often based on surface similarity (e.g. between objects and story lines) rather than on structural similarities between the solution principles. Gentner and Landers (1985) tested a large set of stories and found that surface similarity was the best predictor of memory access, while similarity in relational structure was the best predictor of ratings of inferential soundness. However, Novick (1988) suggests that experts in mathematics are superior to novices both in making use of structurally similar prior mathematics problems and in resisting misleading surface similarities. Given the importance of analog in transfer, working out the determinants of analogical access is an important research problem.

Conclusion

Analogy is a particularly interesting cognitive mechanism, in that it can serve to import a complex system of interrelated knowledge from one domain to another. Recent research has advanced our knowledge of how people align representational structures and compute further inferences over them. The study of analogy leads us to deeper insight into the role of similarity in human thought.

❐❐❐