PHYSICS AND PHILOSOPHY OF THE MIND:

SOME NEW INSIGHTS IN AN EXTENDED NEURO-PROJECT

 — ISBN 0 9577737 1 4

Ondwelle Publications,
Melbourne

 — ISBN 0 9577737 1 4

PHYSICS AND PHILOSOPHY

OF THE MIND:

SOME NEW INSIGHTS IN AN EXTENDED NEURO-PROJECT

'Book B'
in a series on
Mechanisms of the Mind
(After the introductory 'Book A')

ROBERT R. TRAILL
BSc, BA, PhD

Ondwelle Publications — Melbourne 3193

 — ISBN 0 9577737 1 4

PHYSICS AND PHILOSOPHY OF THE MIND:

SOME NEW INSIGHTS IN AN EXTENDED NEURO-PROJECT

First published in December 2000 by Ondwelle Publications, Melbourne:
29 Arkaringa Crescent, Black Rock 3193, Victoria, Australia.

Digital version (with minor amendments in violet), February 2005.

Traill, Robert R.
Physics and philosophy of the mind:
some new insights in an extended neuro-project.

Bibliography.
Includes index
ISBN 0 9577737 1 4

1. Mind and body. 2. Brain. I. Title

128.2

Originally Printed in Melbourne

CONTENTS

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

PREFACE

> Something has been missing from our explanations
> of both mind and brain.
>
> We need to look for a more coherent overall explanation; but how? At this stage, not even well-funded experiments are likely to get much further — that is, not unless we first get some clearer idea about the brain's actual strategies.
>
> Without that, we don't really know what we should be looking for — so perhaps any well-based clearcut suggestion will make for progress. Even if such ideas turn out to be wrong, they should nevertheless give direction for future investigations — and especially if they really are based on sharply definable concepts, like *gene* or *atom* in other fields.
>
> At the other extreme, *pure philosophy on its own* is not likely to help much either. But if we borrow some of its eagerness to ask "what if?" — and then apply that to the tiresome details of information technology, then we might get somewhere. After all, we do know much of what the mind/brain can *perform* and what its basic "digital" ingredients are — so it should not be beyond us to "re-design" the mind-brain's most fundamental mechanisms in a plausible way.

Following on from an introductory *Book A* (1999), this present volume is *Book B* of a series to explore this mind/brain issue from as many angles as possible, from psychology to physics. This includes topics ranging from freewill and causality, to social issues such as research policies.

This present book sets out to establish what the basic mechanisms of the mind must be. One important clue is that the mind/brain is *one of several knowledge-gathering systems*, including the genetic-code system; and each of these others clearly depends on a one-dimensional digital coding-system. Moreover, if we stop to think about it, there are very good reasons why such a system should be one-dimensional — especially if a subsidiary two-dimensional system is *also* allowed as an extra. Moreover if the coding is indeed one-dimensional, it is difficult to see it as being anything other than a population of linear macromolecules like RNA!

The main trouble is that this seems to clash with current views which depend on the observable neural network of synaptic junctions and suchlike. That need not be a problem though, as these two systems are here seen as acting together as a symbiotic partnership. They would share some of the same physical structures, and also share the tasks needed within any non-trivial brain:

The conventional network can already explain several tasks adequately: • input and output interfaces with the environment (basic senses and muscle-control); • controlling parameters, including emotions (adjusting the "volume and brightness knobs" on the more intricate equipment); and • some types of pattern-recognition in two-or-more dimensions (like the now-common commercial use of "neural-net" software).

However the other system — the one supposedly based on one-dimensional digital coding — is needed to account for such human abilities as • detailed memory, • accurate *or* "intuitive" classification (including social skills), and • logical thought.

A second trouble caused by having to consider this new extra system is that its communicational needs would almost certainly be different. In fact it would probably require channels for *infra-red* signals: "IR", or "NIR" (Near-to-visible Infra-Red). However, as explained in Part II, the myelin insulation-sheaths around nerve-fibres would probably be suitable media to serve as fibre-optic channels for these signals. (This entails possible technical problems, and these are also examined). Meanwhile this need not compromise the other acknowledged function of these myelin sheaths.

Book C will digress into embryology, taking its cue from this question of IR signals. If IR signals are indeed routinely present within the myelin, certain other possibilities arise for the control of myelin geometry, as explained briefly in *Book A*. This seems to offer an explanation for a problem which has gone unsolved since Donaldson and Hoke (1905) — with considerable postwar activity until the early 1980s, involving researchers such as Friede and Bischhausen (1982). For the current project, the main value of this digression is that it offers unexpected corroboration for the proposed molecular mechanisms.

Book D will take us back to where the project started — trying to make sense of psychology in biological terms.

Acknowledgements

This has been a long project, and some of those to whom credit is due are regrettably no longer with us. I have already mentioned in *Book A* my indebtedness to I.A.Boyd, F.H.George, and Sherrill Hallmond. I am also indebted to the Department of History and Philosophy of Science, at Melbourne University for providing research facilities during a significant part of the project.

In addition I would like to acknowledge useful discussion or correspondence with (in alphabetical order): Don Campbell, June Factor, David Finkelstein, H.G.Furth, T.E.Gambling, H.A.Gebbie, D.O.Hebb, Mary Hesse, G.L.Mallen, Sarah Mallen, Peter M.Milner, J.Mittelstrass, A.G.S.Pask, G.Rossman, Bernard Scott, T.A.Sears, T.F.Slater, Christine Thomson, D.G. von Keyserlinck, and Norman Wetherick.

I am also most grateful to Bronwyn Oster for assistance and advice on design matters, and to Declan Traill for technical assistance regarding computer facilities.

R. R. Traill

Melbourne, November 2000

Any meaningful changes from the original printed book are coloured violet in this online version. The only important alterations have been (i) fixing an obvious units-error in footnote 146, and (ii) to re-name the short-cut titles for the various books in this series: Instead of "Books 0, 1, 2, 3" (with zero signifying "introductory"), these are now referred to as "Books A, B, C, D" respectively — because the zero turned out to be rather confusing.

R. R. Traill

Melbourne, February 2005

 — ISBN 0 9577737 1 4

PART I: ON SCIENCE, MIND, BRAIN AND MECHANISM

Each *brain* tries to map its own main environment. Likewise *science* sets out to map-or-model society's environment[1]. Since both share this same formal task of capturing knowledge about their respective environments, we might expect some formal similarity in their activity so that if we understand either one, that insight could then help us to understand the other. For that reason alone, we should find it instructive to study that process we call "scientific method" if our main goal is to understand the mind.

In any case we are obviously engaged in a scientific task, so this gives us a second reason for understanding the methods of science — including their shortcomings! In fact it is well known that scientific method has been drastically revised since the *Kuhnian-and-postmodern revolutions* of 1962-1978 which downplayed the previously-privileged role of observation and experiment. This rethink is laudable, and oft discussed in some circles. But talk is cheap, and it is not clear that this lip-service has translated into practice as often as one might have expected.

As a community we may *say* that we recognize that experiments are no better than the theories they are based upon; and we may even acknowledge that this theory-ladenness is unavoidable (yet still hidden away as somehow shameful). We may *say* all this, and yet still not know how to do any better in practice, and therefore we often continue to depend on the same old processes for assessing hypotheses.

For that reason, I spend some time in this Part I outlining some post-Kuhnian ideas of how scientific knowledge is actually acquired within society. (In this, experimentation still plays an important part, but no longer the *all*-important part). In brief, we can draw on principles of *self-consistency* within our existing belief-systems, and on principles of *free debate*. (Self-consistency is estimated by "coherence" criteria, see section 8.2; and the specialized free debate is identified as "hermeneutics" in section 8.4). Or if we are uneasy about seeking coherence, we can sometimes harness mathematics instead — though I would claim that this turns out to be a disguised way of doing the same thing: hence the popularity of maths for doing those essential jobs which orthodoxy forbade us to do openly.[2]

Meanwhile we have here a clear opportunity to demonstrate how the same principles appear to apply also within the mind/brain in *its* own task of private knowledge acquisition. In fact, as already suggested, the two cases help to illustrate each other.

We might note that even some physicists adopt this post-modern trend:

> **Interviewer** … *do you feel that there will be a stage in the future where it will be possible to do experiments to discriminate between these different interpretations?*
>
> **David Bohm** *I think there will be, but there won't be if you don't first consider these ideas seriously in the absence of experiments.*
>
> (Davies and Brown, 1986, p.134)

Ultimately the mind-model *may someday* come to be justified by experiment in the old tradition of the 1880-1960 period; but it clearly does not meet those criteria yet, despite some partial evidence in that direction. But then, without the theory given here, who would even think about doing the relevant experiments, which do not fit well in the existing paradigm? — And who would think of funding such otherwise-pointless investigations? They would, after all, be quite expensive.

[1] Knowledge is needed within the *individual brain*, and this "knowledge" consists of some sort of mapping or modelling of that world outside — an inner encapsulation which the mind should find instantly accessible (unlike the outside world itself which is obviously less cooperative). Likewise *society* also needs its own knowledge-store. It too must set up maps-or-models of *its* environment — models in the form of written-descriptions, films, formulae, laws, oral-traditions, working models, land-maps, wall-charts, and so on.

Just what mechanisms are needed to accomplish these tasks of capturing knowledge? This process may seem simple until we really look closely at the detail, but then we begin to appreciate some of the difficulties. (If you doubt this, try your hand at writing a computer program for any "simple" everyday task in an elementary *machine language*). Indeed it is the main task of this book to dissect such key processes of this knowledge-acquisition; and such studies go by the name of "epistemology" — the study of truth and knowledge, and what these terms really mean — though I am more concerned with the physical underpinnings than has been usual amongst most epistemologists.

Note that these processes do not necessarily require consciousness, nor are they restricted to humans. We should also beware of readymade concepts like "copy" and even "observe" — because we then need to spell out *exactly* what that entails, without begging the question. More on this later.

[2] I am referring mainly to mathematical models which reflect the actual structure of the real system being modelled. However as McCollum (2000) points out, mathematical modelling is all-too-often just number crunching without any good claim to represent the substructural makeup of reality. Such a distinction is also implicit in the bibliographical classification system used by Dutton and Starbuck (1971), with their use of qualitative and quantitative criteria.

 — ISBN 0 9577737 1 4

1. On Asking the Right Questions about the Mind

1.1 A scientific view of consciousness

> "Neuroscientists and others are at last plumbing one of the most profound mysteries of existence. But knowledge of the brain alone may not get them to the bottom of it."
>
> D.J.Chalmers (1995).

This article in *Scientific American* next points out that there are both *easy* problems and *hard* problems in trying to understand the brain. Unfortunately though, despite all the splendid work that neuroscientists have done, it seems that they have really got no further than the "easy" part: that is, the more obviously mechanical aspects of brain functioning like signal-processing and pattern-recognition. Meanwhile, such obvious and important human issues as consciousness, humour, cunning and reasoning seem totally beyond such analysis. These then are the "hard" problems which seem to baffle any serious attempt to solve them.[3]

It is indeed difficult to see how such features of the human "spirit" could possibly be embodied in any merely mechanical system; and I suspect that most bystanders are tacitly very sceptical that computer-like systems could ever truly be able to match human abilities in any comprehensive way. The point is that there is an existing credibility gap which we need to look at, regardless of whether the public doubt is fair or not. We need to be able to look beyond the admitted successes of the "easy problems" and take a sober look at our prospects for solving the "hard" problems of consciousness and mind.

Of course in any historical period, it is easy to be sceptical about the capabilities of its own contemporary technology, or of the biological processes which have to be assessed in these terms. — The clockwork technology of the 1600s and 1700s certainly had its admirers, but it was never to be taken seriously as a means for explaining the human spirit. Nevertheless it is significant that people like Descartes, Hume, and Kant were still *seeking explanations for mind*, even though the prospects continued to be poor.

Modern neural theory goes back to about the 1880s and 1890s, notably in the microscope investigations of Ramón y Cajal; but it was still very difficult to see how mere connections between nerve-like elements could possibly account for the more subjective "soul" problems of mind and consciousness. Today we obviously know much more about the details of these nerve fibres and their synaptic linkages; yet it is not at all obvious that we are significantly closer to explaining consciousness or subjective mind in these terms. Clearly these elements are involved in the brain's activities in various ways — but there is no agreed clear explanation which takes us beyond the "easy" problems into the "hard" domain.

1.2 Hidden causes of mind? — "supernatural" or what?

Let us look at three dubious current belief-systems which interact and, somewhat surprisingly, offer a formidable system of mutual self support for each other — not total support it is true, but enough pseudo-support to disguise the weaknesses in the overall picture.

(1) Firstly there is the view held by *the hard-science subculture* that living things are ultimately "just" machines, a view which finds plenty of support if you confine your investigation to hard-science evidence and pretend not to notice your own consciousness[4] and other such "spooky" psycho-mentalistic phenomena.

(2) Secondly there is *the human-spirit subculture* which ridicules the idea that human attributes could be mechanistic an any thoroughgoing sense. — Now since these two groups usually have to live together in the same world, they have to paper over their differences; and this is where we come back to:

(3) The third view. This is the tacit assumption held by most of *our human culture as a whole*, that there is a **dualism** which allows the first group to be right within its own lab-orientated domain, and yet simultaneously allows the second group to be right within the domains of the clinic, or religion, or the arts.

I see this compromise as a good way of keeping the peace, yet it is only a poor and ambiguous way of understanding the mind. But is there any viable alternative to this well-entrenched view? Let us look again at the two views within the dualism. On the one hand, (1), the **hard-science** accounts of the brain seem to make some important inroads into some aspects of the problem, calling upon a variety of mostly-compatible approaches. Indeed these hard-science accounts of *simple* nervous systems may well be near enough to the truth. However when it comes to explaining the extraordinary special features of the human mind/brain, almost all existing accounts are sadly deficient in various important aspects.

Nor is the outlook very encouraging! Most accounts treat the synaptic junction (or some similar aspect of nerve-cell functioning) as the basic element of memory and mental

[3] There are similar hard problems in physics too. — We still do not *really* understand the fundamental nature of subatomic particles, nor gravity, nor electric fields; and perhaps we never will, despite our formidable ability to calculate and predict their behaviour. — Does that matter? Arguably it is less important in physics if only because we *do* there have the mathematical models to fall back on — a luxury denied us in the complexity of most mind-problems.

[4] Here we have the paradox of the "subjectively thinking positivist"!

 — ISBN 0 9577737 1 4

activity. But these cell-properties seem to have none of the rigidly-reproducible coding potential that we have come to expect, at least in our commercial digital communications. Nor do they offer any clearcut and flexible means for coping with unforeseen structures of our environment — in one, two or three dimensions — surely a vital requirement for us to cope with this fast-moving real world?

But then on the other hand, (2) those who see the human **spirit** as something special that transcends matter altogether, are also left with an unpalatable dilemma:

> If our mental-spiritual attribute is immaterial, while the brain is material,
> *Does this mere bio-brain have any spiritual-mental role at all?*

• If not, then that seems to deny the possibility of a scientific explanation for mind altogether.

• If, on the other hand, the brain *does have some auxilliary mental role*, then how and where could it *interface* with the immaterial spiritual component? — That is a problem which troubled Descartes himself, the original main proponent of this dualistic model; (Descartes, 1649)[5].

1.3 Could we possibly reconcile these rival views?

In short then, *we are looking for an explanation of mind which does justice to the best aspects of both the "hard-science" and the "human spirit" traditions*. Unfortunately though, at present there seems to be no commonly accepted ground on which these two traditions could meet. Was there anything we could have done to remedy this during the past decades? Is there anything we can do *now*?

The present situation does now looks promising, as we shall see. But before that there was a formidable obstacle to such developments — namely the well-intentioned (but ultimately fallacious) procedures of "the scientific method" which have held sway throughout the "modernist period" (1840-1978 perhaps[6]). This question of scientific method is important in two ways. Firstly for the practical task of acquiring knowledge within society, but also as a possible analogy for how the brain itself might go about similar knowledge-acquisition tasks. So let us now look briefly at the methodological doctrines of this Modernist period:

[5] Such matters are discussed in Papanicolaou and Gunter (1987); and also in sections 9.4 to 9.5 below.

[6] Railways, telegraph, and the works of Faraday and Comte, all date from about 1840; and meanwhile the high noon of romanticism passed with the deaths of its post-revolutionary generation: Keats 1821, P.B.Shelley 1822, Byron 1824, Schubert 1828, Mary Shelley 1851, and Heine 1856 (though of course the trend continued as a somewhat secondary force). This then might be seen as the ultimate start of modernism and positivism. *Logical positivism* however comes later, mainly with the Vienna Circle in the 1920s, though its immediate roots probably lie in Bertrand Russell's work a decade earlier (Whitehead and Russell, 1910-1913).
1978 was the year in which A.J.Ayer admitted, on TV, that positivism's pretensions to infallibility were misguided — "nearly all of it was false" — though "the attitude was right", (Ayer 1978). Kuhn had published in 1962; but such messages take time to gain acceptance, even just at lip-service level.

 — ISBN 0 9577737 1 4

2. POSITIVISM — THE FAILED HOPE FOR MODERNIST SCIENCE

2.1 Where did we go astray in the pre-1978 period?

Certainty is a comfortable feeling, so we should not be surprised if people go to great lengths to find it; and positivism purports to exclude uncertainties (though unfortunately its rationale has *itself* now been exposed as uncertain, and therein lies its ultimate failure!) This word-based quest for certainty has its good and bad aspects. It is surely desirable to find more precise and concise ways of understanding the world; but as soon as we make a single definite mistake in any highly structured logic, we are likely to end up in trouble.

Just consider the structure of a typical logical argument in geometry or algebra. Originally these were attempts by the ancient Greeks to arrive at certainty, and of course they are of considerable value when properly applied; but then, when is this use *im*proper? I can see four ways in which this supposedly ideal rational procedure could occasionally mislead us, and perhaps they have other weaknesses as well. Let us look at these in turn, in the next four sections, *2.2* to *2.5*:[7]

2.2 Supposedly infallible logic — or "the doctrine of the holy language"[8]

These arguments include the tacit assumption that no respectable user of the system will make any logical mistakes — that any page of closely argued logic will be accurate and acceptable at face value. But note that such an argument is supposedly organized in a *single straight line*, and like any single route it is very vulnerable to any mishap along the way. In fact though, this single-path nature is an artefact foisted upon us by the ancient Greeks' preconceptions about perfection. As Mach pointed out in 1905, such theorems must originally have had a large scaffolding of corroborating evidence (such as straightforward experiment and practical knowhow), but this untidy scaffolding was later removed and the very thought that it had ever existed came to be suppressed as an inconvenient embarassment.[9]

Today we may think that such fallibility-in-logic is behind us, for "surely computers will always produce the correct logical conclusions from the given data, without any calculation errors?" Well maybe to some extent. It *is* possible to have highly reliable systems if the elements and rules are valid and stable enough. But note that such *validity* is ultimately an empirical matter on which programmers will have spent considerable effort to get it right (and program debugging is just another of those "scaffolding" procedures which we like to pretend isn't really there).

Then again the much-prized *stability* of procedures and symbols within a modern computer did not just happen out of the blue. In the early days of computers it was quickly realized that, whenever the user needed *stable reliability*, it was necessary to have some sort of double checking and correction procedure — preferably within the computer itself, but otherwise by a manual desk-check just to make sure! Such logic-checking is now standard practice within current computer systems, so the user does not have to give it a moment's thought — but that does not mean that such redundancy can simply be omitted.

Accompanying this, there has also been a move toward digitization: the use of switch-systems which are increasingly definite in their on/off-ness, with no room for half-measures. Ultimately these systems depend on feedback channels to achieve the desired stability, and note that feedback is yet another way in which the supposedly straight-line nature of the logic is circumvented.

In short then, the Aristotelian ideal of a pure linear logic turns out never to be strictly attainable. There are always side-loops at least; or to put it another way, real live logic is more of a *network or lattice* of interrelating inferences, allowing for alternative paths and corroboration, but making no real provision for any all-powerful authority to ensure fair play — not even for the rules of logic themselves, which are therefore ultimately only as valid as experience shows them to be!

This sort of argument had originally been clearly established by Gödel's famous theorem of 1931, (see Penrose's exposition of it in his 1989 book); but its indigestible implications were swept under the carpet until the acceptance of works like Kuhn's book of 1962. These then spelt the end of misplaced self-confidence and supposed-certainty. That was effectively the end of Modernism, and the start of a breaking down of the questionable logic structures — the "deconstruction" characteristic of *postmodernism*.[10] Perhaps it is now time to reconstruct anew; but it will help if we can

---

[7] (§*2.2*) Logic seen as God-given; (§*2.3*) "Seeing is believing"; (§*2.4*) Hierarchical axiom-approach; and (§*2.5*) Distrust of analogy.

[8] quoted from Gunter, in Papanicolaou and Gunter (1987), page 4.

[9] Similar pretensions are with us still, notably in the verbal antics that many a scientific journal will demand its authors to perform, as if to prove their objectivity — resulting in such linguistic absurdities as "*the present author*" instead of the simple pronoun *"I"*!

[10] It is difficult to find any concise overall account of postmodernism. Most texts seem narrowly overconcerned with some specific discipline such as literature, so they not only lack comprehensiveness, but they also call upon the reader to have a prior detailed knowledge of that discipline.
For texts directly relevant to our present discussion, see works opposing logical positivism: works such as Hanfling (1981) and even the provocatively titled early work: *The Metaphysics of Logical Positivism*, by Bergmann (1954 !). Against this, the orthodox views may be found in the anthology edited by Ayer (1959).

 — ISBN 0 9577737 1 4

better understand logic as it really works, and not just as Aristotelians saw it.

Note however that there are some *very special circumstances* in which strict traditional logic seems to work well — cases like modern computers where the subject-matter is simple enough, and where the logical side-loops have been well worked-over, so the system holds no surprises as long as troublesome innovators are kept in place — or at least there are no surprises for the accepted "experts". But these do have to be comparatively simple systems. Perhaps the main tragedy of the twentieth century is that it has tried to apply traditional "straight-line" logic and maths to *social* problems, without realizing how especially important *those inconvenient loops and networks* and the *variable-elements* all are for complex systems like society, or even just for the individual human. Modernism seemed so full of promise, but in many ways the twentieth century has not been a happy period. All too often this has been through the use of inappropriate tools — tricks which usually work on simple clearly-defined systems, but which cannot be trusted for complex biological or social systems, except as mere tentative suggestions.

2.3 *Experiment as king — or "the doctrine of the immaculate perception"*[11]

Paradoxically scientists seem to have been blessed with two conflicting trends: Mathematics and logic have been pretending to be God-given and thus *beyond* any need for experiment. But meanwhile the rest of science has been labouring under a supposedly rigid need ultimately to judge all hypotheses *by experiment and observation alone.*[12] That is an odd inconsistency.[13] We find experts asserting the value of experimentation; yet, in the spirit of "self-evident" mathematical dogma, they see no reason why they themselves should have to experimentally justify their own assertion.

Not that this should be seen as conscious hypocrisy. Whenever one's views seem so obvious and natural that one never has the slightest thought of challenging them, then of course it is easy to overlook any resulting inconsistencies, and just as easy to confidently attack the wrong target.

Thus we usually naively conceive observation as: *"a direct unprocessed input — a ready-made 'model' of the reality situation out there"* — something so "obvious" that we don't even need to think about it.

Small wonder then that, when we are trying our hardest to get a true picture of reality, we will fall back on what we think is a foolproof method, even at the cost of efficacy. If we deserve blame for that, it is only for being so incautious as to assume that we have thought of everything that could go wrong. But then once someone has demonstrated a fallacy (Gödel in this case?) and its implications have been spelt out (by Kuhn, Habermass, Hesse, Laudan, Piaget, and others) then the politics of the situation are perhaps due for a change; and the unconscious double-standard needs to be made quite explicit since such apparently-trivial fallacies can have a powerful effect.

More importantly though, this near-total reliance on observation begs the question of what the observation process is all about, and how it actually works. Now it's not surprising that the physicists and mathematicians who have dominated this question of methodology, have almost always taken the *observation process* at face value — as some God-given capability which simply "comes naturally" and does not need to be questioned. After all, *"If I see it, then it must be there!"* — That is an assumption which is usually true, but not always: Thus we now know enough about optical illusions, hallucinations, dreams and unconscious bias to know that we can never be exactly 100% sure about what we think we see.

Then again, the reverse case of *not* seeing is even more of a problem: *I cannot see it*, and therefore believe it to be "merely metaphysical"[14].

Of course for simple systems involving straightforward everyday objects, we adults (*though not infants*) probably will interpret things correctly most of the time; but such skills rapidly begin to fail us if we think we can accurately observe all the nuances of a *social* interaction. That is not to say that we should never try such observation or other assessments, but we should not blindly assume that this observation is always the best approach for gaining knowledge about a complex system like society. Nor can we be sure of *ever* getting reliable results on some issues by any method available to us.

To put it bluntly, most physicists and other "hard-scientists" would seem to be less-than-expert on matters such as perception and observation, so we should have some reservations about their pronouncements on such matters — just as we would need to be equally circumspect about accepting

---

[11] a second quote from Gunter, in Papanicolaou and Gunter (1987), page 4 again.

[12] There are other criteria which we might use — such matters as whether the ideas make sense, or whether they form a self-consistent whole. However under instrumentalist or behaviouristic policies of modernism, these important considerations were allowed to play no more than an "advisory" role.

[13] Actually it seems likely that this inconsistency was essential for scientific progress! Given that there was one fallacy, it was also necessary to have another fallacy to compensate for it. Thus, as the following discussion will show, *overall self-consistency* is perhaps *the* most important workable criterion for truth (even if it is fallible) — so if we try to suppress it by official policy, it will have to creep in the back door under some other guise. Hence one unacknowledged value of mathematics: — Whatever else we may say of mathematical systems, they do go out of their way to be self-consistent. So if we mathematize our ideas of a particular real system (thereby also unconsciously *asserting an analogy*, see section 2.5), we will be saying in effect that we have tested our model in additional ways which are not just experimental.

[14] See subsection 3.3(2), and section 4.1, for other comments on metaphysical issues. The word "metaphysics" itself is commonly misunderstood, and in our immediate context it is likely to be taken as a disparagement — unfortunately!

a psychologist's opinion about nuclear forces. Either interloper *might* be right, but we would do well to check their claims with special care.

In this present side-discussion within chapter 2 we have actually been concerned with how *science as a community* gains its knowledge, and the suggestion is that the reliance on experimental observations has been overly emphasised in the past. We have also seen that, as luck would have it, this technical side-issue is actually closely related to our main topic — how knowledge is gained (ultimately from scratch) *within the individual human*; so this present section may bear re-reading after considering all the brain-related issues below and in *Book C*.

There remain two other attitudes which have accompanied modernism and positivism in a less rigorous way, and perhaps they are not quite so important; but it might nevertheless pay us to keep them in mind when looking for new solutions.

2.4 *The feudal axiom-attitude — a dread of circular reasoning and other loops*

In its simplest form, *circular reasoning* is like this: "Proposition A supports proposition B, which in turn supports A again, and so on!" — A and B uphold each other, but neither has any really independent means of support. Neither of them has the authority of a fixed unalterable base on which to build.

I suggest that there are two grounds for attacking circular reasoning. The first criticism is *partly valid*, but can be profitably circumvented with some re-adjustment away from the purity of strict circularity. The second criticism may now be seen as misguided. Let us look at these in turn:

Firstly if the logic path is indeed a *single* loop, then the system is effectively "sterile" with no inbuilt dynamic to push it towards either truth or falsity as a verdict. There is no room for growth, nor any sign of self-organization in the past. The claimed relationship may or may not be valid, but there is no independent way of checking; and that can lead us astray if we take it for more than it is.

However almost any real-life system will have numerous causal loops associated with it as stabilizing feedback. So if we set up our abstract proposition such that it has no more than a single loop, we have probably overdone the purity of our abstraction — over-idealizing and oversimplifying so as to tidy away most of the "messiness" and interconnectedness of reality.

If we were now to add this messiness to our abstract model, it would no longer be the sterile single loop but rather an interlocking network of loops. We can then conveniently visualise this in three-dimensional space as a *globular network of feedback loops*. In real life, that sort of arrangement can only be stable if there is a measure of overall "cooperation" between the loops. (Speaking mathematically, this amounts to identifying this with a set of simultaneous equations, and indeed that is what we would get if this were a network of wiring as found in a complex-circuit problem.)

This extension built upon circular reasoning is what I tend to call "**spherical reasoning**" — as a lighthearted way of saying "beyond mere circular", though it would be slightly more correct to say "*network* spread over a sphere …", and even more correct to say "*a richly-connected self-stabilizing network, in as many dimensions as we need"!*

The point is this: Because of the extra constraints from the interconnection, it becomes possible to conclude more about the system *collectively* than was possible in the neutral single-loop case.[15] When we apply this to logic, it opens the way for us to use logic loops, and to do so constructively as long as we look beyond single loops and look for other loops which might collectively corroborate or contradict our original premise.

If the loops substantially *fail* to corroborate this premise, the logical argument can be said to be "false"; and of course we do need a logic-strategy which can deliver us both types of proposition-verdict, both true and false.

Note that there is no requirement that any of these loops (which purport to accord with reality) to be actual empirically-verified "fact". Indeed the very idea of *strict empirical verification* is nowadays seen as misconceived. Some of the claimed relationships may be relatively well founded experimentally, in which case this experimental justification should be represented by extra loops which proclaim this justification). Other loops may represent pure guesses which will hopefully survive if valid, but be eliminated as inconsistent whenever they are false.

This brings us to the second criticism of circular reasoning — on logical positivist grounds, which would apply just as much to the reformed "spherical reasoning" as to the plain simple circular case:

Traditionally all such circular-based arguments have been treated with ridicule since their chains of reasoning seemed to have "no visible means of support" — that is, no solid factual basis to underlie them. However what if we are forced to admit that there can never be such a thing as a completely reliable "solid factual basis" for anything, not even a "clearcut" observation? What then?

The Solar System has no truly fixed base either, no matter what Ptolomy *or* Copernicus may have thought, and yet it survives. In fact this circulating feedback is characteristic of self-organizing systems; and I would include: *living individuals, natural brains, societies, and social knowledge-systems* as all being examples of such self-organizing systems. None can really depend greatly on outside help (and any help they do receive is probably by virtue of their being part of a larger such system, such as a species, and *that*

[15] This collective approach to the system as a whole, and conclusions from the resulting holistic models, is very closely related to the concept of *"coherence"* which we will come to again in this section, and in sections 8.2 and 8.3 below. (It is also discussed in the previous volume, *"Book A"*.)
In this project, *"internal coherence"* is a preferred synonym for "coherence" — see footnote 77, on page 41.

larger system will ultimately have had to cope without outside help). So they must depend on the tricks of self-organization — notably through the feedback loops and redundant pathways discussed earlier; (page 4). These systems have much in common with some types of circular or "spherical" reasoning, as we have just seen; — systems with *many interlocking loops* giving stability to the overall ensemble; (page 6).

Of course other rival belief systems will also form their own similar "spheres" of self-supporting notions *including observational data*; and if we want to choose between these rivals, all we can do is assess the comparative overall tightness or *coherence* of their internal supports. If we then get it wrong, well — that's just tough luck! As long as we also try to restrain any undue bias, and to consider all aspects, that's the best we can do until further data comes in — but there is no foolproof independent arbiter, no matter how much we might wish to find one. That was the mistake made by Modernism, and it has been the scourge of the twentieth century.

(Simplistic Postmodernism within the humanities takes this argument to the opposite extreme, pessimistically asserting that there is no arbiter at all, and overlooking the helpful-yet-fallible selfconsistency criterion discussed here.)

As for the "experiment is king" policy; we should note that it is still commonly followed *in practice* — both in science itself and within its administration. In contrast, nearly all *philosophers* can now tell us why this view is wrong and too restrictive; but hard-headed administrators will pay little heed.

In fact though, such conservatism is not entirely misplaced; so we should at least understand their position: There is probably a real need to hold back the horde of "flat-earthers" who would flood the market with their pet ill-founded theories if they were let loose. We do need some check to unfettered speculation, and an insistence on experimental evidence is one way of doing this, even if it is inefficient or unjust.

A strict ban on debatable speculation would stifle all change, so we cannot swallow this argument uncritically. But no doubt hard-headed administrators will continue with such thoughts unless the imported advice also happens to suit some pre-existing political agenda.

The main practical objection to philosophy is its predilection for *abstract concepts* which seem to defy any of our attempt to embody them in material terms. One major task of this present book is precisely to remedy this shortfall — to seek plausible physical embodiments for philosophical or psychological abstractions, and see how they stand up to such "de-wafflizing". More to the point, this is done in the present book as a prelude to *Book D* in which some of the confusing abstractions of current philosophy will then be re-examined in the light of this new sharper focus.

2.5 Distrust of analogy

Analogy is another much maligned approach for scientific argument, and yet every mathematical formula embodies an analogy. In effect:

> *This real system behaves in an analogous way to what I achieve artificially whenever I do 'these calculations' — (i.e. these esoteric repeatable actions with my pencil and paper). But note that I also necessarily 'interpret' my written version in a more-or-less subjective way at the start and end of this modelling process.*

This analogy process is even more powerful when the same general formula turns out to be applicable to several apparently quite-different systems (like the inverse square law, or those very general Hamiltonian equations which apply conservation of energy to many different problems).

In short, analogy too has its place; but we should be aware of its strengths and limitations. Perhaps the most important thing to look for is a formal similarity in substructure (preferably in its fundamental elements or fields, if any) such that equivalent identifiable inputs in each microsystem will produce equivalent identifiable outputs in the matching microsystems. Perhaps that alone will not guarantee a useful analogue, but it would be a good start.

Probably the main danger here is in claiming a meaningful analogy prematurely on the basis of mere superficial resemblance, without proper attention to substructure and dynamics. For complex systems like society, this will be a grey area because there we will never be able to gain a *full* understanding of any such substructure; and even if we did, our chances of finding an informative faithful analogue would be quite remote. Instead we just have to make the best of various partial analogies, and remain aware of their limitations.

 — ISBN 0 9577737 1 4

3. HARD-SCIENCE VERSUS THE ARTS

3.1 *Hard-science — as criticized within the humanities*

These criticisms may perhaps fall under three headings: (a) Hard-science as unduly *tidy* — and therefore inhuman. (b) Disillusionment, politics, power, and fashion. (c) Distrust of science's brain-models whenever they are seen as *machines without freewill*.

Let us look at these in turn:

(a) "Hard-science is frequently obsessed with *tidiness* rather than humanity"

As a physicist or engineer, it is easy to get enraptured with the precision and clarity of your own *hard-science* approach without realizing the pitfalls which are all too obvious to others. In the real social world, that clearcut picture is muddied by such messy and unpredictable things as humans and their emotions!

Mind you, hard scientist or bureaucrats are not like that all the time. But that tends to be their fall-back position; so that in an emergency, or when called upon to report to their peers in a "responsible professional manner", they will often have little option but to conform to the hard lines of their own established subculture — especially as they would be hard put to it to argue in other terms anyhow. Thus even within academic psychology (which has pretentions to the respectability of the hard sciences) there has often been a regrettable tendency to ignore or deny all subjective phenomena as "unscientific" and not worthy of serious consideration — a view taken to extremes, and quite explicitly, by the late Professor B.F.Skinner and other prominent behaviourists.

Of course to many other people, and especially those in the humanities, this view has long seemed manifestly absurd; though in the past they may not have had the courage of their convictions (nor perhaps the vocabulary) to pass judgement openly. If they had, their words which might have been something like this:

> *"Skinnerians can talk their 'logic' at me till Hell freezes over, but I know that I have consciousness and subjective feelings whatever they may say".*

Or perhaps less obviously:

> *"Economic rationalists like Milton Friedman may have perfectly tidy equations for their trade-balance or whatever; but what's the good of that if those equations cannot properly capture the important issues of life?"*

In other words, there is little sign that the hard-sciences (on their own) have anything much to say about how Tom or Mary's mind tackles messy ethical problems, or deals with those other human issues found in any competent novel. — Meanwhile Tom and Mary may frankly find that *supernatural explanations offer a better fit* for their messy real-life needs, even if they are not quite happy with such occult accounts either.

In other words, arts-orientated people and others will tend to see engineers, economists, physicists, brain-researchers and other "hard-scientists" as being *cold and heartless mechanists* (and hence assumed to be evil, greedy, careless, or incompetent). This charge sometimes has some basis, though it would be more accurate to say that often hard-scientists just do not adequately understand what emotional life is all about, nor even its merely biological necessity. Either they just cannot see its relevance at all, or they don't know how to handle it professionally (and they are hardly alone in this). So they simply *ignore* the "impossible" human emotional reality at considerable peril to society and its wellbeing.

(b) Disillusionment, politics, power, and fashion

Advertising hype is often used to build our expectations. If these hopes are not fulfilled (or if the "cure" is actually worse than the disease), then sooner or later public opinion will swing the other way. After World War II, great things were expected of technology including endless cheap energy from atomic reactors; but as we now know, such confidence was somewhat misplaced, and the end of that dream helped to mark the end of modernism — a transition which went along with student radicalism of the 1960s, the decline of logical positivism, and other interrelated features of the postwar period.

(This change of attitude to *technology* need not necessarily have applied to *pure science*; but as the lay citizen has not been encouraged to see any difference between the two, we must expect that both science and technology will normally be tarred with the same brush — so both could be expected to lose prestige together.[16])

Taking a historical view, we might have known that this fall in popularity could happen. Anti-scientific feeling does recur from time to time, perhaps like other fashions. In ancient times, the burning of the library in Alexandria was no mere accident; and Dhombres (1988) cites similar sentiments after the French revolution.

Then again, German science-and-technology of the 1920s (together with rationality in general) had fallen very considerably from its pre-war status — not because Germany was

[16] Pockley (1988) casts some doubt as to whether there really had been such a swing in public opinion in recent times. He claims that (in Australia at least) this supposed swing was merely asserted without evidence by a government Minister for Science in 1974 and then "gained credence simply by its frequent repetition". Of course it might have been true whether any formal evidence was there or not; and then again such an official pronouncement can easily become a self-fulfilling prophecy.

doing bad science, but rather as a consequence of unexpected loss and humiliation in the war, (Forman, 1971). This period saw a *growth of irrationality* as an understandable but dangerous response to a perceived failure of the rational. Unfortunately we may also note some similar trends in recent times — developments such as the growth in fundamentalism and the more extreme excesses of deconstructionism.

Of course if our old rational-framework really does conspicuously fail us, we might have thought it better *not* to descend to unpredictably treacherous irrationality, but rather to seek out a new and improved system of rationality. But, alas, that may be easier said than done — in the short term at least.

(c) Distrust of brain-models whenever they are seen as *machines without freewill*

This question entails some tricky philosophy, and we shall return to it in chapter *6*. Meanwhile however, we can sidestep the debate itself and merely note what the two main schools of thought are, and who belongs to them.

The mechanistically minded hard-science people are placed in an uncomfortable ambiguous situation here. They are still human after all, and they will doubtless place some value on their own feelings of freewill even if they do not explicitly recognize they are doing so. On the other hand they have a professional committment to the more-or-less rigid predictability of formal physico-chemical systems (which seems to preclude freewill in those systems), and they are also inclined to see biological systems as being built up from those same physico-chemical systems.

They may not all feel compelled to opt against freewill when we force them to choose, but they may well betray their ambivalence in much that they say and do. At any rate, their colleagues in the humanities are likely to *see* them as being faint-hearted or downright treacherous on the important question of freewill. And occasional extremist movements like behaviourism do little to help reconciliation!

Meanwhile the humanities people themselves will feel no comparable ambiguity about their own situation. They do see *freewill* as important, but they have no such strong committment to *rigid predictability within systems* — not even physico-chemical systems, which are usually well outside their area of interest anyhow.

In short then, people in the arts and humanities see the hard-scientists as somehow sabotaging the notion of freewill — either as advocates or as bearers of unwelcome news. This may be unfair to the hard-scientists themselves; though perhaps we should remember that hard-science has often been placed within the power of well-funded corporations and lobby-groups who *will* usually have a vested interest in minimizing the freewill of others! The distrust then may be justified in this respect. If so, it should perhaps be recognized explicitly as such, but not applied blanket-like to other aspects of the rivalry.

3.2 *Hard-science and its usual world-view*

Before looking at this particular set of attitudes, let us first consider the more general world-view that the typical hard-scientist will tend to have. Of course this is only an informal subjective assessment, but it may be of some crude help in our attempted reconciliation of views.

Maybe it will help if we list likely characteristics, and then try to decide which are most fundamental to their world view, and therefore the main driving force when they evaluate any rivals. Here are some suggested attributes, both good, bad and neutral:

(1) Failure to understand emotions, both as actual human needs and as often-valuable survival strategies evolved naturally over countless generations;

(2) Distrust of emotion-based decisions in *any* context;

(3) Preference for the explicit and structured rather than the vague or poetic;

(4) A heightened fear of losing face by producing a clearly-false discovery or theory;

(5) A readiness to attack other colleagues who seem to have taken any such wrong turn;

(6) Near-obsession with the need to find causes, (that is to say *structured* causes);

(7) A curiosity about substructure, and hence about organizational hierarchies;

(8) Distrust of supernatural explanations (invoking anything deemed to be outside normal space, time, or conservation laws);

(9) Pro-experiment;

(10) Anti-theory unless it is extensively mathematized;

(11) A rather blind faith in the power and relevance of mathematics.[17]

(12) …(see below).

This is not the place to discuss such suggestions in any detail, but as a montage they may serve our immediate purpose. The eleven items are not necessarily independent, and indeed you will often be able to see the train of thought as we move down the list.

What else could we add to this list of hard-science attitudes? *"Mechanistic approach"* perhaps? This is often used as a disparaging label for the hard sciences, but I prefer to see it as a secondary derivative, emerging out of various

[17] Krips (1996) and Beller (1992) discuss some of the sociological reasons behind these latter three attitudes in Western science (and how Soviet science took a different turn in response to the same forces). Thus *"This 'monocracy' … did not indicate a generalized acceptance of Bohr's ideas by physicists but rather a rhetorical and political exercise in the management of dissent"*.

 — ISBN 0 9577737 1 4

other items when taken together — especially items *3, 6-9* (pro 'obviously-real' structure), and *11* (pro-maths).

Given the biological bias of this book, it is also quite important to note the significance of item *7*, referring to hierarchies and substructure — a feature which has only blossomed comparatively recently in this context, and which arguably makes a major change to the debate as compared with pre-war philosophy: Thus *extended* and *in-depth* organizational hierarchies enable us to shift any blatantly mechanistic issues down through several layers of substructure — away from our world of whole-body experience where hard-science is badly out of place, down through cells, and further still until we reach the atomic-and-molecular level where mechanistic physics and chemistry *do* make proper sense at last.[18]

In other words if we are compelled to accept an ultimate mechanistic base even for biology and society, then maybe we can (to put it in psycho-computer terminology) compartmentalize such mechanism away from human values by burying it deeply within *subprograms within subprograms*, embedded within yet further subprograms, and so on. Note that this is indeed the way most present-day computers are programmed, starting at the bottom with very mechanistic switching, then "micro-programming", then "machine-language", and so on up into human orientated "user-friendly" interfaces. If we did not know in advance about this connection, we would find it difficult to guess at it because the two ends of the hierarchy are so different in their manifest properties.

[18] If we have to reconcile some ultimate mechanism (probably deterministic) with apparently non-mechanist human qualities, then a long "meta-" chain like this will probably be the most painless way of linking them in the same system whilst keeping them at ample arms-length conceptually.

Here we are talking about hidden *linkages* between structures, without necessarily saying anything about the *direction of hierarchical control*. Nevertheless our present knowledge of info-technology remind us that "*small is beautiful*" — that ultra-miniaturization of contol units is often a definite advantage (allowing for speed, compactness, economy of construction, and an access to strictly digitizable substrates). On average then, we might expect that any powerful controllers will often be found at the ultra-micro level, while the "muscle-exerting slaves" will tend to be much bigger. But the conceptual distancing would seem valid regardless of this directionality.

Of course the pre-industrial view was quite different, often seeing control as belonging intrinsically to the large and *physically* powerful — overlooking any control centres underlying this power except as some God-given charisma that the king-like leader might possess. Accordingly, the medieval and renaissance idea of arms-length separation of barely compatible ideas was to place one end of the chain on Earth, and the other in Heaven — way up amongst the unreachable and ineffable stars.

Finally we should note that there was previously no clearly formulated view that brain/mind systems might have anything more than the two control-levels of master and slave, as in Descartes' dualism — though we do find multi-level notions in social organization, and later in mathematics (especially in relation to Gödel's theorem). As far as I know, Wiener (1948) and Ashby (1952) were the first to introduce the *multi-level* hierarchical-control idea explicitly into *brain theory*.

The above list is perhaps fairly obvious in what it contains, (except perhaps the parts concerned with internal politics, items *4* and *5*). Let us now add another less obvious characteristic as suggested by Pask and Scott (1972).

Their study identified two main cognitive styles which had to be matched (for both teacher and student) before efficient learning could take place in the classroom. For the most part, according to Pask and Scott, humanities-oriented people are *"holist"* thinkers[19]. In contrast though, engineering-orientated types were:

(12) *"Serialist"* thinkers — preferring to confine their attention narrowly to comparatively precise relationships within isolated parts of their topic-of-study, but linking them in a linear sequence (as in language or in a mathematical proof).

It would seem likely that both these skills are necessary; and in a complex society like our own, we may well end up in dire straits if we cannot coordinate the two approaches successfully. Viewed superficially, the need for a holist overview seems to be the more obvious. After all we live in a world where real phenomena are often profoundly interconnected, so that we will be very poorly equipped if we can only contemplate supposedly-disconnected parts of that whole. Moreover the associated artistic imagination and flights of fantasy associated with this style enable us to find new solutions and prepare ourselves for the unexpected; — and note that games and play amongst children likewise fulfil this same vital role.

On the other hand, serialists evidently see themselves as the serious-minded workers who keep to the point and do not play round with irrelevancies or wild-goose-chases. Such an approach may be fine when designing a machine or arguing a point of law, but in real life we cannot always tell what is relevant to what, especially as we cannot possibly know or attend to everything at once. So it will be quite misguided for serialists to expect anyone to keep to the point in all circumstances.

Yet it is the natural tendency of serialists to do just that: — expecting a strictly logical approach to all professional matters, and meanwhile glossing over any aspects which do not fit into this mould. In past times especially, they would not have seen any sense in the "chaotic" word battles of their humanities colleagues (and worse still within their own ranks!) Today they may have a better in-principle appreciation that their own premises may not necessarily be as firm as they would wish; they may even have some practical appreciation of how to conduct a hermeneutic discussion over vaguely perceived notions; but this is not their natural territory and, when pressed for a quick answer, they will retreat to the strict formulations which they are at home with. Moreover, some of their fears have some real basis to them:

[19] Moreover, some holist thinkers prefer to take in additional *redundant* embellishment-information about a problem, as if to lend colour to it.

 — ISBN 0 9577737 1 4

Emotional and other inborn biases do certainly affect our perceptions and evaluations, and sometimes do clearly lead us astray. For this reason, hard-science policy has often sought to remove *all traces* of such subjectivity from its procedures. This blanket *reject-all* policy is, I suggest, an unfortunate over-reaction which well and truly throws the baby out with the bathwater. Moreover any such anti-subjective policy is bound to fail if it is rigorously pursued, because such non-logical processes are actually essential for knowledge-acquisition[20]; — and if we think we can manage without them, this just means that we have not yet identified all the ways they impinge on us, either as individuals, or as a species in which the encoding for such biases and strategies are inheritable.

Natural selection has admittedly given us some powers of logical thinking, but it has also given us much more powerful biases and emotional tendencies. We may well disparage these latter as *Animal instincts* or *Original sin* or *Human weakness*, but often they offer quite good survival tactics, at least for anyone living in pre-civilization societies.

In infancy before we can even talk, we will normally learn the concept of what an object is, *using certain perceptual biases*, even though the "logic" of this process is questionable. And yet it works! Then again, if someone tries to belittle us materially or socially, we have an inbuilt tendency to respond instinctively with *anger* and whatever flows from that — something that a robot or a worm would not necessarily do.

Of course the trouble is that these rules of thumb, like any others, were only *statistically valid* in the first place; and secondly they can sometimes become spectacularly inappropriate *when the circumstances alter*, as when human communities become cities, and when they then industrialize. Concerning the two above examples: Industrialization meshes quite well with our *object-discovery* strategy in infancy — but it does not fit too well with our instincts toward *anger and revenge*, which can now escalate out of control, as the twentieth century has amply and horrifyingly illustrated many times! Logically minded observers are understandably appalled at the folly of the latter, but they are generally not even aware of their inescapable debt to the former object-invention tactics of their infancy.

Science too is in the same ambivalent position, though usually the consequences will be somewhat less extreme. Mistakes can be embarassing or expensive, so it would be nice to think we could always avoid them by appropriate logical planning. In real life this ideal can only be partially attained, so it is a tricky question to decide in advance which ideas are wrong.[21]

In the history of science there have been honest mistakes, like the supposed discovery of *N-rays* by Blondlot; (see Eysenck 1965, pp.128-9; and Wallace, 1996). These form part of the folklore of science, but I suspect this mainly presents as a *post hoc* rationalization of later developments. Of course, rather paradoxically, emotions play their part here, in the politics of both sides!

More to the point though, there have also been many cases of politico-social interference — as with Galileo versus the Inquisition, Darwin against church fundamentalism, or Soviet geneticists versus Stalin. Such traumatic experiences have understandably led the majority of scientists to distrust religious doctrines in general, and of course dictatorships as well. The main trouble is though, that this tends to lead (irrationally and emotionally) to *guilt by association* — a rejection of anything deemed to be supernatural or even just *potentially-explicable in supernatural terms*.

That immediately impacts on studies of the *mind* (rather than *brain*). The mind has so far defied any serious attempts to explain it except those couched in supernatural terms. Hence "mind" became a dirty word among scientists, and even many psychologists tried to expunge the concept from their work — though that still did not stop some other scientists from using the word "psychologising" as a term of disparagement. Yet this witch-hunt against the intangible is somewhat misplaced.

Thus, while Skinner and Watson[22] explicitly *denied mind* and others chose to *ignore mind*, this simply admitted defeat, and paradoxically left the field in possession of their rivals — the supernaturally-minded. Surely that is counter-productive in scientific terms, and it betrays a supposedly inadmissible emotional response amongst scientists — a response which might bear re-examination, both by the hard-scientists themselves and by those in the humanities!

In summary then, the hard-science outlook has been to place great reliance on direct observation and on logic, but without much effort to understand the heuristic methodology of "arts people" in the humanities, and mostly without any inkling of how much their own thought processes were ultimately dependent on those very same heuristic principles. These were simply accepted uncritically as so obviously self-evident that they did not need comment!

Nevertheless hard-scientists did have some major grievances here, to which we now turn:

---

[20] As we shall see, both coherence-testing and hermeneutic trial-and-error play a much greater role here than was formerly realized. See Hesse (1980), Traill (1999) and, e.g., chapter 5 below.

[21] This is also made much more complicated because of ambiguities in key words like "wrong" and even "truth" — (ambiguities which, as we shall see, are probably inherent in the epistemological processes which we are actually constrained to use in any real world). Does "wrong" mean (i) giving us a false idea of what systems are like? — Or does it mean (ii) leading to undesirable consequences within society? — And that in turn raises the question of what one means by the term "society". All too often this term probably does just mean *the powers that be* within that society.

[22] J.B.Watson (1878-1958), behaviourist psychologist and pioneer of psychological advertising — not to be confused with J.D.Watson, the DNA theorist (of *Watson and Crick* fame).

 — ISBN 0 9577737 1 4

3.3 *The humanities-and-arts — as criticized within the hard-sciences*

(1) "The Arts-faculty is a talking-shop which never actually gets anywhere!"

Of course that overstates the case, though there is some element of truth to it, as we shall see.

To start with, we should clear up a minor confusion between "Humanities" and "Arts" (in the academic sense). The first refers to *subject matter*, while the other invokes a *methodological approach for dealing with* subject matter. Commonly the two go together quite happily in practice; but here we are dealing with theoretical matters, so some caution is called for.

Thus sociology is a subject matter within the humanities — a subject-matter to which we may apply this-or-that method. In principle we could try just using the methods of hard-science on their own, and regretably that is by no means unknown; but the results of such "social engineering" are usually disastrous, even if that does not hurt the perpetrators themselves. Alternatively we might use just discursive methods: committees at their round-table discussions, polemical debates, or hermeneutic processes over wine-glasses or coffee.

That discursive approach is something which many dedicated hard-scientists really do not understand even though they may indulge in it themselves — usually seeing it as mere recreation[23] or a time-wasting beaurocratic necessity, or as a means of correcting their "idiot" rivals! But let us consider what it can achieve when it works well.

Many real-life problems defy any clear structural formulation (at least for the present, and maybe forever); but life must go on, and meanwhile we often have to manage in some other "less-scientific" way. So using whatever vague concepts we can find, we construct a tentative pseudo-model of reality; and meanwhile we take measures which seem likely to improve this makeshift model: We keep trying to juggle, interchange, replace, and re-arrange those vague concepts in some sort of trial-and-error process such as may take place in a constructive discussion group.[24]

[23] Of course psychologists will recognize that recreation or "play" is not just a meaningless indulgence. Instead it plays a vital cognitive role for the individual. While our immediate discussion here is more concerned with society rather than the individual, there is nevertheless a clear connection which many hard-scientists would fail to notice.

[24] As we will see in "(2)" below, the resulting model of reality will obviously be imperfect, and it may well keep changing; but it may be the best we can get in present circumstances. Even if our model is pretty hopeless, it may arguably be better than no model at all, which would leave us completely without any basis for making certain decisions. But in more fortunate cases it may well steer us toward those more structured "scientific" formulations so prized within hard-science. (Of course such formulations may not be possible in our chosen area: They did work for alchemy, converting it ultimately to modern chemistry; but it is most unlikely to work rigorously for sociology, if only for the reasons given in footnote 27, (page 13).

This is all rather like trying to put together a largish jigsaw puzzle where most of the pieces are fuzzy and where we can never be quite certain which pieces belong to this particular set. In this difficult task, it may sometimes help to discuss the options within a group of colleagues. Note that as long as the fuzziness remains (and it may well do so forever in the case of the humanities) we can never quite believe in our model. But if that is as good as we can obtain, we will have no option but to make the best of it, and maybe even learn to enjoy the ongoing challenge.

If hard-scientists criticise this as a mere "talking-shop", they may be missing the point and not seeing the true nature of the problem. However if they then go on to claim that this never gets anywhere, they may *sometimes* have a point: Of course if a problem really is insoluble, then we cannot truly expect our discussion to get anywhere further; but then maybe there are occasions where we tacitly give up too easily, and *miss the occasional opportunity* to remove some of the fuzziness. Such a step would have turned the problem into a proper clearcut jigsaw puzzle which the hard-scientists would then have acknowledged. — Nay, more! — They might even have claimed it as their own, just as they did with chemistry!

Arts people do usually "enjoy the ongoing challenge" of endless discussion with fuzzy concepts. They often do it well, and it is a necessary skill in this world where many problems can never be realistically tackled by any other means. But then not all of their problems are quite as intractable as that. Obviously they may be blind to the possibilities offered within other disciplines, but then they may also be ill-equipped to deal with the sharper concepts which would then emerge. In short we may have something of a demarcation problem here — or perhaps a reluctance to even contemplate that ideas might "grow up", leave the parental fold, and move to another area where less-discursive skills prevail.

There is no danger that genuine Politics or Cultural studies will ever follow that route into the hard-sciences (though dictators like Stalin may delude themselves that they have achieved this extreme "neatness"). However there are many aspects of the life-sciences which have indeed been acquiring hard-concepts ever since the Darwinian revolution of 1859 — just as physics or "natural science" had done earlier, thanks to Galileo and others.

What about philosophers and their endless discussions of *the Mind?* That of course is the main topic in these several books which do attempt to change this "getting nowhere" verdict on this very issue of the mind/brain. Philosophers generally seem quite content to continue their discussions using fuzzy concepts — seeing no future in trying to identify any of their ideas with any feasible biological structures truly capable of transcending the fuzziness.[25]

[25] In recent times, the unit of brain-activity commonly accepted by hard-science has been the synaptic junction. When looked at realistically (rather than as an ideal switch input) this inconveniently remains seriously fuzzy. If there had been

 — ISBN 0 9577737 1 4

Why this reticence amongst philosophers? Apart from their own inclinations mentioned above, we might also consider the disdain which was likely to be heaped upon them for daring to venture into the domain of hard-science even tentatively — *even if* these philosophers happened to be well-informed on hard-science matters. Moreover this disdain or ridicule would have been particularly likely (from both sides) whilst the policies of positivism and overriding empiricism were at their height — from about the 1930s to the 1970s.

So on this *Mind issue* at least, the charge of *"the talking-shop which never actually gets anywhere"* does have some actual substance to it. However I suggest that here the problem is social rather than inevitable, with both sides sharing the blame; and that a proper mutual understanding and cooperation between the Arts and the Hard-sciences could work wonders!

(2) "Dangerous inaccuracy in the arts; — and wrong decisions which are difficult to fix."

Heuristic trial-and-error principles will inevitably produce misleading results from time to time, and these may then be very difficult to revise once they become embedded in the culture, even when their original role has become obsolete. Many scientists have been made painfully aware of this in the past, and that partly explains the rise of positivism with its ideal of accepting only demonstrable facts — and its readiness to assume that any supposed-knowledge which could not be achieved by this means was either not worth knowing, or it was a downright illusion.

This policy seemed to work well in mainstream physics and other such fields, though not in others. However it is now evident that its rationale is faulty, offering an impossibly stringent test standard which, *fortunately*, is routinely breached by all productive scientists! This is usually done unconsciously though, and therefore in an unsystematic way — unnecessarily hindering progress in some areas, and yet doing nothing effective to stop or correct genuinely wrong decisions.

This is particularly likely to occur in fields which have been misunderstood and rashly declared "non-existent" — areas such as *metaphysics*. In practice, that ban conveniently cuts dead any criticism of those supposedly non-existent metaphysical assertions which the positivists themselves are busy making! Note that their agenda was inevitably based on their own metaphysical beliefs about knowledge-acquisition processes, and about metaphysics itself! (Also see section 4.1).

Hard-scientists are almost certainly right *to aim at* banishing vagueness and any appeals to unknowable causes; but two points are worth remembering. Firstly (as just mentioned) in the social sciences and elsewhere there are many issues on which it will *never be possible* to banish the vagueness of our knowledge[26] — so we had better keep or re-invent any feasible procedures for handling those vaguely-defined issues which social life will always throw at us, whether hard-scientists like it or not.

Secondly we should remember that even the hard-sciences started off with vaguely-defined concepts, and their clarification did not come just from experiments alone. Nor indeed has all the vagueness been banished even here: We may be very precise technologically in our inter-planetary rocket travel, with an apparently impeccable grasp of the mathematical formulations — but, as already mentioned in *footnote 3*, we are still unclear about *what gravity is really*, or how it operates through empty space, and it is even possible that some aspects of this vagueness will remain with us forever!

But if we may have to face vagueness even in physics, then how much worse will this problem become when we turn to the much more complex systems of macrobiology and the social sciences? Here it is effectively certain that we can never fully overcome this vagueness, so social life will always have unpredictable aspects.

To remove all vagueness — that is, to build and apply a fully predictive model of society and all its connections — we would have to satisfy a lengthy list[27] of requirements. A failure on *any one of these conditions* would suffice to sabotage the attempt, but in fact they will probably *all* fail us — and in most cases, for always and forever! In practice then, it is clear that such certainty is totally unattainable for truly complex systems like humans and their societies, though we do seem to be quite good at finding workable

proper cooperation between the disciplines, philosophers might have been able to suggest a better alternative. In fact such an alternative did surface, but its proponent evidently felt the need to be very circumspect in promoting this idea. That of course is Piaget's *"scheme"* suggestion which will engage our attention in much of the present project, including the introductory book (Traill, 1999).

[26] Not even if society itself were highly deterministic and known to be so. Even then, in that especially "favourable" case, the complexity of society would still prevent us from ever knowing enough about its detailed workings for us to make an exact science out of it for reasons summarized in the next footnote. We can of course *strive* to make it more exact, but we should also have the humility to recognize that today's best knowledge about social "laws" may soon be obsolete as the salient conditions continue to evolve over time.

[27] The list would include: (1) Overcoming the apparently-causeless Heisenberg-indeterminacy within the physics which we may suppose underlies all social and biological activity — or, if you attribute this physical effects of non-physical agents, then you might substitute an indeterminacy due to those occult causes. (2) Measuring every single variable with even the slightest influence on the system! (3) Collate all those variables. (4) Calculate their consequences, and do so fast enough to be useful, i.e. faster than the real processes modelled! (5) Exclude yourself and your computer from the study because you cannot rigorously model yourself, and yet you must include yourself because you cannot fail to be in causal connection with the system! (Landsberg and Evans, 1970). (6) Devise some way of infallibly extracting appropriate conclusions from this system, without being sidetracked by irrelevancies, nor missing any relevant detail — and of course this requires a proper understanding of the terms used here.

 — ISBN 0 9577737 1 4

second-best strategies for coping heuristically with this sort of problem.

In short then, the humanities will always play an important role in the social sciences and other areas which hard-science has failed to master. Hard-scientists may not like that conclusion, and indeed their criticisms sometimes have some merit. However it would be helpful if they could acknowledge that inappropriate use of their hard approach has often been counterproductive — and sometimes disastrous.

4. "SPIRIT" versus "HARD-SCIENCE"

4.1 *Metaphysical assumptions as obstacles to this reconciliation*

Could we ever find common-cause in some agreed joint policy between these unlikely partners — or is there some element of thought which ensures that their two world-views will always remain opposed? It would be nice to know now whether such a merger is, or is not possible.

Any thorough search for *the hardware of the mind* is bound to be looking for supposed mechanisms underlying everything mental — even those most subjective and apparently-spiritual aspects of our mental life; and that looks very like a hard-science approach. Against this, many arts-and-humanities people may well be inclined to ridicule any attempt to "mechanise" our subjectivity, our feelings, and our soul or conscious-awareness. — Overall then, that looks like a stalemate with no resolution; and, if need be, we should be prepared to accept this gracefully, at least for yet another generation or so!

However my feeling is that (i) both sides offer a good case on at least some points, suggesting that some sort of synthesis might ultimately serve us better, and anyhow (ii) such dualistic solutions are usually a sign of mere temporary expediency which we choose to accept pending some better solution — a solution which probably exists even if we cannot find it. So, if I am right on these two points, it would seem to follow that there is something about the fundamental assumptions made by either arts or science people, or both, which generates a needless misunderstanding between them. But what could that something be? And if we found it, could we apply if profitably here and elsewhere?

Some disputes continue endlessly because the participants base their reasoning on different deep assumptions over which they have no conscious awareness, simply taking them so much for granted that they are in no position to question or revise them. Some are deeply ingrained in this-or-that subculture, like notions of "rights, duty, truth, honour, justice and legitimate-power". Different subcultures will often have *different perceptions* about these abstract concepts; yet each culture may well take these parochially-defined ideas as very basic to its way of life — so much so that its members will sometimes defend their approach with a violence which can amaze and shock the neutral bystander.

As things stand, we are left with an uncomfortable combination of inbuilt *difference* over *often-emotive principles* whose origin is *poorly understood*. So disputes over such issues are only to be expected in a pluralist society; and in a monoculture we are likely find stagnation into odd belief-systems, since there is no incentive to examine one's traditional beliefs, nor even any apparent need to defend them.

Now the concepts of *"right, duty, and justice etc."* are all *abstract* notions — intangible, incorporeal, and apparently all in the mind[28] — so their validity cannot be demonstrated physically like the laws of motion, but only through consulting man-made edicts and procedures which are themselves abstractions, all taken in conjunction with our individual emotions and feelings, which also appear to us in an abstract mental form. In other words such concepts do not refer directly to the physical world outside the mind/ brain, but to the intangible mental world of advice or intuition on *how to deal with* that physical world. — So we are not talking about *physics* or other such sciences, but rather ***metaphysics***, that issue already raised in subsection 3.3(2).

For most of the twentieth century "metaphysics" has been a dirty word in most fields. This has been an unfortunate mistake, arising partly due to logical-positivists and their misunderstanding[29] of the how knowledge is acquired, and partly through some confusion amongst the four common meanings of the word. Of these four meanings[30], one was "supernatural", and that was quite anathema to scientific modernists who then seem to have roundly condemned all four meanings indiscriminately without pausing to look closely at what they were doing. Moreover the chief irony

---

[28] I.e. intangible and apparently-immaterial even if they do *refer to* material things like chairs, coins, and houses.

[29] For example: *"the impossibility of any metaphysics which tries to draw inferences from experience to something transcendent which lies beyond experience … Since rigorous inference can never lead from experience to the transcendent, metaphysical inferences must leave out essential steps. … "* — giving us *"mere illusory concepts which are to be rejected … No matter how much they are sanctified by tradition and charged with feeling, they are meaningless words."* — Rudolf Carnap (1930/1959).
Of course most philosophers now accept that this criticism applies to *all* inferences, even our "straightforward" conclusions about what we think we see! So there *must* be some other component to the knowledge-acquisition process which the logical positivists had overlooked.

[30] The four common meanings for metaphysical are, in brief:
(1) Relating to assumptions underlying an enquiry — (*the most central meaning, which directly concerns us here*).
(2) based on abstract reasoning — arising when the above assumptions are applied.
(3) abstruse or over-theoretical — a criticism of *style* rather than content, though if we then apply the 'Occam's Razor' criterion of simplicity, this will implicitly condemn the content also.
(4) supernatural or irretrievably mystical.
— *adapted from the Collins English Dictionary (1979).*
It *looks* as though all four meanings imply a lack of experimental varifiability and that this is what should link them. However this linkage is debatable as a defining theme, especially as it is now realized that nothing is genuinely and fully testable by experiment or observation anyhow, so then maybe we could include absolutely *everything* in this frowned-upon category!

 — ISBN 0 9577737 1 4

of this move was that this rejection of metaphysics was itself a metaphysical assumption!

Now while it is quite defensible to question the existence of any supernatural realm, it does not make sense to assert *as a fundamental assumption*, that fundamental assumptions cannot exist!

Such a ban on using basic principles is frankly incoherent, and is therefore unenforceable; and any refusal to allow discussion of such basics is surely dangerous. Yet the ban was taken seriously during the 1930s-1970s period, so the inevitable consequence was that discussion of these basic points was often forced underground; and like any illegal drug trade, it then operated in an uncontrolled way without proper public awareness or monitoring.[31]

The outcome, whether intended or not, has been a general refusal to consider fundamental aspects of such tricky processes as knowledge-acquisition (the very process that we *should* have been investigating in this context!) — or how these processes *might possibly work in principle*, regardless of whether we could actually observe such detailed activity.

If we can accept that we now know better, then this will clear the way for new developments previously tabooed by the old policy; and therein lies some promising hope for reconciling the best of hard-science with some perhaps-compatible ideas on what the "spiritual" nature of our minds really entails. That is, this would give us some hope for resolving the dualism.

There is a further interesting twist here. The thing that motivated the observability-emphasis in the first place was presumably a reaction against supernatural explanations because they were deemed to be an unscientific obstacle. Of course it was often religious critics (with an existing metaphysical commitment to supernatural explanations) who most challenged new alternative scientific ideas. Their criticisms sometimes have some validity in human terms, though I believe most of these can be solved by teasing out the metaphysical issues properly.

Part of the problem may have been • that these critics were chiefly motivated by their *abstract religious convictions* about the metaphysics of reality, • that these notions had evolved over the generations and come to symbolize real-but-perplexing human problems, and • that the connections between these two domains had never become properly explicit.

Indeed I would suggest that supernaturalism itself is not the real problem — that *certain types of supernatural system* could, ***if*** they existed, fit in quite happily with scientific systems[32], though other supernatural types would not. Two of the more crucial issues are, it seems to me, whether we see real-world as obeying *causal laws* and whether we may consider it as always having *structure*[33]. Both these properties just might be found in some hypothetical supernatural domains; but then either of them just might conceivably be missing from certain natural earthly systems. So maybe we should concentrate on these more-basic properties of structure and causality on their own, and simply not worry about whether a given system is natural or supernatural as such.

In this section I have argued that the difference in outlook between Arts and Sciences is probably due to fundamental differences in basic metaphysical assumptions. But because metaphysics supposedly did not exist, such problems were scarcely even recognized or discussed in the postwar period. So then, what are these basic assumption-types which underlie our thinking? I have just mentioned *causality* and *structure*, and I shall return to these issues shortly in sections 4.4 - 6.7; but first I would like to look at two assumptions about how the mind operates — issues which will be particularly relevant to this present overall project on the mind/brain. These are assumptions on how we "record" memory and how we "store" it.

4.2 Assumed memory-recording method — Lamarck and the tape-recorder?

When we learn something, what actually happens? For the moment let us forget about any official scientific views, and concentrate on the lay popular belief as we may infer it from everyday conversation and attitudes, including our own unguarded comments. Surely the common perception is one of *recording* our experiences, as if by tape-recorder or movie camera. These are obviously attractive metaphors in some ways, yet there is something seriously misleading about them: Nature is self-organizing and robust, while gadgets need painstaking design and maintenance. Then again, the natural brain is a bit inconsistent in its precision but very good at reconstructing from fragmentary information, while gadgets score well on precision but relatively poorly on their interpretation of their input.

But I would suggest that the really important problem with this "**tape-recorder**" **model of perception** is this: It is very difficult to offer any plausible-and-searching account of just how our own natural speech *etc* could be: *received* AND *stored* AND *retrieved* — and in a way which both behavioural biologists and hard-scientists would find acceptible.

---

[31] Likewise in Freudian theory, when we subconsciously try to *repress* unwelcome thoughts, these are likely to break out elsewhere — sometimes in bizarre ways.

[32] E.g. as we shall see in chapter 6, postmodern science might well be able to accomodate some types of supernatural entity (*if* they happened to exist) merely by postulating a few extra unseen dimensions, which would preferably also have some common interface with our normal three-dimensional space. Moreover this thought also points the way to another less-extreme solution involving a *hidden-but-not-actually-supernatural* domain, which I shall refer to as "**crypto**natural". — We will return to this in section 4.6. Also see Traill (1999, Appendix C).

[33] Preferably a dynamic self-sustaining structure, and *not* just the traditional idea of a rigid immutible structure which hardly seems to fit in with biological developmental concepts, and which has been condemned in the social sciences by "post-structuralists" for this very reason.

Reading is not the same thing as writing — and decoding is not necessarily the same thing as reversing-the-coding-process, especially if we expect something more perceptive than a mere *playback* at the reading phase. And even where a mere reversal will suffice, we still have to learn what it is to *reverse* something. (*Undoing* an action is not something which comes immediately to the young infant[34]; and I understand that sharks cannot swim backwards even though other fish can).

Moreover the brain would have enough trouble working out where to write its memory codings without also having to keep *devising new codes and techniques* to record new experiences — not to mention the need then for retaining these new code-conventions in some place accessible to conscious attention; and some critics might claim that this is merely creating still further coding-convention problems!

At this stage then, I suggest we drop or shelve this simple "tape-recorder or camera" model of brain activity, and try to find a better account. Here it is helpful to note a formal similarity to a problem in a different area of biology — the task of explaining how living species come in time to adapt themselves to be better able to cope with their environment. Of course this is now usually explained in Darwinian terms, and we will return to it again (next column) for the more common neuro-synaptic version — or *sections 5.3 and 9.4* for a Piagetian account which seems to offer some additional advantages.

But meanwhile it is interesting to note that before Darwin's *Origin of Species* (1859), there was a rival view which now looks suspiciously like the "tape-recorder view of memory which I have just been criticising. According to this **Lamarckian view**[35], at least some of the characteristics we acquire through personal experience are then *somehow also written into our genetic code* and passed on to our descendants.[36] There is no need to claim that this *never* happens, but it would surely require very special arrangements which would be very vulnerable to disruption, especially during successive generations.[37]

We have just been considering the popular view in two equivalent forms; but what about the *current scientific* view of how memory is stored and processed? In brief, there is widespread acceptance of the *trial-and-error* approach, and mechanisms have been proposed and investigated along these lines with some modest success. So far so good, and one interesting typical account is given in the book ***Neural Darwinism*** by Edelman (1987). What we are shown here is an initially random arrangement in the synaptic linkages between nerve fibres or their branches — followed then by "success-or-failure" feedback which serves to close down those linkages which are probably inappropriate, and hence we achieve adaptation.

I have no quarrel with this account as far as it goes. Such processes clearly occur, and they offer the same formal solution to *causality problems* as Darwin offered in the species-adaptation case — both depending on trial-and-error. However I have serious misgivings about its *structural* aspects. Here we have the sort of vagueness and slowness which we might expect of the tuning knobs on our TV sets. Certainly these are important (especially if we include the knobs hidden at the back), but they are by no means *all*-important. Their role is actually peripheral to the much more highly structured components and circuitry in the deep interior of the TV set.

Likewise I can happily accept that revision in the connectivity of my own synapses might well correct my manic mood, cure my migraine, resolve imbalances, channel information correctly, and perform impressive feats of pattern recognition; *but* I nevertheless find it very hard to believe that the loosely-structured synaptic adjustments alone could account for my ability to calculate 3.4×1.1 by mental arithmetic — or drive a car — or understand rapid speech. Moreover I suspect that most lay people (including those in the Arts and Humanities) would tacitly share this doubt, and not really take the adaptive-synapse theory too seriously as *the* supposed answer.

[34] Systematic reversibility and undoing — negating one's previous action — is actually quite important to the theory of how we come to understand *objects* (Piaget 1949, 1950). But such moves are often not as straightforward as we might suppose, as when we come to look at the fine detail of *actual transactions* (rather than such abstract idealizations as "-" for anti-plus, or "÷" for anti-multiply). Indeed we can readily discover this when we try to write machine-code subroutines to perform such action/undo pairs of activity in a computer environment. Or to take a simpler example: "down" is not just "anti-up" if it makes us use a *different* muscle. — (Section 9.4(3') discusses how possible muscle-instructions may be read from memory-code.)

[35] J.B.P.A.de M. Lamarck (1744-1829), the father of modern invertebrate zoology. In his book *Zoological philosophy* (1809), he pioneered, at professional level, the concept that evolution might be occurring — though we now know that his ideas about *causality* were wrong. (Asimov, 1964/1975).

[36] Of course this deliberate writing-down would seem to be in line with our *social* experience which *is* recorded on tape or penned into diaries and history books. Indeed society is yet another context which calls for similar information-building tasks, though in a rather different environment of possibilities. Actually I suspect that even here the process may be less Lamarckian than we suppose. After all, as any publisher, politician, entrepreneur or archivist will tell you, a lot of what is written down passes unheeded into oblivion or destruction; and some of us might see this as a sort of Darwinian natural selection. But that's another story!

[37] Anyhow the evidence does not usually support this Lamarckian view. Instead it seems that (always or nearly-always) adaptive change actually arises through blind spontaneous mutations or recombinations followed by the usual trend towards elimination of the least adaptive individuals then on offer — a systematic *trial-and-error* procedure, broadly conforming to Darwin's original view, though we might quibble over the details.

 — ISBN 0 9577737 1 4

4.3 Communication within the memory — analogue, digital, or what?

I am tempted to say we have been looking at the contrast between analogue and digital mechanisms, with *analogue* synaptic adjustments serving to fine-tune some unrecognized deeply *digital* processing in my mind proper. On the other hand there also seems to be a third category half way between them — the semi-digital millisecond pulses or "voltage-spikes" which travel in a chain-reaction across the surface of cell membranes, and most notably along the membranes of those lengthy wire-like axons, each of which grows out as a long hollow "finger" from the surface of its cell body thus linking it to distant synaptic contacts.

These *voltage spikes* do have some of the properties expected of digital technology, notably their fairly definite *on-or-off* nature (unlike the undisciplined analogue-response found in some primitive animals, with multiple gradations of *"x% on"* being allowed). However I have come to feel that these *on/off* pulses are not yet as disciplined and definite as we might wish a digital mechanism to be. Their graph-shape is less rectangular than one might hope for, but more importantly there is (as far as I know) no digitally precise pattern to the timing of successive pulses, apart from general statistical trends.

Perhaps then we have been unduly optimistic about their supposed signal-clarity, and that optimism might be due in part to the historical development of the topic: In the postwar decade or two, each neuron cell was seen as an on/off switch, fed by definite on/off inputs and delivering resultant on/off outputs — all in strict and neat accordance with the formal logic which underlies computer theory (Hebb, 1949; Blum, 1962). By now it has become quite obvious that this is an oversimplification since neurons are too extended and diffuse (each with its multitude of synapses separated by significant distances), so the simple on/off notion leaves a lot unsaid — and there is little to inspire confidence that, in real nerve systems, the same mechanism could ever be counted on to behave in exactly the same way next time the same input pattern arose.

For many purposes that will not matter and may even be an advantage, especially if there are other parallel mechanisms so that statistical smoothing can take place. But where we really need dependable precision, we might have to consider other mechanisms *if* any can be found. In fact Hebb himself forestalled this development by insisting that he was talking about the "*formal* neuron" — a hypothetical switch-like entity which might, or might not, happen to coincide with the actual neuron. Yet what other switch-like mechanisms could there be? That is a question which I shall come back to later on when I consider the case for molecular-level encoding in chapter 10 — and in *Part II* which then follows.

But let us return to the question of whether the millisecond voltage spikes should be counted as neither analogue nor digital, but something in between. Here we should note that these spikes are not an actual *storage* medium, but rather a means of *communication* — and this communication is clearly in contact with the accepted storage-medium: the synapse-and-whole-neuron memory system discussed by Edelman, Crick, and others since Blum (1962). But what if there really is an extra alternative storage system more attuned to digital requirements? Could it be that the voltage-spike signalling is just as much in touch with this extra system as well, and thus offer a bridge between the more digital and the more analogue systems, so justifying the "in between" status suggested for it?

Attractive though this idea might be, I am inclined to discount it; or at least argue that this action-potential spike system would not suffice for routine communication *within* the postulated digital system, even if it can serve for interface purposes. I will give my reasons in more detail within the next few paragraphs, but in brief: (i) the millisecond spikes would be much too long and slow, and not offer an appropriate technical match to the digital system if it happens to depend on molecular-level coding (as seems likely); and (ii) there exists a much more attractive theoretical possibility for how these signals could travel — a postulated system more akin to fibre-optics, and much more in keeping with molecular-level switching and with the high communicational *baud-rate* which we might expect within a busy system like the brain.

In our present context then, this discussion leaves the impression that there are two likely modes of memory storage: Firstly the analogue-orientated synapse system involving identifiable cell parts, communicating mainly through millisecond spike pulses. Secondly the postulated digital system, perhaps molecular.

But how could such digital-molecular code-sites be communicating with each other, with the outside world, and with other systems? I can see no credible way in which the traditional millisecond pulses could do this job *efficiently* and with the required *finesse*. In fact though, as we shall see, there is a strong case for implicating an alternative system which would use short-range infra-red emission patterns — emissions having frequencies governed by quantum activity. More on this later, especially in chapter 9, and in *Part II* (chapter 11 onwards).

I shall just end this section by noting that any molecule-based digital system would be particularly well placed to offer the Darwinian trial-and-error we were discussing earlier. Such evolution is a cumbersome process which needs a large population of candidates *and* lots of time compared to the time it takes for a single mutation to express itself.[38]

The synapse-modification model does seem to modestly fulfill these requirements; but if a parallel molecule-based system also exists, it would probably meet these require-

[38] Species evolution is thus only feasible if we allow for millions of years. This caused a problem when many thought that the world began in *4004 BC* on scriptural grounds, or a similar figure based on Earth-cooling calculations which overlooked subterranean radioactivity; see for example Nahin (1987) — chapter 11: "The Age-of-the-Earth Controversy".

 — ISBN 0 9577737 1 4

ments many times over, on both counts[39] — potentially allowing for much more spectacular feats of rapid adaptation; and so perhaps allowing my brain to understand your speech, *and* avoid daily car-crashes! — and generally *seeming to be supernaturally adept*, far outperforming any conventional robot.

4.4 Our assumptions about mental storage and retrieval[40]

How do we remember our mental picture of the typical chair in all its *three*-dimensional splendour? Or indeed how can our brains manage those simpler *two*-dimensional icons like roadsigns? It seems most unlikely that our brains could cope with actual *3D* models of chairs as physical structures within the brain. Nor is it entirely clear how our brains could deal with an internal *2D* model of *the Mona Lisa* painting. Yet we generally do remember the look of items like chairs and paintings quite well — if only for recognition purposes.

We are perhaps even better at recalling *one*-dimensional items like lists, sentences, action-sequences, and road-routes; not to mention learning the stage-role of *King Lear*, or the whole of *the Koran!* Moreover for this simpler *1D* case, it would be much less farfetched to imagine it mirrored by some matching physical structure within the brain. Thus a mere linear string of coding would seem to be quite manageable, whereas a *3D* physical chair-replica in the brain would not. In any case we have ample precedent this time since we now know that such *1D* coding *is actually* used for similar purposes elsewhere — as DNA and RNA in particular.

Suppose we went one further and considered isolated fragments of coding with no readymade collective structure. We would then be looking at *zero*-dimensional items — rather like brief notes on separate scraps of paper, which are then just stuffed into a bottom drawer. This time we would have no great trouble envisaging a brain-code model, but we might well ask whether such as disordered system could serve any useful purpose even if that system did exist.

This is actually a matter of some importance since the traditional "neural net" of formal switch-elements seems, at first sight, to have much in common with this simple disorganized *zero-D* system of "paper-scraps". In neither case is there any clearcut formulation as to which fragment of coding should follow after which. In other words it has no *inbuilt systematic* topological structure for sequence.[41] Against this though, it may be possible for each scrap to have its own private signpost to nominate the next scrap in line, and do so in any sensible-or-bizarre way it chooses.[42]

This is best illustrated by looking at how a normal computer program is written (either in machine code, or in simple versions of BASIC or FORTRAN without the more sophisticated techniques). Each instruction may be seen as an element or *pseudo-point* which produces this-or-that effect before allowing the next element to take effect. But the big question here is: *Which* element is to be next? We have learnt to take it for granted that it is to be *"the next element down the list unless we are explicitly told otherwise"* — and indeed that is how computers have been predesigned to behave, thus tacitly giving them the same *1D* property as speech and DNA

As for the question of being "explicitly told otherwise" about where to focus next, this is provided by specialized "jump" or "GOTO" instruction-elements (like similar instructions in text, music-scores, and RNA). We are accustomed to these *jump* instructions being comparatively rare, but in principle such an idiosyncratic signpost could be part of *every* element's repertoire.

It seems to me that this is effectively what we have with a traditional "neural network", whether literally composed of actual neurons or simply of elements with some of these properties ("formal neurons" in Hebb's sense). Such complex and potentially adjustable connectivity seems to be exactly what is needed for certain types of neural activity, be they natural — or be they artificial, as in the proven commercial use of this approach for pattern recognition. Obviously the *jump* instructions here take the form of physical connectivity — a direct "wiring" which links each element to one-or-more other elements which are then to be next in line (at the whim of the present element, though probably also partly at the whim of any other currently-active elements which are also linked to it).

---

[39] That is, very much faster, and with a much greater population of basic units.

[40] of skills, lists, pictures, and object-concepts, *etc*.

[41] *2D* systems offer before-and-after neighbours in *two* different coordinate directions; and *3D* systems offer them on *three* coordinates; — arguably too much freedom of choice for the brain to handle easily! In contrast, a *1D* code-string normally offers a single unambiguous way forward (as long as we can distinguish *forwards* from *backwards*!); and the genuine *zero-D* case offers us no systematic progression at all.

[42] You may have noticed a further complication here. These *scraps* are masquerading as mathematical points in our discussion of topology and dimensionality, and so too are any coding elements in the *1D* strings, or any *2D* and *3D* equivalents which may exist. Clearly these would have an internal *ultramicro* structure of their own (probably in *3D*) but we need to ignore this in our present dimensional discussion, except when it comes to the question (just raised in the main text) of each element being potentially able to have its own private signpost to the *next* element — and that linkage ability betokens a substructure for each "pointlink scrap" which we could not ignore, though we could play it down.
This ambivalence over dimensionality (depending on scale) shows the need to view such systems within a *hierarchical* framework with at least two levels of resolution: one which views the code-elements as pseudo-points, and another which treats them as *3D* microsystems in their own right.

 — ISBN 0 9577737 1 4

Of course with such a network of *linkages*, the elements are no longer isolated zero-dimensional "note-scraps in my drawer", but then they are not necessarily any sort of *1D*, *2D*, or *3D* system either. In fact they will generally have a complexity which does not fit neatly into to the notion of *n*-dimensional space at all.[43] Herein perhaps lies their strength and their weakness — well suited for some purposes, but lacking the rigid discipline and dimensional-sequencing needed for precision and reproducible ordered actions, especially if we also need a massive trial-and-error capability on a second-by-second basis. For these latter properties we may thus have to turn back again to structures which do have these linear preset-order properties, as in a standard computer program, and in the RNA copied from DNA.

As for the question of understanding images of *2D* paintings and *3D* objects, this is probably a matter of such subtlety and complexity that I suspect it needs all the help it can get from *both* approaches simultaneously. Thus, as part of the process, we may well use our neural-net systems to achieve recognition in a broad sense, maybe directly coupled to emotional centres and hence involved with subjective feelings; — and meanwhile our *1D* systems enable us to trace outlines (visually and/or manually) and hence achieve some idea of exact proportion and geometry.

However, in this present book, I shall not have much to say about such figurative memory systems. Clearly visual processing is very important, and a hugely disproportionate part of our brains is devoted to such tasks; but I believe it is preferable to concentrate first on the many simpler tasks of the mind/brain, like speech-recognition or action-sequences. These will be challenge enough, and anyway many of their most fundamental principles are likely to be the same as for the vision tasks, so the one will surely illuminate the other.

4.5 Reconciling views on substructure, mechanism, and reductionism

On the whole, physicists and chemists have no problem appreciating the role of structure in nature; but even hard-science people like these have some ambivalence over substructure at the quantum level, and perhaps even more when it comes to discussing whether there is any rigorous structuring within their own minds.

So we need not be surprised if the Art-and-Humanities community has even less trust in structural notions — and of course the same distrust also falls on any other mechanistically orientated approach. As already mentioned in the last footnote of section 4.1, there has even been an explicit rejection of structure, in the form of *"post-structuralism"*. That trend may immediately make good sense to those confined to disciplines like anthropology or literature; but to those of us concerned with physiological biology, such an anti-structure movement may seem perplexing at first. In fact there seems to be a major area of mutual misunderstanding here:

To many, a structure is best exemplified by the Eiffel Tower or the Sydney Harbour Bridge — or rather idealized eternal versions of them. They are seen as being composed of a countably finite number of rigid components like girders, nuts, bolts and rivets, all rigidly locked into place like the *constants* in a mathematical equation. If change is tolerated at all, it will be carefully circumscribed and controlled, as with the rigid-but-rotatable expansion-rollers on the bridge — or even the conspicuous motion of wheels or pistons of a steam locomotive moving under well defined laws. Such allowed changes might perhaps be seen as being like *variables* within the mathematical equation, but otherwise well-controlled by a host of constraints within that equation.

From this quasi-static viewpoint, such messy things as biosystems would not generally be proper structures, though they might *produce* structures like shells or spicules. Perhaps one might grudgingly concede that the adult eye (for instance) could pass as a structure, but one would then be discomforted by talk about its embryological growth preceding that. Needless to say, such a rigid model is not very applicable to systems which really are "messy" and dynamic, like sociology or literature. Small wonder then, that social scientists have mostly come to reject this approach and hence the word "structure" which they attribute to it.[44]

The *second notion of structure* might more helpfully be called a "dynamic system" or a "self-organizing system"[45];

---

[43] The formal-neurons in a network *are sometimes* organized in a relevant spacial sequencing, as when they map areas of the body surface onto areas within the brain, though that could just be a matter of arbitrary convenience; and note that there are some notable breaks in the map, as when it is split between the brain's two hemispheres. In fact, in principle such formal neurons should be able to be sequenced in any way at all; though that is not to say that this free re-sequencing could be done rapidly or efficiently, or that there would even be any plausible influence which could bring it about. And anyhow, too much freedom is often simply debilitating (Ashby, 1952). So these flexible networks may have partial dimensionality, but it is questionable how useful that could ever be for producing precise or reproducible action patterns of the sort we are more concerned with in this current project; and that, of course, gives urgency to the search for faster and more reliable mechanisms.

[44] Economics is one social science which does tend to keep to the undesirable quasi-static approach. Economic-rationalists in particular do place great faith in mathematical models which have often been known to omit many of those important factors which do not fit into the neatness of the equation system, nor into the logical positivists' experimental-evidence doctrine. The consequences are sometimes most unfortunate; but before we condemn these economists, it could be helpful to understand the pressure they are under to produce policies which *seem* well founded — and bear in mind that there are not yet any well formulated and well-publicised alternatives which can be made to seem plausible enough (and safe enough) to gain acceptance by the power-brokers of this world. Maybe such advances will happen some day!

[45] Norbert Wiener (1948) originally developed the study of such self-organizing systems, for which he coined the name

 — ISBN 0 9577737 1 4

and examples would include the Solar System, a live *amoeba* cell, an individual mouse or human, a viable social community, and a natural eco-system in the wild. Not one of these systems is rigid or fixed in its configuration — and not one could survive as a system if anything seriously hindered its internal motion. True, there may be conceptual difficulties because the system's *components* might seem to be rigid and eternal, like the planets within the Solar System, for instance. Of course we do now know that these components (planets, cell membranes, or whatever) are ultimately made up of atoms and molecules, and that these are actually less than perfectly static, so they too are often part of a dynamic self-sustaining ensemble.[46]

In general then, there will be a *hierarchy* of substructures within substructures; and usually we will find some self-perpetuating dynamics happening within the substructure and perhaps giving it any superficial appearance it might have of being rigid and static, as viewed at low resolution. Presumably any such hierarchy *must end somewhere* in some truly basic building-blocks, though in the nature of things, such extreme entities are likely to be very difficult to detect and identify. In practice however, we can usually call a convenient halt at some stage where the subsystems really do *seem* to be stable and eternal (though with "rule-books" which perhaps betray a yet deeper world which we choose not to disturb). — Thus in chemistry we may stop at atoms or ions, and not normally dig too deeply into nuclear matters. And for the Eiffel Tower, we might well stop at girders (*along with* an engineers' reference-book on elasticities, densities and other such details).[47]

But what can this tell us about the misgivings that many humanities people feel about a structured *wholly material mind/brain?* Does it show why they distrust the idea of a complete structured-matter base for the mind? I have suggested that their main objection is against the *static* meaning of "structure"; and they would have some justification for supposing that the synapse-modification model is just a minor tinkering with this questionable static approach, since it says nothing about wider issues such as gross development and higher intelligence.

Could these critics accept a more flexible model which is structural in the second sense — i.e. self-organizing and dynamic? That remains to be seen. I suspect that the suggestion has not generally been put to them hitherto — at least not in any coherent and user-friendly way.

Suppose these humanist critics still do *not* accept the now-orthodox notion of a structured material mind, what other alternatives are there? We are left with several other possibilities to explore if we vary our normal assumptions — dissenting options which seldom appeal to scientists or materialists. Such hypotheticals need to be looked at, if only to clarify their ultimate weaknesses, and we shall take up that question in the next chapter.

4.6 *Possible building-blocks — for building any structure*

Before looking into these mind-structure options themselves, we should perhaps digress into a brief generalized look at *ontology — basic notions of existence* — existence of matter, or any forces (natural *or* occult) which anyone thinks might *somehow be out there* impinging on us. Presumably any mind/brain will owe its properties to some combination of whichever basic influences actually exist; so it is of some importance to have some explicit opinions about them, though of course not everyone will agree on which ones can be safely discarded as non-existent.

We should perhaps also have the humility to recognize that any such fundamental ingredients probably exist quite independently of any minds which may now be built up from them.[48] It is in that *independent role* that we shall first consider these conceivable components or media, (see table 4).

Why should we worry about the ultimate building-blocks of all structure? In the present context there are two reasons. The one concerns those other mind/brains of *other* beings which we choose to study; and the other concerns our *own* thought processes.

Firstly consider the key task of any particular "third person" vertebrate mind/brain ***x***. The key task of this brain is to come up with some *workable* map-or-model of its real environment. Note that "workable" does not necessarily mean *true* or *faithful*, though it does imply that it should lead

Cybernetics, before that word was seized upon by futurists with a more hard-science agenda — notably by robotics-enthusiasts and science-fiction writers. This later usage rather misses the point of the original innovative self-organization idea: Thus the typical notion of robot might *allow* it to be self-organizing; but it surely *does not require it* to be so. Moreover the development of the further fashionable concept of "cyberspace" simply adds to the confusion.

Instead of the now-corrupted word "Cybernetics", the alternative name *"General Systems Theory"* has sometimes been used, though the implied meaning is perhaps slightly different (Klir, 1970). There are one or two other recent terms like *"Complexity Theory"* which also seem to apply to the same subject area.

[46] We might even like to debate whether those girders of the Eiffel Tower form a dynamic ensemble in this sense, or whether they really are a static collective structure made up of a fixed crystal-like lattice of iron, carbon and other atoms!

[47] In fact even our explication of computer programs has come to be like this, with guidebooks describing the gross properties of standard subprograms which the programmers are then expected to use unquestioningly, even though such artificial systems *could* theoretically always be explained in every last detail; (and in the early days they often were).

[48] Today that seems fairly obvious, but note that it would seem to clash with the views of philosophical *idealists*. Over several centuries these philosophical idealists have believed that mind-and-ideas are more fundamental than matter, so that matter is merely a derivative. That *could* be seen as causing a circular argument here.

(As for that mystifying special use of the word "idealist": — It would be less confusing to the non-philosopher if this word were: "idea-ist"; but I suppose philosophers are unlikely to agree!).

 — ISBN 0 9577737 1 4

Table 4

WHAT COULD THE FUNDAMENTAL BUILDING-BLOCKS OF MATTER-AND-MIND BE?

	Ineffable spirit (non-sci?) 0	Pure basic continuous medium 1	Continuum, but with Knots or Vortices, etc. 2	Fundamental particles with "properties" 3	Rigid law-free particles (Democritus) 4	Any others ? n
Natural	Totally unstructured!	waves in 1800s Fresnel, Maxwell	Kelvin, **Bohr** standing waves?	Dalton, **Bohr** chem, nucl.phys	anti-Newton!! not coherent	
Crypto-natural	in medieval sky-Heaven?	the æther? vital forces?	**ultra-micro** or macro	**ultra-micro** or macro		
Super-natural	God, viewed traditionally	none?	**maybe!** hyperspace?	**maybe!** hyperspace?		

⟵ Modern wave-particle dualism of the quantum theory ⟶ (columns 2–3)

x to produce adaptive behaviour sufficiently often (whatever that might mean)! [49]

Even as better-informed investigators we have no guaranteed omniscience about ultimate reality either; but we may perhaps decide on a short-list of possibilities, and what strategy it would be best for ***x*** to adopt in each case. The columns in the nearby table offer us several rival possibilities; and of these, columns *3*, *2*, and *0* probably give us the most common tacit assumptions about the basis of reality — namely (*3*) *particle*-like, (*2*) *wave*-like, or (*0*) ineffably mysterious and *spirit*-like.

As all three have some claim to be taken seriously as guides to success in practical matters (either material or social), it seems that ***x*** would be wise to become skilfull in working with all three of these mental outlooks, according to the prevailing conditions. Thus:
(*3*) cogwheels, forks, electrons-inside-your-TV-screen-tube, and even fried eggs, are all best thought of as either *particles*, or objects made up of linked particles (with gross *objects* also being particle-like in some ways); —
(*2*) magnetic fields, electons in an atomic orbital, and free-travelling light are best thought of as wave-like; — while
(*0*) personalities, emotions, and human foibles will often be best treated as "spiritual" and beyond explicit understanding — provided that ***x*** has a sufficiently developed *intuition* and *savoir faire* to cope with such situations.

I emphasize again that the choice of any one of these modes, or any set combination of them, does not necessarily endorse them as being ultimately true. However it may betoken a successful mode of interacting with reality in its current form. Accordingly it will not generally be a serious sin if the person or animal we are observing, is *inconsistent* from one situation to another; such as using one mode to cope with electron beams, and another to work with their atomic orbitals — indeed that ambiguous *wave/particle* dualism has long been official policy within the quantum theory of physics!

This question of building-blocks will surely arise whenever we are considering anything which seems to have structure or any self-sustaining performance — that is, almost anything other than fleeting transients which have already disappeared before they can be of much interest. That will apply to happenings in nature which we might want to observe or speculate upon, and it will surely also apply to any code-structure which may develop within our mind/brains.

That is not to say that the type of structure within our minds will necessarily have the same *type* of structural units as those units in nature which it purports to model; but surely both domains must have some structure, made up from some sort of building-blocks. Perhaps that is just my opinion, but then I really do not see how anyone can proceed scientifically with a totally unstructured natural world — nor with a totally unstructured mind! So, as I see it, that seems to rule out case (0) with its ineffable spirits, as a creditable basis for either domain. That leaves us with a choice of either

• *particles-with-properties* (case 3), or something like
• *standing-wave "knots-in-a-force-field"* (case 2) which could be considered as fuzzy particles —

and both are acceptable to contemporary physics.

In either case then, we may *provisionally assume the existence of structures made up from building-blocks of some sort* — both within the mind/brain and in nature generally (and even including any supernatural components if they happen to exist for either mind or nature).

That prepares the ground for the next chapter where we will go on to consider the causal interrelation between these

[49] ***x*** may not need this insight; but what about our own needs as investigators of ***x***? It will be useful for us to have a deeper understanding of the reality that ***x*** is grappling with. That way we will be better placed to understand ***x***'s successes and failures.

two domains — mind and its environment — given that both are now assumed to be wholly-or-mainly *structured* in the dynamic self-organizing sense.

That will introduce some new problems relating to mechanism and reductionism — both of which (as commonly defined) have a bad name amongst those in the arts-and-humanities. It remains to be seen if some mutually-satisfactory re-conceptualization can be found.

 — ISBN 0 9577737 1 4

5. CAUSALITY, MECHANISM, AND KNOWLEDGE-GROWTH

5.1 Causality and information — compared to energy

Explaining the terms

What is *information?* Is it just something passed down a causal chain known as a communication channel? If so, is it merely conserved like total energy? — or perhaps diminished like the amount of "usable" energy? Alternatively can it grow, transcending any conservation laws? At any rate its technologists are confident that they can *measure it*, using units like bits, bytes and baud-rates. Thus we have a ready made basis for discussing its quantity — though there are some traps for the unwary like "Do we *double* the amount of information when we take a photocopy?"[50]

Then again, what is *causality*, and what natural laws must *it* obey? It is not entirely clear whether we can speak of *quantities* of causality without laying down some new rules. We would probably agree *that causality does keep growing* when some trivial accident triggers a skirmish, and that then escalates out of control into riots and worse; though actual measurement might be a problem. Likewise we can envisage causal-chain disturbances *dying away*, even though we may not always agree on what constitutes "the separate events" which can then be counted when we are seeking some exact figures for a before-and-after reckoning. However we can tolerate that degree of quantitative vagueness in the present discussion, as long as we are aware of it enough to avoid sophistry.

Anyhow we are left with the picture that neither information nor causality are conserved. They can both diminish, and it seems they can both grow (though probably *NOT* starting out of nothingness, as we will see). In contrast of course, *mass-energy* cannot be created nor destroyed, and neither can *total momentum*.

In one special set of cases we might say that causality *is* conserved, locally:
A solid rigid planetoid rotating in isolated space will have a certain state-and-orientation at this instant (t = 0). Shortly afterwards (t = 1, or whatever) it will have rotated a certain amount which need not concern us, but the important point here is that its new configuration will be fully and unambiguously causally connected to its previous one. It won't be a growing or transforming image; and *it certainly won't* be one of a completely unconnected set of images following each other subliminally, as if placed arbitrarily on successive frames of a movie film.

In fact the system will be dead boring, and fully predictable — not quite so boring as a complete static inactivity, but not far from it:

$$\mathbf{A_0} \to A_1 \to A_2 \to A_3 \to A_4 \to A_5 \to A_6 \to A_7 \to A_8 \to A_9 \to \mathbf{A_0} \to \cdots$$

$$\mathbf{B_0} \to B_1 \to B_2 \to B_3 \to B_4 \to B_5 \to B_6 \to B_7 \to B_8 \to B_9 \to \mathbf{B_0} \to \cdots$$

$$\mathbf{C_0} \to C_1 \to C_2 \to C_3 \to C_4 \to C_5 \to C_6 \to C_7 \to C_8 \to C_9 \to \mathbf{C_0} \to \cdots$$

$$\mathbf{D_0} \to D_1 \to D_2 \to D_3 \to D_4 \to D_5 \to D_6 \to D_7 \to D_8 \to D_9 \to \mathbf{D_0} \to \cdots$$

Such neatness is never fully the stuff of life (though we might sometimes like to get closer to it when our daily round gets too hectic). However we may fairly say that causality *is* conserved here, and we will expect that to happen whenever a system seems to exist eternally-but-dynamically — *always fully caused by former incarnations of itself* (and doubtless using negative feedback from its own internal activities, and exerted such that no parts will "fall out of line"). Of course such regularity and constancy will not be true for causality in general.

Nevertheless it seems to me that there is a hidden popular *expectation* here even if it is contradicted from time to time. This is a subjective feeling we seem to have that at least the number of causes will roughly match the number of effects over any given time interval. Leaving aside the occasional mistake of believing in "one cause for each effect", we may nevertheless sometimes feel that

- a typical *cause* has *n* effects
  e.g. "a typical adult has 2 children", while
- a typical *effect* has *m* causes,
  e.g. "a typical individual has 2 parents"

and generally speaking the two trends will balance out, with $m \approx n$, or maybe even "*$m=n$ on average*".

This would betoken a comforting steady-state in the overall complexity of the system considered — (an assumption which fits in with the simple world of Newtonian particles colliding in a frictionless ideal environment).

Of course physicists themselves have long accepted the Second Law of Thermodynamics (Clausius, 1865) which says in effect that any closed[51] system will eventually lose its diversity, running itself down to a state of ultimate dead-and-boring uniformity — or, to put it in technical terms, the closed system will reach a state of *maximum entropy* (i.e. *minimum variety* — or equivalently *minimum "negentropy"*) for the system. This amounts to saying that the *many causes*

[50] This present discussion of *information and causality* (and changes in their *quantity*) serves as an introduction to:
• Section 5.6 where we will look at the knowledge/information held within a system, and whether it can grow from scratch — from a starting point where there is initially no knowledge at all; and • chapter 6 which looks at the supposed possibility of *"uncaused events"*, and what connection (if any) they might have with *freewill*.

[51] Here "closed" means a bar against any new stimulation from outside the system. This run-down of a system can be seen in computer models too, as discussed theoretically in Ashby's *Introduction to Cybernetics* (1956/1964).

 — ISBN 0 9577737 1 4

that we started with, will ultimately degenerate to a *single static effect* or something very like it.

Entropy in information theory

We have just been talking about energy/matter and the usefulness of how it was organized physically. However in the field of *information theory*, the situation is a little different — with systems "open to energy but closed to information" (Ashby, 1956; Klir, 1969).

(How, in detail, do we store meaningful information? That is a difficult puzzle to solve, as the present project attests; and yet *energy* conservation is not a major part of this problem. In principle, masses of data can be stored and retrieved using minuscule amounts of energy; so it is customary simply to assume that adequate energy will be freely available, and that any significant constraints will come from elsewhere. Such approximations can sometimes cause us trouble later when we forget their limitation; but for our immediate purpose this seems to be an acceptable assumption.)

Anyhow the concept of entropy was co-opted into information theory by Shannon and Weaver (1949); and it has been a crucial feature of that subject ever since, with negentropy serving as a measure of information content. Shannon's original context centred on telephone cables and other *passive* signal-channels; and there the important issue was to minimize the loss of signal content (negentropy in the informational sense, regardless of energy considerations). So at a formal level, this conserving of precious informational-organization, looked almost identical to the older conservation of precious energy/matter-organization — a mathematical equivalence between the two types of negentropy.

However now we are *not* just studying passive cables, but rather how the brain develops knowledge for the individual, and how science develops knowledge for society. In both cases we hopefully have an *increasing body of knowledge*, and not just the rote relaying which is all that a cable is expected to deliver.

So note the important new departure, disowning any analogue to the Second Law of Thermodynamics within the wider information-handling field. Unlike the energy-system degradation of Clausius, and unlike the cable-loss of Shannon and Weaver, a properly operating brain or science-policy will actively *generate* knowledge — *increasing informational negentropy* — seemingly out of nothing but an ill-defined exposure to reality! But this does not fit at all well with the modernist-positivist views of rigorous logic; so how is this knowledge-growth actually achieved?

5.2 The Growth of Knowledge

Basic Knowledge-handling.

Here we have again that old question of how knowledge is generated. That is one of the key issues of the current project, and we will return to it repeatedly so we need not dwell on it right here. Suffice it to say that the process appears to depend on self-consistency tests which seem not to be available to energy/matter systems nor to simple cable-transmission.

However here we need to consider another vexing set of questions: Can this self-instruction process really start from nothing but the mere presence of an environment? (In which case, why can't every stone and water-drop do the same?) Our mind/brains do seem to learn a great deal through some sort of effort of their own, starting with comparatively little; but did they (or their original ancestral form) really start from *nothing-at-all?* And if not, then where did they get their starting-instructions from so that they *could* start?

Here we have the makings of some emotive differences of metaphysical opinion:

• The conservation fallacy, as if we were dealing with energy

Firstly there will be some who *deny this growth of knowledge altogether*, though they may not put it in those terms. Thus it seems to me that some religious fundamentalists could be saying in effect:

> "There is no way mere mortals can wrest valid knowledge from nature, and no way for such knowledge to *grow*. — Such wisdom is in fixed supply (like your mass/energy) and what there is can only come from God. — Moreover, any dissenting views serve only to cancel out some of that precious limited supply, so such dissent is a sort of vandalism".

We need not agree with such notions! However if we do meet people with such a view, we should be aware that their outlook may stem from a very basic metaphysical assumption which we might not share. (In this example, the assumption is that knowledge is a conserved commodity like our mass/energy.) If we can at least see what their world-view is, then this insight may take us one step closer to peaceful co-existence — maybe!

(More surprisingly perhaps, I would suggest that this need for metaphysical insight also applies in part to traditional formal logic. After all, here one starts with "premises or axioms" as the given knowledge, and anything else is then a mere transformation of that same *conserved-but-non-growing* body of knowledge. All impeccably valid within its own terms of reference, but not making any real progress. It is for that sort of reason that Piaget (1949, 1952) criticized the work of such *"logisticiens"*.)

• Pump-priming.

The *second* metaphysical debating-point is this. Given that knowledge does usually *grow*[52] (unlike its energy ana-

[52] Encoded knowledge may grow in at least two different senses. (1) It may *replicate*, creating extra copies of codings which already exist. But here we will be more concerned with the more creative aspects: Thus (2) We may also see the *creation of new knowledge* — whether by Darwinian or Lamarckian strategies (see section 4.2) — and either from scratch or by adapting some pre-existing code. The following discussion may help us to choose between these competing details.

 — ISBN 0 9577737 1 4

logue), and given that mere unstructured contact with the environment will not alone produce meaningful knowledge; what else then *does* the knowledge grow from? We seem to need another cause — an internal cause — but what bare-essential properties must it have? We shall consider both internal mechanism and pre-existing coding:

Does this *"internal cause"* need some sort of *organizing mechanism* within the being-or-society which acquires the new knowledge? Presumably yes, no matter what the process involved. Thus if (i) we envisage a Lamarckian "tape-recorder" type of learning process, then the need for a mechanism is quite obvious — and very demanding. If instead (ii) we adopt a Darwinian-selection view of the learning process, there is then much less call for elaborate mechanisms; but surely there must still be some sort of structure to handle the "success-or-failure" feedback-signals at least?

And surely any such purpose-built structure must owe its existence to some *pre-existing* knowledge (even if only implicit in structure) — a pump-priming "investment" which should then bear "interest".

5.3 *Media for the storage of memory*

Moving on now from creative and transmissive mechanisms into means of storage: This entails looking at possible ways of organizing the *coding-media* — the "note-pads" which might be used.

Does this "internal learning mechanism" need any sort of *"writing-paper"* to work on and modify? Would a "blank page" or "blank tape" suffice, or maybe a "ruled-up page" like an official form? Or can it encode its innovations entirely from scratch, without anything but some irreducible beadlike elements which it can somehow string together? Here our needs seem to depend once more on whether the process is supposed to be Lamarckian like a tape-recorder, or a Darwinian selection process. Let us consider these two cases:

The Lamarckian case. A truly ideal tape-recorder would tirelessly manufacture its own tape as it recorded! Our less-ideal existing tape-recorders do need a blank tape; but perhaps we may still allow that as being free, or *mostly*-free from pre-existing coding. Likewise with any similar "write-it-down" learning system postulated within biology: Any such bio-recording system would presumably need to generate at least some ordered linear codings, either by *selecting and assembling* the supposedly-correct units in real time (like a very fast type-setter of pre-computer days), or by *setting "switches" along some pre-existing blank* coding-strip. Neither is entirely impossible, but the logistical problems would be formidable.

Anyhow the immediate point is this: *If* Lamarckian recording were genuinely in use, then maybe this recording *could* theoretically be done with no pre-formatted coding, not even any "blank tape" (something which a Darwinian system could not do). However prefabricated aids like this pre-formatting might still be very helpful here, even if they are not strictly essential.

How then are we to interpret the evidence for any pre-coding? If we could demonstrate the *absence* of such pre-coding, that would indeed be evidence favouring the Lamarckian model. However the *presence* of precoding would be compatible with either model — it being *perhaps-desirable* for the Lamarckian model and *essential* for the Darwinian process.

The Darwinian-learning case.[53] This envisages trial-and error selection within populations of pre-existing candidate codes. Most members of this population are then discarded during the learning process — leaving only the apparently-correct encoding *as if* it had been written down. By definition then, this strategy does definitely need at least some pre-coding to exist beforehand.

In other words, the Darwinian approach. always needs some preliminary good-or-bad hypotheses to get it going — and it cannot just write down a new encoding out of nothing. This has important implications for the nature of originality under this strategy: According to the Darwinian explanation of thought, we *can* be original by making a suitable selection from a set of spontaneous "mutations" to pre-existing ideas; but we cannot be 100% original because we are always somewhat dependent on those raw idea-codes as being the source of the mutations.

5.4 *Causes which may be minuscule, yet never entirely absent*

This all leads to a general postulate which I shall call **the principle of seeded knowledge-growth**: *Knowledge is not a conserved property like momentum or energy, but rather it can grow open-endedly. However it cannot start from zero, so it must have some encodement of supposed knowledge to begin with — even if this encodement is a mere fragment, and even if it misrepresents reality!*

This calls for several comments. For the first few, I shall continue to assume the Darwinian-learning model; but I shall eventually return to re-consider the Lamarckian approach in the next section.

(1) Note the formal similarity to life and its reproduction. This should not surprise us *if* we agree that both are operating on the same Darwinian-selection principles even though they may be using quite different mechanisms. In either case, proliferation of copies can sometimes be very quick (ideas[54] can spread very rapidly, and so can bacterial

---

[53] Introduced in section 4.2 above; and also see *Book D* for an account of the likely build-up of a hierarchical system.

[54] Such ideas can once again be thought of either within the mind of the individual, or equally within society. The latter is discussed by Cavalli-Sforza and Feldman (1981), including reference to Richard Dawkins' (1976) concept of the "meme" — a term fulfilling a role roughly comparable to Piaget's "scheme", though primarily operating in a different domain. Of course both also have some formal similarity to the "gene" concept of genetics. (Actually Dawkins (1976, 1982) uses "meme" more-or-less indiscriminantly for both social and brain based knowledge, though the emphasis seems to be on the social

 — ISBN 0 9577737 1 4

populations); but in neither case can this happen without some sort of "seeding" operation to start with — a single thought-element, or a single bacterium would suffice — but that minimal contribution cannot be bypassed.

(2) But where does *that original encoding* come from? For the individual we can easily see that it could be inherited along with the encodings for bodily growth; and of course the earliest societies could acquire their "memes" from individuals. However for the case of our earliest microbial ancestors we run into a bit of a problem since the above *seeding-principle* does not allow them to start from zero. In fact there is nothing new in this problem. It is essentially the same as the mystery of life itself. In both cases some previous encoding is needed before the process can get going, but how could that prior encoding arise?

The answer seems to be that such first-time encoding has to be simple enough for it to have formed spontaneously by chance. Of course this is a much discussed issue, at least for the origin-of-life question. As the encoding for life-activity is now known to be normally stored in DNA, and transcribable to RNA (with occasional cases of transcription the other way, back to DNA), it has been usual to speculate on the spontaneous formation of viable encodings in one or other of these two nucleic acid types. I do not propose to go further into this well worn debate except to repeat the suggestion that the first "seeding" for thought could well be part of the same process, or at least have followed a similar path.

However there is now a new possible variation on this theme. PNA is a more primitive type of nucleic acid, and it has recently been shown that it can exchange codings into RNA, and the reverse; (Böhler, Nielsen, and Orgel, 1995). This opens up new possibilities for both: the first life-code, *and* the first thought-code.

(3) A few pages back I argued that the minimal requirements for a thought-system were *"internal mechanism"* to do the code-reading, as well as the "pre-existing coding" which we have just been looking at.[55] For normal bodily life processes, it is the ribosome which acts as the mechanism for reading the RNA's code, normally using this code as protein-building instructions (Miller, 1973). Should we envisage similar ribosome-like mechanisms for reading molecular encodings for memory or thought, assuming such molecular encodings exist? Actually I will argue that, for such communicational purposes, the mechanisms could well be *simpler* — and indeed they will probably *need to be* simpler!

The point is that signal-handling is a rather different process from body-building, with rather different logistical requirements when it comes to actual performance — and indeed these requirements would still be different *even if their instructions turn out to be encoded on the same medium*, using similar-looking coding conventions. Thus protein-building instructions arrive as coded patterns on RNA; and *maybe* elements of thought (such as Piaget's "schemes"?) arrive in the same way, also on RNA or at least on some sort of linear nucleic acid substrate. But what then?

Protein-building obviously entails the handling of the solid "bricks and mortar" of the body's architecture — along with the need to acquire *the right* solid ingredients (amino acid molecules) in the right sequence. But probably there will be no desperate urgency about the task, so an unscheduled wait for materials would be no great calamity.

In contrast, signal handling is not like that at all. Incoming signals must often be attended to immediately in real-time, for otherwise they will often be lost forever. But then if I am not building a wall, I do not need to wait for bricks, nor do I need the help of a concrete-mixer or spirit-level. Instead I am doing the white-collar job of processing messages — a task which need not use heavy equipment, nor routinely depend on material supplies, but which does often depend on the prompt response provided by an efficient office with a well-organized filing-system.

At the molecular level then, even if action-messages are encoded in the same sort of way as building instructions, the related mechanisms might well be quite different, and probably less materially cumbersome. Thus it may be that the ribosome (the "protein-factory") would be superfluous, and a simple "desk-or-switchboard" might suffice instead. Indeed it is perhaps not too farfetched to imagine that the linear coding-strip (RNA or PNA or whatever) might serve as its own switchboard or computer-chip, even "gating" signals which "attempt" to travel along its own length — signals in the form of phonons, or electrons, or (more plausibly) some patterned combination of these — a matter we shall return to briefly in subsections 9.4 (3') and 9.5 (i).

But my immediate point is that, if such mind/memory encoding is held as linear molecular codes, then the mechanisms for accessing it need not be any more complex than those needed for constructing proteins — and indeed they could be *simpler* as we have just seen. Hence there is a *prima facie* case for taking seriously the idea of a molecular-basis for mind.

5.5 But what about Lamarckian mechanisms, if any?

(4) If mind or memory are matter-based, then they will surely entail a material encoding regardless of whether the encoding is achieved by Lamarckian or Darwinian techniques. So let us provisionally assume that any notions we may have about linear-molecule encoding will apply comparably in both cases. Any differences then will presumably reside in the read/write mechanisms, or selection processes, for these codes:

(5) As already mentioned, one serious difficulty with Lamarckian learning is its need to assemble perhaps-unruly materials under the pressure of "real-time" urgency. This

aspect — but then he is not concerned about questions of its physical embodiment.)

[55] It might have been more parsimonious to class such mechanism structures as further *coding* — for arguably all meaningful structure is a sort of self-coding (as hinted at in the last paragraph of section 5.2). However I doubt whether that would have made the explanation any clearer, so let's just keep to the existing dichotomy!

 — ISBN 0 9577737 1 4

Table 5 — THREE RIVAL APPROACHES TO ACQUIRING KNOWLEDGE (EPISTEMOLOGY)

	Lamarckian using linear-code	Darwinian using linear-code	Darwinian using synaptic system
Pre-existing coding needed ?	No, but desirable	**YES** ?DNA, ?RNA, ?PNA	**YES**, as arbitrary synaptic network ?
Pre-existing *clever* mechanisms needed	**YES** ribosome-like?	No, except as own inbuilt "chip-logic"?	? Presumably Yes; as protein ?

problem could perhaps be lessened by having materials already lined up ready for last-minute encoding — either by some sort of *switch-setting* for each "bead" along the "necklace" of linear coding — or else by some sort of *pre-ordering or sorting* of the ingredients so that they could be instantly accessed as required.

(For instance, for the sake of argument[56] we might consider a prefabricated "recording-tape" of RNA with a very long *repeating* nucleotide sequence of

AGCUAGCUAGCUAGCU-…, etc.[57]

In principle this could serve as a general-purpose ingredient — a queue of candidate nucleotides with about 75% wastage. Thus if the system required a new coding of UUGA, it could select and use those items in bold-underlined type:

AGC**U**AGC**U**A**G**CU**A** *GCU-…*

and discard the rest, leaving the *"GCU-…"* as a residue when it has finished. That *might* at least ensure that there would be no long waiting for materials; though it does not explain how the correct nucleotides would be selected, nor how the resulting code would be read back meaningfully when required.)

5.6 Comparing three rival systems

We are now in a better position to evaluate the above-mentioned *Principle of seeded knowledge-growth*, as it might apply in various situations; (see table 5):

> "Knowledge within bio-systems cannot start from a state of zero-coding, and so needs *some* coded information to get it going — to *bootstrap* it, as computer buffs might put it?"

Is that correct? Well the answer seems to be *Yes*, and perhaps obviously so (except to some supporters of the occult); but the *Yes* takes different forms according to the strategy thought to be in use.

For the unlikely Lamarckian case, strict pre-coding would surely be needed to construct the "recorder-head" mechanisms which would surely be at least as complex as ribosomes. In fact we might judge their specifications as rather too complicated and precision-demanding for us to find any comfortable natural explanation, or even devise any in-principle deliberate design for such a strategy. (We may also note that any such pure transcription system offers no facility for originality or creativity. So as a model for human capabilities it offers a very bleak picture — a metaphysical view which can be disastrous if applied systematically by political leaders.)

For the traditional axon-synaptic Darwinian systems, the situation is probably no better; though it is difficult to comment on how precision-tasks might be handled since I am not aware of any such detailed and comprehensive proposal. — (See *e.g.* Edelman, 1987).

For the Darwinian linear-code model however, it is arguable that the most essential codes *are* basically simple and primitive enough for them to have originated spontaneously by chance. Indeed they would appear to be very like the codings for life itself — life codings which are, it would now seem, just a rather special form of knowledge.

In conclusion then, I have argued that no event is strictly uncaused, so it seems there can be no *truly 100%* original thought; and what originality there is, will occur through trial-and-error selection processes which usually occur below our level of awareness. That verdict has implications for our freewill and autonomy — issues which we will consider in the next chapter:

[56] It should be obvious that I find it hard to believe in *any* Lamarckian explanation for memory encoding, and this example is no exception. However I am trying to play devil's advocate, and offer as plausible an account as possible.

[57] For those unfamiliar with the details of genetic-code theory, Asimov (1963) is a good starting-point. The classical autobiographical account of how the structure of DNA was discovered by Watson, Crick, and others, is given by Watson (1968); and a more feminist-orientated account is given by Sayre (1975).

 — ISBN 0 9577737 1 4

6. RECONCILING "SPIRIT" AND "HARD-SCIENCE"

6.1 *Reconciling our views on* **freewill** *and* **determinism?**

Any purely mechanistic interpretation of my mind seems to deny that I have any freewill; and it likewise seems to assert that all my mental activity is totally driven by external causes, leaving me as a mere pawn of outside forces in every way. But how true is this deterministic impression? Would it relieve me of all moral responsibility; and how relevant is this issue anyhow?

It seems to me that the mind-causality problem has three layers:
(i) Mind-events which are seen as *genuinely uncaused* thanks to quantum indeterminacy, or for some more occult reason;
(ii) The problem of how to classify any postulated *hidden causes* which could never be traced in detail experimentally (see section 6.4 below); and
(iii) *The relevance of* "*clockwork*" *models* in which every event can be traced in full detail.

6.2 *Clockwork models and determinism*

Let us look first at the latter — the Newtonian clockwork model. That is surely *mechanistic* in every sense, and *deterministic* too if the machine is ideally obedient to the relevant laws. That is the sort of model which we usually associate with the word *"reductionism"* — the notion that any system can be seen as made up of irreducible components, like eternally ideal cogwheels and springs, or planets in a Newtonian Solar-system. Certainly we can think of such systems as reductionist. But is all reductionism necessarily like that?

Here we might distinguish between such *crude reductionism* and the more sophisticated types I shall return to shortly. Let us take it as the implied defining characteristic of crude reductionism that the "basic elements" of any such system will be supposedly eternal items (like ideal cogwheels or planets or diodes) — items which in reality do have a substructure and a degree of mortality, but these realities are assumed away. For many practical purposes, such assumptions are perfectly reasonable; but it is always dangerous to import approximations into a *theoretical* argument without explicitly acknowledging the liberties that have been taken.

Such approximations are reasonable when the "cogwheels" seem *stable*. That means, firstly, that any minor perturbation within them will simply die away and never become manifest at the more macro level which concerns us — in other words our stability requirement surely demands that any substructural dynamics will be kept in check by its own local *negative feedback* or self-damping. Secondly, we must assume that the "cogwheels" *will not wear out at all* within the timeframe under consideration.

Given these assumptions, we do have systems which are deterministic, and yet which can sometimes surprise us with their "emergent properties" when the components are allowed to interact in certain well-balanced complex systems. One example is the ability of some computer-based networks to learn how to perform visual recognition tasks. Provided that we had a record of the exact digital input, we could re-run the test many times and get exactly the same results each time despite its apparent ability to innovate. We could also, in principle, trace through every tedious step in the process using pencil and paper, and get the same answer — perhaps after a bit of our own negative feedback in the form of error-correction! This then is clearly a deterministic system, and yet its pseudo-intelligence might be seen as impressive.

Pseudo-random numbers are also of this type. These are numbers which we can get the computer to generate quite simply, and they *appear to be* random.[58] In fact for many purposes, they are perfectly acceptable as random numbers, and yet they are nevertheless fully determined so that we will get exactly the same output sequence every time we invoke the same recipe, as in the footnote. This then borders on the edge of the paradoxical: Some outward appearance of freedom from control — a token of freewill — and yet actually pre-determined in every detail.

6.3 *Departure from this determinism*

Pure ideal randomness. This would be such that there would be no way of predicting the next number, not even in principle; and there would be no chance of doing an exact repeat run. No thought even of any imaginable antecedent input or mechanism which might affect the output; and that means we are talking about *uncaused* effects — a clear absence of determinism — something out of nothing! Such acausal effects may or may not exist, and therein lies considerable scope for debate; though we could perhaps avoid this controversy by devising explanations which do not depend on such postulated acausality.

Practical randomness. Suppose we retained the same general recipe as for the pseudo-random case, but now added some further input — perhaps regularly altering (say) one digit of the third number in some deterministic way according to minor fluctuations in the ambient temperature

---

[58] *One recipe for a series of pseudo-random numbers*: Start by listing a dozen six-figure numbers taken arbitrarily from the phone-book, but then keep this same list as a constant starting point on every test-run. Next choose some procedure of (say) adding the first to the last, discarding any overflow of the most-significant digit (so the answer will still have six figures), then "writing-in" the new number at the right whilst discarding the number at the left. — For each run, initially repeat this procedure (say) 20 times, but then start "harvesting" a pseudo-random digit each time, taking the most significant digit from (say) the right-hand number.

 — ISBN 0 9577737 1 4

outside. In a sense the output would still be fully deterministic, but we have now let some of the control slip out of our hands and so we can probably no longer fully predict or replicate the results.

The role of continuous quantities. If that new input had been today's date, or the current cricket score, then we could probably still manage to replicate our computer's "random" output. Dates and scores are conceptualized digitally, so there is no room for minuscule variations in their value which could lead to ambiguities in input value. (There may still be ambiguities about the exact timing of a score change, but let us temporarily define that away by using a digital watch.) However once we face the realities of genuine continua like temperature and time, and apply them as inputs to our deterministic systems, then we really do have to accept a loss of our control and predictability.

Note that such a system based on continuous quantities may still be rigorously deterministic in the sense that *IF exactly* the same input values were ever to be replicated, then exactly the same outputs would follow. That is important theoretically, and we should take full note of it. But of course in practice we could never get a true replication here because of the infinity of almost-the-same possibilities which would produce quite-different-results, at least in the long run, as Chaos Theory now tells us. And yet this high instability in outcomes should not obscure the abovementioned theoretical point that the existing system may still be fully deterministic — its path fully laid out in full detail, *even though we have no way of predicting that path, nor replicating it*, since even the slightest deviation will lead to a *different* deterministic system; and of course such deviations would be absolutely unavoidable.

6.4 *Hidden causes, and reductionism*

This brings us back to the "case (ii)" mentioned in section 6.1 at the start of this chapter: "the problem of how to classify any postulated *hidden causes* which could never be traced in detail experimentally". We have just seen one likely source of hidden causes — deterministic happenings at some ultramicro level which turn out to have macro effects — micro-variations which we could not possibly measure, (and we probably could not properly process such data independently anyhow) — invisible micro-causes whose effects escalate upon themselves in a run-away positive feedback which does burst into our everyday visible world.

A second suggestion is that *supernatural forces* drive the emergent activity and defeat our attempts to predict or replicate. However, as was suggested earlier, near the close of section 4.1:

"… supernaturalism itself is not the real problem — that *certain types of supernatural system* could, if they existed, fit in quite happily with scientific systems, …"

The point was rather that we should decide whether there were *causal influences* of any sort (whether we understood them or not), and whether they had *structure* of any conceivable sort. Thus if Zeus and Athena existed and chose to intervene, they would clearly be causal influences and it seems likely that they would also have some sort of structure, even if were bizarre by our standards and hidden in some strange *n*-dimensional space.[59]

So I find it hard to see that any such *supernatural-versus-natural* status really changes the problem of explicating causality:

— whether reality (of all existing sorts[60]) could be *strictly deterministic*,
— whether *freewill* really exists (naturally *or* supernaturally),
— and whether determinism and freewill could be *compatible despite appearances to the contrary*.

Varieties of strict determinism

I think the point that is beginning to emerge is that "strict determinism" can mean at least two different things according to the type of system envisaged:

(A) *Digital-and-comparatively-macro*. Here the determinism is assumed to apply within the designed system itself with its well-nigh immutable components, but not to the messy "indeterminate" environment outside. So we are talking about systems with rigidly stable cogwheel-like components (see above), a relatively limited behavioural repertoire so they will be fairly predictable; and it will also be possible to obtain reasonably faithful copies of them.

That also sounds like most of our technological devices — artefacts which we like to think we (collectively) can still control. Such physical systems sometimes replicate very limited aspects of brain activity (like visual pattern recognition), but they fall far short of offering any comprehensive model, despite such interesting emergent properties. Finally, and significantly, **this seems to be what critics have in mind when they speak disparagingly about anti-human *reductionism* as a poor model for biological systems**.

(B) *Rigidly determined by fine detail in the continua which underlie reality*. As we saw above, such continuum features are inherently beyond our powers to measure or process in any full way — so we poor mortals cannot determine the outcomes, and in our self-centred way we are tempted then to assert that the system is indeterminate. (Indeed to a pre-Kuhnian empiricist, anything which defied measurement must have something non-existent about it, by definition! — and we will come back to this below in relation to the quantum theory). However it now seems clear that we have no full justification for asserting such indeterminacy even if

[59] Please note that I see no actual merit in believing in such supernatural arrangements; I am merely saying that if one does happen to believe in this sort of influence, then that is not necessarily incompatible with science. In general though, such hypotheses seem to be superfluous, adding nothing of explanatory value apart from simply shifting the problem from *A* to *B*; so I would advise the use of Occam's razor, and tidy such interesting-but-obsolete ideas away into their historical archives. However, like anyone else, I could always have got it wrong — or half wrong!

[60] including supernatural, if need be!

 — ISBN 0 9577737 1 4

we suspect it — so we should at least allow for the possibility of rigorously deterministic systems like this.

This would tend to put us in something of a double-bind. On the one hand it offers us the disquieting notion of predestination — seemingly the very antithesis of freewill; yet on the other hand it offers to break the constraints of conventional reductionism with its "cogwheel" components, and allows us a new and deeper reduction with new possibilities. So then, can we reconcile these two views?

6.5 *Quantum theory*

Some, like Eccles (1986,1989) have invoked the peculiarities of the quantum theory, and I shall consider it too, though I should first point out some misgivings. For one thing, it has a certain mysticism built into it, ostensively to explain wave-particle duality and also why complete measurements can never be made; but it has long remained framed in such a way that it does seem to offer an "escape-clause" on our present topic of freewill. Maybe all this is above board, but we should at least be aware of ideological agendas which may have biased doctrines in this area.

Bohr, for instance, was influenced by the theological ideas of Kierkegaard and Høffding (see Honner, 1994, p142; Röseberg, 1994, p341-2); so maybe one should think twice before using Bohr's physics to support modern theological theories. That could just turn out to be one of those circular arguments which may look impressive, but actually get us nowhere.[61]

Bohr's "Copenhagen convention" left the unmeasurable aspects of ultramicro phenomena unexplained, or rather disclaimed as non-existent. That seems to mean that the path of (say) an electron emitted and then detected one metre to the North-East, had no sufficient cause to direct it there — no adequate cause at all, it just somehow happened. This denies any hidden variables or causes, and instead espouses a causeless or *"acausal"* effect.

Eccles effectively offers an alternative interpretation. He ascribes such missing causes to supernatural influences — and in particular, he sees a supernatural mind exerting its influence on the bodily brain as the hidden quantum influences taking place at neuro-synaptic junctions. In effect this is claiming that the supernatural operates through the unexplained aspects of any "true" random generator.[62]

So, having already looked at cases (iii) and (ii) as listed at the start of this chapter, we now find ourselves addressing case (i), which considered the possibility of "Mind-events which are seen as *genuinely uncaused* thanks to quantum indeterminacy, or for some more occult reason" — the Bohr solution or the Eccles solution.

[61] Unlike "spherical arguments" which may be rather more valid as paths towards the truth; see page 6 within section 2.4.

[62] That idea is not exactly new. One reason for some religious opposition to gambling is the belief that it is a profanity to use God's power-over-randomness in this way.

Unfamiliar causes, posing as "uncaused"?

In fact I do not see the Eccles formula as genuinely acausal.[63] *If* his view did happen to be correct, then surely all quantum effects which he deemed relevant would then be controlled by supernatural components of mind such as one's personal spirit or soul, or more directly by other supernatural forces. In short, they would be *caused* — and these *causes* would still be of potential interest to theoretical scientists, notwithstanding their otherworldly nature. (Of course this does shift the problem into a more inaccessible domain,[64] and we would then have to cope with the question of whether there might be acausality in that domain. This would be less than helpful, but I do not see it as materially altering the original materialistic puzzle, *as seen theoretically rather than empirically*).

That then leaves us to deal with "true" acausality, whether linked to the Bohr-Heisenberg indeterminacy or anything else. The remaining issues then seem to be whether true acausality does-or-can exist; and even supposing it does exist, can it then help us to understand freewill?

Of course I cannot *prove* that all happenings have causal origins — that every effect must have been triggered by some sort of preceding event or presence — and yet I find it very difficult to see how any other course could be possible. Such genuine acausality would require the spontaneous creation of something out of nothing at all, and with no help from anyone or anything — not even divine intervention! Moreover it is well known that the apparently-uncaused results of quantum activity do fall into well defined statistical distributions, and that does at least seem compatible with some hidden organizing principle — something which is difficult to see as being anything other than a cause! This is not conclusive either way, but I would suggest that it is up to the current supporters of acausality to bring their arguments up to date if they are to convince us doubters.

But even suppose we were to accept some acausality as genuine, what then? How would that explain any freewill I might have? Take the hypothetical case of my having some *totally* original thought (if that were possible); that new thought would presumably come to be encoded somewhere within the matter of my brain — but just where would it have come *from*? If it came from nowhere at all, as an acausal event, then how could I take any credit for it? If, on the other hand, it came from my mind/brain (be it material, or structured supernatural, or some combination of both) then it would surely have been *caused*.

As far as I can see, the only remaining slight hope of rescuing acausality (despite this dilemma against it) would be to picture some part of the mind as being neither material

[63] We may not share his quantum-source vision anyhow, but that is not the point.

[64] In fact, it will be argued below that it is the postulated *molecular-IR system* which is in a position to usurp this power — the mystique that Eccles attributes to the supernatural. However this still leaves us in a comparatively inaccessible domain, where thorough experiment and observation may well be impossible.

nor structured-supernatural but rather some sort of unstructured and ineffable spirit such as that considered earlier in section 4.6 (though *even then* it is not entirely clear that this would allow for acausality). Of course there has been a long mystical tradition of viewing the mind as an unstructured spirit in this way; but if true, it is difficult to see any way of reconciling this view with any detailed scientific approach.

If all else fails, we might be forced to come back to this position. However I believe we can now **do better by pursuing causality into the fine structure which is in plentiful supply at molecular level**; and of course the present project is directed toward that end, aided by such concepts as hierarchy of control, trial-and-error, and the Piagetian "scheme".

6.6 Freewill despite determinism?

So where does that leave us on the issue of *freewill in a deterministic universe?* Suppose the universe is strictly deterministic in sense (B) as discussed above on page 30 — highly sensitive to slight variations in the continua, but not admitting of any "outside power" which could make those slight exogenous changes, and so actually set on an unshakable pre-set course despite the sensitivity. Does this also allow us to have freewill in any acceptable sense?

Let us put the question another way. Suppose we could someday produce realistic humanoid beings within the virtual reality of some future computer system; and suppose also that these beings were capable of making subjective decisions within their own virtual-reality domain. The big questions would then be: Might these beings *believe* that they had freewill? In what sense could that be true? And to what extent would they be deluding themselves?

This is a matter which deserves some debate, so it would be premature to try to come to any firm conclusions at this stage. However my tentative suggestion is that there have to be two concurrent answers — one from within the virtual domain — and one from any pure-observer (a godlike computer operator outside that virtual world; someone who does not intervene[65], but only observes in a totally non-invasive[66] way).

To the godlike observer, the virtual beings do not really have freewill. In principle the whole scenario could be re-run, using the same pseudo-random series of numbers, and the virtual beings would then behave in precisely the same way in response to exactly the same circumstances.

Freewill an illusion perhaps, but impossible to test?

However the virtual participants could never be in the position of studying all the factors influencing their decisions; and if they were truly humanlike, some of these factors would be tactics learnt *unconsciously* by their species long ago so that their ability to use these tactics would seem like some magical virtue of their own. If all such factors were then added to what they had *consciously* learnt, they might well have a self-image as beings with a large degree of autonomy. In other words they would see themselves as having a fair measure of freewill — and more to the point, their peers within the virtual domain might well see them that way too, and thus bolster their self-image. Perhaps indeed, that is all we require of our subjective feeling of freewill — as some sort of social tool or status for our dealings with our fellow beings? And anyway, who are we to say that this is not freewill within the limits of that domain or culture?

So much then for that virtual world of cyberspace; but what about our own situation? If there were any *totally non-intervening divine observers*, fully aware of all our circumstances, they might well see that all our activities and thoughts were strictly-speaking predestined — as if run on a deterministic computer system. But then, by definition, such godly observers can never share that insight with us (and we may well doubt whether they actually exist anyhow).

Surely then, for all practical purposes, we may allow ourselves to believe that we do have at least some freewill — and do so secure in the knowledge that no other mortal can possibly have the omniscience necessary to clearly demonstrate any fallacy in this.

So, speaking for myself, I will continue to live my life in the belief that I do have freewill on at least some important matters (in line with the Arts-and-Humanities way of thinking). But at the same time I am prepared to accept determinism in our fundamental mechanisms *if* that is what seems to be technically called for; (a sentiment which is more in line with the hard-sciences). — *Yes*, I do want to have it both ways; and indeed I suggest that this sort of reconciliation is essential if we are to get to the bottom of such matters.

In other words, here is one possible **solution of our metaphysical dilemma** on whether to believe in the world-view *of hard-science*, or that *of the poet* — and this first answer is that we accept both, in most respects. The price we pay is an acceptance that our claim to freewill could not convince any omniscient godlike observer. However we do not often meet such august beings (and anyway they would not be able to talk to us without becoming part of the system themselves); and meanwhile no mere mortals would be omniscient enough to reliably debunk our pretence — not our colleagues, and not even we ourselves. So there seems to be nothing to stop us believing in at least some freewill at the practical level, despite the apparent predeterminism of the system.

[65] As soon as our godlike observer *does* intervene, the situation changes radically. The two domains then effectively become *linked into one system* — with the observers being effectively in a "supernatural-but-reachable" domain. But as we have seen, any such interactive supernatural domain can be seen as compatible with the science of the ungodlike "natural" beings, if they have wit enough to make sense of its clues.

[66] This would even forbid the interference-due-to-measurement which underlies the Heisenberg indeterminacy principle. However that need not be a problem within a virtual reality domain, as the programmers could presumably have independent access to all such data — access within their own realworld domain, and leaving the virtual-world parameters quite untouched.

 — ISBN 0 9577737 1 4

Not-quite-strict determinism

As a second variant on this theme, we might prefer to just slightly relax the supposed strictness on causality and determinism. As we know from chaos theory, it only takes the slightest variation in the causal network, and the results within the system may soon become irrevocably different from what they would have been — and that would presumably get rid of the embarrassing hint of predestination. But that then leaves us with a different price to pay. We now have to accept some lawless physical behaviour — some fraction of events as happening without any cause whatsoever — as *truly* random happenings.

We may well be prepared to pay that price. After all, quantum theory has by now conditioned physicists to accept a certain well-defined degree of ignorance about micro-systems, and also to accept that this ignorance actually betokens acausality — that there is no cause there to measure, so we can hardly succeed in measuring it. But having accepted that price, does it really help us to bolster our belief in our freewill and autonomy?

I repeat the question raised above. If my beliefs and decisions are ultimately driven acausally — by nothing-at-all — then how can I personally identify with this nothingness, and accept the praise-or-blame which freewill would normally imply? In other words is this pure randomness really any closer to offering freewill?

A compromise mixture?

But there is a third variant on this theme; one which draws on both of the others. Consider a two-stage process in which the encodings for *potential* decisions are generated abundantly by a more-or-less random process, but only very few are chosen for actual implementation. This secondary choice process could be governed by some *deterministic formula peculiar to the individual concerned*; and, seen socially, that individual might then be praised or blamed according to the consequences of his-or-her decision.

Notice first that the initial random generation of ideas does not need to be particularly sophisticated. It *could* use a generator of pure randomness if such were available, but it could usually work just as well with a pseudo-random source, or maybe sometimes even a carefully ordered systematic scan! (The variation on which Darwinian selection is based, *could be* generated by any one of these three approaches. It hardly matters which.) So there is no pressing need for acausality here.

On the other hand, strict determinism need not be a bar to evolutionary development, despite appearances to the contrary! We can see this if we model Darwinian processes on a computer (Dawkins, 1986) — a procedure which can operate using a particular pseudo-random number series, so that we get identical results each time — or we can use *different* pseudo-random inputs and get different detail, but a common tendency to look like real-life evolution in its various forms. The important point seems to be that the deterministic part be kept in the background as "mere" substructure[67]. In practice then, that seems to allow us to have our cake, and eat it too!

6.7 Conclusion: Reconciliation and other benefits from rethinking the fundamentals

• Freewill and determinism seem to be poles apart.
  But then also:
• The humanities often seem just as remote from the world of hard-science.

This chapter has tried to grasp the nettle of explaining why these two emotive divisions might exist; and in fact I have sought to show that the two conflicts are closely interrelated.

This overall issue of *freewill-versus-determinism* is no mere philosophical backwater. It is an issue on which people often have strong opinions; though they may not often have those opinions challenged, so the topic may not usually reach conversation level. The point is that it is important to our self-image to see ourselves as effective agents in this world, and we can hardly do that if we cannot believe in our own freewill.

For that reason then, I suspect that this is *the* most important point to be resolved before we can solve the mysteries of the mind in any generally acceptable way — more important than working out nerve-connections or their chemical attributes, or any of the other issues discussed in the various volumes of this project.

Of course I am not actually denying the importance of these latter main issues, but there have been two incompatible perceptions of them — the humanities versus hard-science. — Both sides have important contributions to make, so their mutual incomprehension has tended to lead to a stalemate which has somewhat eclipsed the good work done in both camps.

The idea of reconciling such public differences has had a long history, including such names as Hegel (who was all too mystical about it), and the modern advocates of the *win-win* collective decision-making within Conflict Resolution. Likewise, at the level of explaining perception within the individual, Piaget has often pointed to the need to "decentre" from unduly prolonged focus on this-or-that feature of a situation, so that one can re-focus on the wider general view. Or to put it another way, perhaps it is always

[67] or *sub*-substructure within that, etc.; and of course that implies some sort of *hierarchy*. This hiding within the lower depths of a hierarchy invites an intriguing comparison — between nature unintentionally hiding a "guilty" secret about determinism, and a fraudulent businessman deliberately doing the same sort of thing to cover up his shady dealings. He will no doubt create a tangled web of opaque transactions, each removing the crucial evidence further from scrutiny, and ending up in something like a numbered Swiss bank-account — something which surely does have a meaningful internal structure, but which we poor mortals have little hope of investigating.

 — ISBN 0 9577737 1 4

worth trying hard to dissolve a dualism by penetrating deeper into the reality of the situation in the hope of reaching a much more comprehensive monist overview.

So, to what extent have these recent chapters succeeded in this bridge-building goal? That is not for me to say, and maybe it is too early to tell anyhow. But I feel that it was worth making the effort; and wherever I may have failed to provide a convincing reconciliation, I hope others will fill the breach; for surely this interdisciplinary cold war is a major obstacle which should be laid to rest if at all possible.

7. TWO PATHS TOWARD STRUCTURED UNDERSTANDING

7.1 *BOTTOM-UP, especially in simple systems*

Dynamic systems which interest us will almost invariably be organized hierarchically, whether or not that is evident to the observer. Where should we start looking at such systems if we want to study or analyse them? Our everyday life is likely to throw us into contact with the chaos of the middle sections; but if we are really looking for understanding rather than crude involvement, we might well consider a different approach. In fact we would often do better to demote much of that chaotic "middle-level" experience as being merely impressionistic. That then leaves us free to start afresh with a new strategy — hopefully one based on elements which are more orderly and homogeneous.

One solution is to re-start with the extreme viewpoint of *the whole system* — perhaps seen as having human qualities, even if the system is actually inanimate[68]). Clearly this would be the top-down approach.

Or we may take the other extreme of focussing on the most basic building-blocks of the system and their immediate interactions — the *bottom-up* approach, and that is what will concern us right now:

First we should note that we are usually happy to reduce *molecules* conceptually into their constituent *chemical-atoms*, the building-blocks for most practical purposes other than particle physics. Here we are in the domain of a typical "hard-science" where we may see structure-building as analogous to bricklaying or bolting together struts and plates in scaffolding. Even here though, this reductive notion is not without its problems (quantum ambivalence, emergent properties due to interaction, the need for self-organization, and suchlike); but reduction clearly still makes sense, and moreover we can also visualize clearcut structures in these terms.

Moreover, given discrete elements like this, we can then visualize how the resulting bigger groupings like benzene rings might serve as units on a larger scale — and how such hierarchical nestings of unit-assemblies could conceivably go on for level after level. Each would have "units" bigger than the last, and potentially offer an orderly structure even up into the macro level of our everyday experience and even beyond that. Although conceivable, such a recursive process would seem farfetched if it were not for two further thoughts:

• Whatever the explanation, such structural hierarchies *are actually found* in the real world! — and: • Such subsystems are often stable in a way which we might not have expected, and this local stability in a dynamic system is probably due to some sort of *negative feedback* or self-reinforcement — some fortuitous tendency of the subsystem to automatically cancel any trend which might have caused it to fall apart.

(This then is the opposite to the high instability involved in the "butterfly effect".[69] Moreover it is not too surprising that we *do find* this self-stabilization in many existing subsystems. Perhaps many other *unstable* groupings will also occur spontaneously, and maybe much more often; but then, being unstable, they will not last for long. Typically they will simply self-destruct due to their own arbitrarily-imposed *positive* feedback. In short then, we have something which looks very like Darwinian trial-and-error — another setup where we seldom see the really "bad" examples because they have already mostly perished after a comparatively fleeting existence.)

Such a structural hierarchy will presumably therefore be quite common in various forms; though the following account maybe overstates the idea, querying the notion of any reachable top or bottom!

> Great fleas have little fleas upon their backs to bite 'em
> And little fleas have lesser fleas, and so *ad infinitum*.
> And the great fleas themselves, in turn, have greater fleas to go on;
> While these again have greater still, and greater still, and so on.
>
> *De Morgan (1872; v2, p191), after Swift.*

Unfortunately for us investigators of real-and-finite systems, the further up the hierarchy we go, the *messier* the bottom-up situation becomes for any onlooker. The repertoire of feasible substructures becomes disproportionately

---

[68] As scientists we may wish to shun such anthropomorphic projections, but it is worth noting that this approach to complex systems is part of our legacy from our evolutionary past, and we actually do it rather well (as far as such rule-of-thumb procedures go), and it does give us one practical way of coping with complexity which we may well have to use whilst we are trying to devise something better. Of course it probably works best in helping us to deal with other humans — the task it was presumably "designed" for.

[69] *Chaos Theory* tells us that even the minutest difference between models can result in profound differences in results — the so-called "butterfly effect", where (for instance) some trivial difference in an insect's activity can trigger-*or-not-trigger* some devastating storm on the other side of the world, though in an unpredictable way.

It *can* produce this greatly magnified result, though it is *also possible* that the effect will die away in an approximately exponential way, and never trigger any further observable difference at all. But as we will often have no reliable way of predicting which way things will go, and as there will usually be many such perturbations, we should therefore expect the worst, an average growth in the unforseeable effects as time passes; and hence we should recognize the impossibility of reliably predicting the course of complex systems in the more distant future.

On the other hand, there *are* some occasions in which we can count on perturbations dying away, and we will see their importance later in relation to self-stabilizing systems, including living beings. However it is probably fair to say that this damping-down of perturbations can never be totally counted upon within any grand-modelling system such as we are looking at here — given that we can never expect to understand it fully.

 — ISBN 0 9577737 1 4

larger[70] as we ascend the ladder, so obviously we can then only become less certain as to *which particular* possible stable substructures are actually present.

Perhaps that then gives us a formal justification for the intuitively obvious: that analyses based on reduction and structured hypotheses work best for simple systems (with their fewer choices, as in physics), and worst for social systems which may well be just as structured in their ultimate makeup, but whose immensely complex *secret details* will usually be quite beyond us, and constantly changing anyhow!

Is it applicable to complex systems?

So then, what are our prospects for understanding complex systems if we start from their *elementary components* — using a *bottom-up* approach? We may perhaps compare the situation to navigation by "dead reckoning", where you know your starting-point and then keep measurements of your speed and direction to calculate your subsequent position — but with no independent means of checking it. In practice of course, this turns out to be a rather chancy method if you really do have nothing else to go on.

Likewise we can envisage starting with our basic units (atoms or whatever), and then conceptually building these up into various theoretical structures, in the hope that they will accord with genuine possibilities in the real world. Often we will succeed, especially in the simpler cases; but the increasing effect of unforseen interactions and feedbacks will make such success ever more rare as complexity increases — unless perhaps we can draw on independent evidence to apply corrections as we go.

There is a further complication too. The traditional mariner had only to explore the two-dimensional surface of our globe, so any choices about future direction were kept within bounds. Not so for anyone dedicated to a rigorous development of theory from a fundamental-elements base alone, for here the choices are legion, as if for an explorer of a hyperspace in which any direction had to be specified in (say) 20 dimensions! Small hope then of ever reaching our destination (even with impeccable dead-reckoning) unless we also have a pretty clear preconception of what we are heading for. In our present context of seeking to understand the mind/brain, this suggests that we will also need a fairly clear idea of how the whole system works — as seen in overview using a *top-down* approach.

The point I am leading to here is that we will almost certainly need the help of *both* approaches (bottom-up *and* top-down), and need them at every stage of a complex problem like brain-explanation. Moreover this dual approach will probably be needed *within each stage* or facet of the problem, and not just for the problem overall — bearing in mind that the system is likely to entail hierarchies of substructure as we have just seen. We surely need some concept of the underlying "building-blocks"[71], for otherwise it is difficult to understand how nature could have arrived at any sort of *stable* structure — or any sort of material hierarchy — or any feasible interfacefor memory and signals.

Furthermore we should not forget our problems of *modelling* such systems, either within our brains, or through human-made artefacts (computers, formulae, or mechanical models). In each of these we depend on having some basic entities upon which to build — either models of objectlike things, or representations of behavioural or relational regularities, as in formulae. Such entities are surely there to be found within nature itself; but even if they were not, we would probably have to invent them in order to get some sort of mental grasp on whatever we wished to think about. In fact we clearly do impose such arbitrary boundaries and invented-structure (like "star constellation", or "kosher food", or "deserving poor", or "safe") on our notion of nature whenever we have nothing better to go on, sometimes with unfortunate results. Far better then, to be aware of the genuine substructure *if* we find it feasible to obtain such information, or at least to make an informed provisional guess about it!

7.2 TOP-DOWN *postulates: tentative semi-solutions for macro systems*

If we are studying the human mind as a holistic going-concern, then we may well detest the reductionist bottom-up approach, for the reasons we considered in section 3.1: *"After all, I am much more interested in people as such than in the details of what such-and-such a Purkinje cell is doing, or whether my memory about today's appointment is held by this-or-that substructure in my brain!"*

Such a view leads quite properly to disciplines like psychology or novel-writing — disciplines which tend to concentrate on how *whole people* (or *groups* of people) behave and think within a given environment over time. As we have already seen, it just so happens that most of us are quite good at intuiting correctly about such matters; but that is both a blessing and a curse! The fact that we can do it at all, tells us a lot about our own natural abilities to intuit in a way which far outshines computers, and it is a very valuable asset in this uncertain world. Then too, the fact that most of us are quite good at it is probably a testament to the long evolutionary process that our forbears have gone through, somehow enabling us to intuit "probable truths" out of fairly garbled information. (This is of double interest to us here of course, because our primary task was to consider the mind/brain; but then, in these current chapters (2 to 8), we

---

[70] This will be due largely to the rapid extension of the permutations and combinations made possible as the system grows, though actually the tendency to form stereotyped (stable) substructures within the hierarchy will do much to reduce the otherwise-expected exponential growth down toward something more like a merely linear increase.
There will however also be other exposures to unpredictability as the hierarchy ascends: There will be an increasing number of candidate-possibilities; and that will mean more competition between them, as well as an increased opportunity for unexpected chance events.

[71] We probably need object-like building blocks even if they are just vortices in a continuum; see, for instance, Descartes' *Principles* (1644); and Lord Kelvin's smoke-ring model for "atoms", (Thomson, 1867).

 — ISBN 0 9577737 1 4

are also studying *our own thought processes* during our scientific investigations).

On the other hand, this very ability may make us somewhat complacent about our own abilities. Merely being able to do something will not then mean that we actually know *how* we did it — and because we do not really understand the process, we have only a hazy idea about its limitations, and that can lead us into error; and worse still, into errors which we do not understand, nor perhaps even acknowledge to be errors. After all, our inborn abilities seem to be God-given and unquestionable, so we may sometimes fight to the death for beliefs which actually have a shaky intuitive validity, based on procedures which have "usually" worked in the past, either genetically or within a cultural tradition.

The status of human judgement

At least until about 1962, the promoters of the "modern scientific method" distrusted human judgement for this very sort of reason, though they did not exactly put it in these terms nor put their finger on the exact failing. Their solution was to react against such "subjectivism" in all its manifestations, and try to find a totally *objective* way of accessing reality. This seemed to be a worthy goal; but as we have already seen, it overlooked some inescapable dependence on the hated subjectivism, and so was ultimately doomed in its bid to be an all-inclusive philosophy, despite its partial successes in the comparatively easy task of the hard-sciences like physics and chemistry.

With the benefit of hindsight, we can now see that a better overall strategy would be to *accept the value* of intuition-like processes, but then properly investigate just how they work, and gain more insight on when we should be alert to their likely failure. That might perhaps play a useful role in the future development of science, and perhaps help to fill the vacuum left by postmodernist criticism.

Be that as it may, top-down theorists will mostly have great difficulty in finding any reliable substructure at all; so they will usually have to invent it or tinker with pre-existing notions about the substructure. Thus Freud invented concepts like *"id* and *superego"*; Keynes introduced the idea of *"marginal propensity to consume"* into economic theory (1936, pp.90 and 115); Newton proposed the strange idea of *action at a distance* (see below; and Hesse, 1961); and Piaget promoted the concepts of *"scheme, assimilation, accomodation"*, *etc*., ideas which will concern us from time to time throughout this project.

To some extent there is something unsatisfactory about all these postulates about substructure. At best they may enable us to make good predictions (very good predictions in Newton's case), so they are then useful *practical or clinical* tools. But that is not the same thing as giving us a thorough understanding of any fundamental submechanisms which may underlie these supposed attributes. We are missing the point if we criticise Freud's ideas here because they are not testable to the satisfaction of modernist theorists. Instead, in the present structure-seeking context, the weakness of the Freudian concepts lies rather in their *vagueness or "floppiness"*; so it is not easy to form any clear hypotheses of how a superego might be physically constructed, nor what its actual transactions might be.

Note that this does not altogether damn the *superego* notion. It might still be valid clinically even if it were no more than a placebo or useful myth (though such judgement is not our immediate concern here); but even if we look only at its contribution to a structured view of the mind with all its floppy inadequacy, it does nevertheless offer some sort of first bridge across a mystifyingly wide gulf — the gap between the macrophenomenon of people's overall behaviour or mental life, and the ultimate elementary micro-activity within their minds, whatever that may be.

Perhaps that gap can never be bridged properly; but if we are going to try, then we have to start somewhere. In this case we have virtually no hope of coming to a well-structured solution (even hypothetically) all in one go; so instead our bridge must have one-or-more provisional piers in mid-stream (like the railway bridge across the Firth of Forth, as mentioned in *Book A*). Maybe these piers will ultimately be found seriously wanting and will have to be replaced, but hopefully we will have gained some useful experience in the intervening attempts; and anyhow, I repeat that we have to start somewhere.

Newton's action-at-a-distance led to the inverse-square law for gravity and other related concepts of great practical importance. But for all that, as already mentioned in *footnote 1* and in *section 3.3 (2)*, we do not really understand what gravity is, even today! Certainly we can now make very accurate predictions (now using relativistically-improved formulae), but merely possessing a good analogy called a *"formula"*, does not in itself mean that we understand fully what is going on. So once again we have really only reached a midstream pier, and not the true opposite bank. Newton himself realized this:

> "For I here design only to give a mathematical notion of those forces, without considering their physical causes and seats." (vol.1, p.5).

We may call these forces "gravity", assign formulae to them, and even talk of gravity waves and gravitons; but like much of modern physics, gravity is something which somehow lacks any *psychologically satisfying*[72] substruc-

---

[72] Here the intended implication of "psychologically satisfying" may be approached from two different-but-related angles: (i) as implying "coherence" as discussed at length elsewhere in this project, notably sections 8.2 - 8.3 below, and *Book A* (Traill, 1999); and (ii) the notion Lord Kelvin is reputed to have supported, that an explanation should (at least in principle) be amenable to being modelled visually using already-understandable concepts — a notion sometimes expressed by the German word *"anschaulich"*. This latter has a hint of circular reasoning about it, but that simply serves to associate it with the former "coherence" concept, as discussed in section 8.2 and in *Book A*. Of course there can be no guarantee that nature will always operate in such an *anschaulich* manner, though there are reasons for expecting that it usually will (since incoherently organized systems are likely to disintegrate promptly and therefore not come to our attention).

 — ISBN 0 9577737 1 4

tural explanation. Many will say that there *is* no intelligible deep structure to be found or guessed at (since there probably has to be a limit to lower levels of substructuring, and that fundamental physics is the discipline where such a limit is likely to turn up). This physics-of-matter problem might thus be a special case which need not accord with the other examples. If it is indeed an exception because it is an extreme case, then we may well choose to accept it as such and not pursue that point further here. In general though, it seems reasonable to expect that meaningful substructure will actually exist, even if we do not yet have any idea what that might be.

The special case of Piaget's "scheme"

Piaget's formal concept of *the scheme* (in its more general sense of a shadowy pseudo-mechanism within the mind) looms large in the present project. This concept too is a "midstream pier" halfway across the gulf separating macro from micro — halfway between your activity as a *whole-person* and the *fundamental physical elements* of your mental activity (whatever these may turn out to be). However in this case there does seem to be a plausible possibility of linking this central pier to both sides of the gulf — of (1) stretching the bottom-up approach just far enough for us to link the visualisable micro-elements to "the scheme" as substructures underlying it; and then (2) using this clarification to link schemes to human behaviour and thought in a more clearcut way than as been possible hitherto.

In other words the scheme-concept, once its own sub-structure is suitably explicated, will then be a much more powerful tool for explaining the macro aspects of the mind, thus potentially tidying up and clarifying Piaget's account — an account which already exists of course, but suffers badly from abstract and "floppy" subconcepts.

This present volume will be mainly concerned with the bottom-up, or "hardware" aspects of this problem, "(1)" above. Meanwhile the issues of "(2)", the resulting improved explication of Piaget's epistemology and psychology will appear later in *Book D* as "software" or top-down treatments.

7.3 *On using both top-down and bottom-up — in combination*

There is something conveniently simple about an overview, even if it is rather vague and difficult to apply. Likewise there is something attractively simple about the elementary components of a system, even if we have no idea about how to fit them together. — A simple view from the top — and a simple view from the bottom.

In fact *a fair degree of simplicity* is just what we need if we are to understand our world; after all, we have only limited attention spans, memory is capricious, and life is short. Yet the real world is often far from simple. Reality as we actually experience it, frequently seems disturbingly complex both in science and in everyday life; and when once we do start trying to make sense of it all, we are likely to seek a manageable simplicity through either a top-down global philosophy like nationalism or positivism, or else through some bottom-up approach.

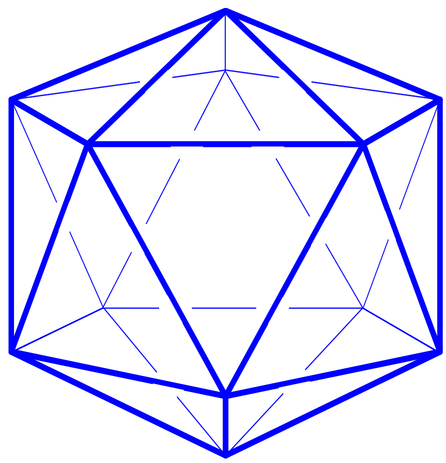

fig. 7:1

One particular crystal-like linked lattice (unrealistically tidy and regular!)

This might perhaps serve as a metaphore for the sort of combined knowledge-structure that we might end up with. At top and bottom it is a single point, symbolizing the comparative simplicity of each — though we should not infer that either is actually a single entity.

Top-down is certainly more likely to give us a comprehensive view, but it will usually lead to endless debate, with little hope of clear resolution precisely because of the lack of any recognized substructure which can fairly be seen as stable. Here we are surely talking about the Arts and Humanities. (As we saw in section 3.3, I am not complaining about endless debates as such. They are inevitable and laudable whenever there is no clear prospect of finding an all-embracing substructure, and maybe even then as well. However it is sometimes possible to aspire to some other "better" procedure which is perhaps a little more reliable, and less likely to lead to warfare.)

Bottom-up is the mainstay of the "hard sciences" like modern physics and chemistry.[73] When it can be applied

[73] Of course during the twentieth century there have been certain reservations about this, caused by the "uncertainly principle" arising from the quantum theory. And before that there was an earlier uncertainty introduced by the statistical basis for the gas

 — ISBN 0 9577737 1 4

rigorously, it works spectacularly well; and even when its validity is only partial (such that the "elements" or their interrelationships are only imperfectly reflected in our scientific models), even then there is often much to be gained from using this bottom-up approach, as long as we keep an eye on its limitations.

If we can find a way to combine both approaches, then we will be well placed to claim that we *understand the system thoroughly* — or as thoroughly as possible in the existing state of science. Here we will need some means for bridging the gap conceptually, and I have already likened this to having to build mid-stream piers to support the bridge except in very simple cases. Alternatively we might liken the connection to a lattice of connections forming a crystal-like pattern of lines, (figure 7:1)[74]:

This diagram is meant to depict the likely complex organization of *our concepts about* some real system out there in the real world. Note the reference to *concepts* — and not necessarily to the reality which those concepts purport to represent. The model might also just happen to depict the real world faithfully; but we could never know for certain, and anyhow that is not the point right here.

In other words, it is *not the real system* which is supposedly depicted in the diagram. The diagram refers instead to *our constructed conceptualization* of that real system, whether within the interrelated notions of our scientific community (our main concern at present), or within the thought processes of the individual (a matter which will concern us more in *Book D*, though it is also mentioned in *Book A*). figure 7:2 offers a rather more realistic diagram of the same idea

So then, here we have a schematic picture of the sort of relationship which might apply when we consider all the deep-structure aspects of our scientific *understanding* of some real system. As yet this depiction is poorly defined — merely a suggestive analogy — indeed one of those vague top-down "arts-discussion" conceptualizations which we discussed in section 3.3. However if we accept the diagram-limitations in that spirit, we can then see certain features — at least in the idealized case of *figure 7:1*:

(i) *a simple top-level* which provides a global overview about the real system, somehow based conceptually on a rather chaotic and complex mixture of sub-concepts found at lower levels in the hierarchy. Moreover these sub-concepts are also but vaguely defined. In short we have a fairly well-defined "house" built on a usable-but-unreliable bed of "sand" — the typical top-down situation. The power of this top-level position if it genuinely exists, is that it purports to tie together all the lower accounts and somehow reconcile their conflicting claims. Of course such perfection is probably asking too much, so we might have to settle for a less exhaustive criterion — a 55% passmark perhaps! And meanwhile we should encourage some level of debate to continue perennially; see section 8.4 (below).

(ii) *a simple bottom level* which offers elementary-and-universal concepts which are directly used by several differently orientated disciplines (five of them in this case). From there, the concepts arising from these immediate disciplines are on offer for other use further up.

(iii) *a complex of interrelationships in the "equatorial region"*, with no apparent prospect of unifying principles once one removes both top and bottom nodes; (and even some of the "midstream piers of our bridge" are only stable because of their support from *bottom* or *top* ideas). It is probably fair to suggest that this is the typical confusing conceptual world that any naive community is faced with. You might say that they enter this complex "building" at a central or "equatorial" level, all unaware of the high-rise above them and the deeply rooted warren of passageways below them, and it is only when they have explored *both* the basement levels and the penthouse levels *and the connections in between*, that they can really understand their world. Moreover the problem is even more perplexing in cases like the more realistic one in *figure 7:2*.

Typically some sort of folk mythologies will be needed for a *top level*, to offer some direction to community endeavours. (At this stage there is little chance that such myths will be literally true, and that may lead to problems later; but one needs to start somewhere). With no possible access to the genuine *bottom level*, hard science too will get off to a slow start — beginning with a practical knowledge of materials and such primitive concepts as "the four elements" (earth, air, fire, and water). Thinking based on such orientations will obviously have its problems, but historically we had no option but to start at the beginning and learn the hard way.

Nor are we by any means out of the quicksands yet. Without *any* top-view, we would have no idea where we were going, nor what the conceivable alternatives might be; and that maybe sounds a bit like some aspects of today's politics and economics!

As for the bottom view, we are by now certainly well informed on many aspects of physics and chemistry; but we are much less clear about the underpinnings of certain biological processes. In particular I have been suggesting that our ideas about the fundamental "units" of mental activity need a major rethink. Accordingly, as already mentioned, such bottom-up "hardware" issues will be the main topic in this book from chapter 9 onwards.[75] Meanwhile *Book 3* will concentrate on the "software" issues which are orientated towards the top-down approach).

laws. However these two uncertainties are both trivial within most everyday phenomena.

[74] This drawing is just a suggestive metaphor at this stage: an *aide memoire*, and *not* necessarily a reliable analogue. Nor is the exact configuration of the "crystal" of any intrinsic importance, though a regular shape like this icosahedron may be relatively easy to visualize.

[75] already previewed in *Book A*, and likely to be discussed further elsewhere.

 — ISBN 0 9577737 1 4

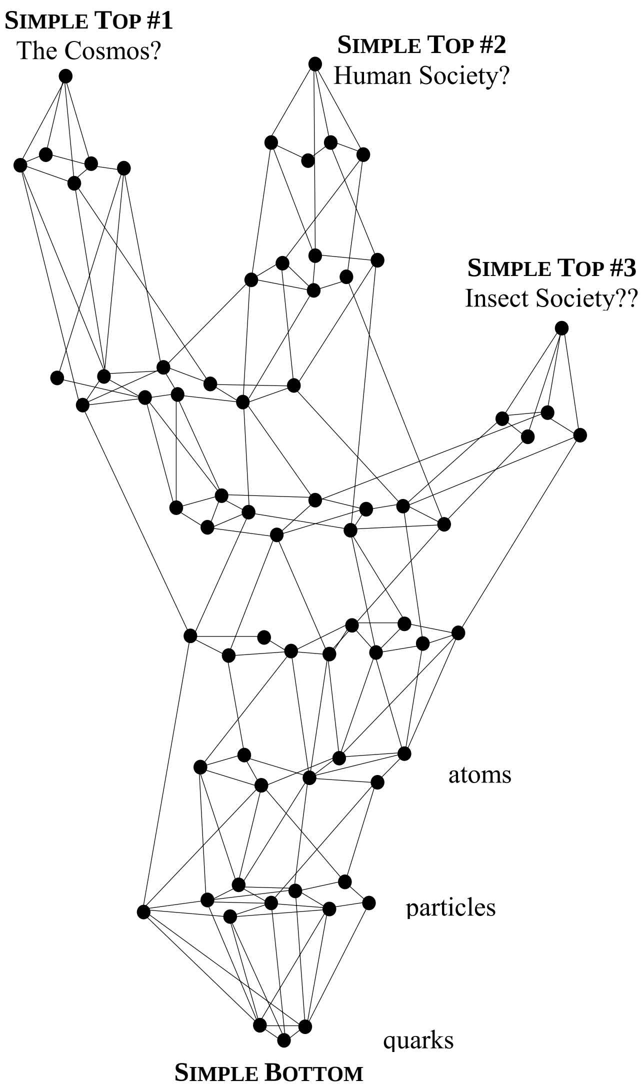

fig. 7:2

A rather more realistic schematic view of the structure and substructure of reality as we know it, but still vastly simplified and probably inaccurate generally.

(Moreover the rather slim central girth does not do full justice to the complexity midway between top and bottom).

Note that connections sometimes stray from the more obvious connections to same-level or next levels.

Finally we should not overlook the need for finding good linkages between our top and bottom conceptualizations. As most of our reality-experience occurs in this complex middle ground, it behoves us to keep it in mind for its own sake. But in any case, a bottom without a linkage toward the top will have no discernible direction (so its boundless possibilities avail us nothing); — and a top without any linkage to a structural bottom is “without form, and void”.

 — ISBN 0 9577737 1 4

8. UPDATING OUR METHODOLOGY

8.1 *Authoritarian versus self-organizing methodologies*

As I keep suggesting, the "modernist" version of scientific method makes certain naive assumptions about how we gain knowledge. These fundamental-but-unacknowledged metaphysical assumptions are then taken as the absolutely fixed rock on which to build. Obviously there must be some validity to this approach or it would never have had any appeal in the first place: It does work quite well when some aspects of the problem (like the nature of solid objects) are so comparatively easy that we almost invariably hit upon the right solution unconsciously. Such assumptions of the universal correctness of our geometrical beliefs will then often be close to the truth for many cases — but this is nevertheless a flawed procedure which can lead us seriously astray when we try to apply it elsewhere; and it does have a hint of authoritarianism about it.

Kings and dictators may sometimes have their uses when there is little to dispute about[76] as in the case of war emergencies; but such authoritarianism has a rigidity about it which does not adapt well with change — not even with changes brought about by its own success.

As already mentioned, philosophers are now well aware of such shortcomings in the still-current scientific method; but what well-formulated alternatives are there? Hesse (1980) offers two suggestions: *Coherence-seeking* and *Hermeneutics*. Let us consider them both:

[76] Concerning the role of kings, here is another interesting example of a misguided appeal to authority: "What…could we make of Copernicus' inference that the sun is at the centre of the planetary system because it is analogous to the king at the centre of his court…?" (Hesse, 1980; p.14).
Of course Copernicus was actually wrong if he thought that the sun was an absolutely firm "authoritative" fixture. As we now know, all free bodies within the solar system are attracted by all the others, so there is strictly speaking no totally dominant body — and yet it is a tolerably good approximation if we do take the sun as a fixture because it does play a locally dominant role. This flawed formulation may thus be harmless for many practical calculations; and yet if we took it literally, it would cause huge theoretical problems with such issues as Newtonian dynamics; though it is of course still a notable advance from the earlier geocentric ideas.
As a further twist though, we might note that a king-or-dictator is not an absolute fixture either, though the subjects ruled by Hitler, Stalin, or Henry VIII might have felt it prudent to believe otherwise. Ultimately even the apparently all-powerful ruler needs some supporting feedback from a sufficient number of his underlings — as Caligula, Caesar, and Charles I discovered, to their cost. So maybe Copernicus was more correct than he knew, using a king-analogy which actually depicts relative power and not the absolute power he probably envisaged!

8.2 *Coherence[77] as an indicator of likely truth*

If there is no ultimate authority for us to consult on what is true — nor any guaranteed procedures for finding truth, then we must obviously make the best of whatever *fallible* procedures may be available to us. Here we come upon the *Coherence criterion* for accepting-or-rejecting ideas.

Put simply, this gets us to test any new idea or hypothesis to see whether it fits in with whatever supposed-facts we have already accepted. If it offers a good fit, then we add it to our existing stock of ideas, thereby increasing the "coherence" or self-consistency of the whole ensemble. If it seems to be a serious misfit, then we simply reject it. But if it is a borderline case, we may have to use our skill and "commonsense" to decide; indeed we may do best not to decide yet, and instead await further insights or evidence.

Then again we may even change our mind afterwards. But that may be unlikely if we have already found our structure satisfyingly self-consistent, or if it has become so deeply embedded in our self image that there can be no going back within our own lifetime — or even the lifetime of our culture.

An obvious disadvantage is that this opens the way for bigotry — a locking on to some particular view (which may not seem right to *other* people who have formed different coherence patterns, or a view which may even be wrong despite general acceptance); and this locked-on view produces a reluctance to even consider new ideas or evidence. This is indeed an unfortunate side effect, but doesn't that sound all too true to life? The strategy is clearly not perfect, but then neither is human nature! Of course that does not, in itself, prove that we have the right model for human thought; but it does show that we may be wasting our time looking for perfect strategies if we want to understand the human mind *as it actually exists*. And as it now seems that no perfect strategy can exist anyhow, we might come to under-

[77] In the introductory text (*Book A*) I used the term "***internal* coherence**" as a safer synonym for the simple traditional "**coherence**" used here and in most of the literature. But however we name it, this is contrasted to experimental testing in the outside world — which I choose to call *"external coherence-testing"*, using the word "coherence" again(!) to emphasize the formal similarity between the two, whilst recognizing that there is still an important difference.
If the claimed similarity comes as a surprise, it is because we naturally assume, unjustifiably, that our direct perceptions are automatically reliable (while we are *awake and sane* — though that begs the question). That does tend to be true, *relatively speaking*; but in strict theoretical terms that is not good enough, especially when we get beyond the hard-sciences, physics and chemistry.
Thus we find that our direct perceptions are ultimately effective only because they too depend on coherence tests, either now or in the dim distant past; and without them we would not even recognize objects as such, though we might still pick up their uninterpretable images.

 — ISBN 0 9577737 1 4

stand why the mind has evolved such that it uses a less-than-perfect strategy.

However we can and do use further strategies for reducing the bigotry errors. *Temporary bigotry* (as we may call it) inevitably occurs all the time: • We can seldom wait for *all* the information relevant to any given decision; • we might have trouble in assessing relevance anyhow; and in any case • we might not be able to attend to it all in one go. So our decisions will usually be biased — *centred on some aspect* of the problem rather than the whole issue. In the longer term then, it would be useful to have some strategy to induce us to widen our focus — to become de-centred, or *"decentré"* (to put it in Piagetian terms), and hence make substantial corrections to our initial bias, though perhaps rather belatedly.

Note that we will also need a strategy for assessing the merit or de-merit of ensembles which are only *partially* coherent. That notion does not accord well with traditional mathematics; yet the idea of coherence is largely mathematical. Algebraic and geometrical systems are usually accepted as such only if they seem to be wholly self-consistent. (Just imagine the problems if ($9+2-2$) did not always equal 9, or if a solid object did *not* look the same after it had been rotated through 360°).

Yet if we are dealing with supposed-facts about a complex system such as a society or even just a simple cell, it would be most surprising if we ever found total and exact coherence amongst the "facts" — no matter in what faithful form we chose to express them. Instead, in practice we should probably be assessing the *degree* of coherence in some manifestly imperfect system of alleged-facts about the system (where these "facts" might, or might not, stem from respectable experimentation). This coherence-assessment might then encourage us to play round interchanging some of the supposed facts, or our "centring" of attention upon them, all in the hope of maximizing the degree of coherence. Or then again, maybe we would adopt the simpler procedure of merely comparing the coherence level of rival ensembles of these "facts" — and effectively these would be rival hypotheses, subject to a competitive selection process. (Thagard, 1992).

In fact such messy procedures would surely be the norm. Only in very special circumstances would coherence be genuinely perfect or near-perfect, though those special geometrically tidy cases are of particular importance when, as infants, we initially learn the "obvious" properties of solid objects.

Thus, on the one hand, there can *never* be any infallibly correct way to handle a given *social* situation; nor can there be any infallible way for a wild fieldmouse to escape all predators; and yet there *are* precise ways of coping with many *mathematical entities* and any simple physical systems which are closely analogous to them.

Another problem with the coherence-seeking technique is that it depends on an already-existing set of mentally-held ideas. This "incumbent membership" of ideas keeps serving as a sort of developing template by which to test the fit of any new ideas which are "candidates for membership to the club". But the process has to start somewhere, with some probably-inherited "foundation membership of ideas" (as we saw in chapter 5); and these initial codings may tend to bias all future developments. This tendency *might* well be substantially countered through learning experiences — but this is all a bit difficult to predict in any detail.

8.3 Coherence within a hierarchy

In section 7.1 we saw that there are particular reasons for expecting matter to be organized in hierarchical nestings from subatomic particles-or-waves, through atoms, then molecules, and so on up to tangible objects and beyond. Maybe that does not prove anything about the brain's organization, but it does at least suggest that we should ask whether the brain too might have an intrinsic hierarchical organization. So let us consider (i) what advantages such an organization might offer; (ii) what evidence there might be that such a hierarchy actually exists; and (iii) some minimal physical requirements for implementing such an organization.

(i) Advantages of mental hierarchy?

To start with, some philosophers and psychologists have raised the topic of *introspection* from time to time (and here I am mainly talking about *unconscious* introspection). This introspection is the apparent ability for the mind to monitor what is going on elsewhere within the mind itself — to *"observe" its own activity*, metaphorically speaking.

It seems likely that if *Part B* of the mind[78] is to "observe" *Part A* of the same mind, then there will be something like a master-servant relationship between them, at least temporarily — with the boss (*Part B*) watching over the performance of the underling (*Part A*). Now it is also possible that *A* could sometimes monitor its boss, *B*, in return — some degree of worker-or-voter participation in management, thus offering us one of those feedback loops which I suspect are highly important for ultimate stability. But on the whole we are likely to find that such relationships will normally be one-sided, and the resulting inequality of power is a key ingredient for hierarchy of some sort.

One chief advantage of being able to introspect within a hierarchy is this: The observer/master *B* can stand back "at arms length" and notice when the servant *A* seems to be stuck in an unproductive rut. *B* may then be in a position to alter *A's* defective tactics in the hope that future performance will improve. At its simplest, this could be just a random resetting of *A's* parameters; and such adjustments could keep occurring every so often until *A's* performance appeared to be maximal for the prevailing circumstances. In fact the efficacy of such a simple trial-and-error system was demonstrated decades ago by Ross Ashby (1952/1960), and

[78] These two entities *Part A* and *Part B* are not necessarily to be seen as having definite singular locations. They are *communicational* ensembles rather than anything tied down to specific sites — each being like a contemporary interest-group on the internet. Clearly any two such groups might well interpenetrate each other's space without this interfering with their (perhaps) totally different *communicational "proximities"*.

then investigated within a business-management context by Stafford Beer (1972/1981).

One of the principal ruts which *A* can fall into is a fixation with some small part of the reality it is trying to elucidate; and as we have seen this inappropriately continued "centring" leads to a sort of bigotry which can lead to bad decision-making. *B* can sometimes ensure that the focus of biased attention is broadened out, or at least moved around occasionally. That may very well provide us with a sufficiently "decentred" or unbiased system — a requirement we encountered in the previous section; (*i.e.* halfway through section 8.2).

So far I have only mentioned the entities *A* and *B*; but there seems to be no compelling reason why we could not also have a further *Part C* which can monitor *B's* performance. But then why stop there? This *C* could be monitored by *D*, and that could be governed by *E*; and so on! Of course this chain could hardly go on for ever, and the stages would have to operate on ever-slower time-scales; but there is no clear reason why the chain should stop abruptly at any particular level as long as the complexity of its task did not overrun it.

(ii) Evidence?

In fact it is this multi-layered control which gave Ashby's "tortoise" robots their surprising similarity to living systems. (Their chief defect, pointed out by Ashby himself, is that they are typical "nuts and bolts" models which have to be built deliberately by someone, whereas living beings have to depend much more heavily on self-organization for most aspects of their development, and not just the introspective part).

Perhaps that will suffice to show the potential value of this sort of hierarchy, but I will just briefly mention a few further points:

Primitive animals with simple behavioural repertoires may well have just *A*, or *A-with-B* organization within their nervous systems. Cats and chimpanzees would appear to have more elaborate control layers (as shown by their ability to *play*); while humans go further still, allowing them to use *abstract symbolic thought* — including fantasy and creative *lie-telling!* Moreover it is no accident that such successive developments of mental ability appear to mirror Piaget's list of developmental stages (Furth, 1969): Sensorimotor, Pre-operational, Concrete Operations, and Formal Operations — though personally I suspect that this list is incomplete, perhaps lacking a stage or two *before* sensorimotor; (see *Book A*).

The main point right now is that we could expect that any hierarchical organization within the mind/brain would offer substantial advantages wherever it could be made to work.

(iii) Some minimal physical requirements for the building of hierarchical mental systems

What then do we need to set up and maintain any type of hierarchy? For one thing there must be some way in which one level can be distinguished from other levels; that is to say, we need some *"badge of rank"* or some sort of *"members' enclosure"*, or both. In a mathematical sense then, we are defining various *sets* of entities; either "intensively" (by some recognizable property like a badge, label, or identifier, or descriptive list-of-attributes); or else defined "extensively" (by some sort of physical boundary or tethering). — In due course we will encounter examples of both types.

8.4 Endless debate on unclear issues — or hermeneutics

Traditionally *hermeneutics* meant the art of making sense of a diffuse body of linguistic information — notably scripture. One obvious approach is to seek *consensus through free debate* — and that is clearly a social task involving *language*[79] between the participating individuals.

Of course there is nothing foolproof about such a procedure: We might never reach a stable consensus; and even if we do, our solution cannot be guaranteed to be a true picture of the real world. But then, since it now seems clear that there are no foolproof paths to truth, we would do well to look seriously at all plausible-though-fallible approaches — including this hermeneutic strategy.

In fact hermeneutics is now also often applied within non-scriptural *scientific* contexts. Doubtless this is a response to the post-Kuhnian crisis in Scientific Method as discussed by Hesse[80] and others — though it would seem that (in the English-speaking world) this trend to embrace hermeneutics as a respectable principle is mainly confined to philosophers. Meanwhile, despite some lip-service, *scientists and their sponsors* seemingly remain largely insensitive to the failings of their supposedly-pure experimental approaches; and to that extent they see little need to look to other approaches, even if they happen to use them unconsciously every day.

---

[79] Here we have a trap for the unwary. *Language* is clearly crucial for the hermeneutic process, so it is easy to assume that languages are automatically just as important in all our personal thought processes — an assumption often unjustifiably made by linguistic philosophers, as Wittgenstein realized in his later writings.
I would provisionally suggest that hermeneutics (as understood in the present context) is an *entirely social* phenomenon, and that likewise languages like English and Japanese are also *primarily social*, playing no part in our *deepest and most primitive* personal thoughts.
Nevertheless we will later explore the Piagetian possibility that the individual mind could indulge in an analogous overall strategy. If so, then Piaget's *equilibration* might be seen as roughly corresponding to the hermeneutic process, and his *schemes* could be analogous to the words of our social languages. But such analogues would not be physically interchangeable. They would instead belong to different domains (like the two incompatible domains for "viruses": Bio-systems or Computer systems!) — though that need not stop them from intercommunicating through some suitable *interface*.

[80] Hesse (1980), notably her chapter 9: "Habermas' Consensus Theory of Truth" (pp.206-231). Also see sections 2.1(b) and 8.1 above.

Anyhow let us focus here on that task of *making sense* of diffuse information — the task of hermeneutic *interpretation*:

Given a fairly disorganized mass of scriptural texts — or scientific beliefs-and-data — how should one conduct the task of trying to *interpret* it all? Perhaps the first task should be to resist the temptation of becoming wedded irreversibly to the first pseudo-solution to present itself! — But having then established a modicum of tolerance, one still needs some way of encouraging the more promising ideas in the hope that they will lead to yet more plausible interpretations.

Freud saw his psychoanalytic task as being like trying to solve a jigsaw puzzle[81]; and I will take that as a useful benchmark example here, though we may have misgivings about it in Freud's own case. *Ideally*, any such detective task consists of assembling all the relevant facts, and then trying to fit them together until finally the single unique perfect solution is obvious for all to see. That of course is what happens when we solve a genuine jigsaw puzzle, though real-life conundrums are unfortunately not so tidy and accomodating in most cases.

One problem is to decide which items are "relevant". One must first decide "relevant for what?"; but even then the task is not trivial. There will be misleading jigsaw pieces which do not belong, and there will be gaps left due to missing pieces so that we may have to guess and do a little sketching of our own.

Another slightly different problem arises from distortion. Some of the pieces may be battered or discoloured. Perhaps all the pieces have fuzzy cloudlike edges instead of the clear sharp boundaries of genuine jigsaw pieces, and that may really try our tempers or lead us into error — and maybe mislead people like Freud too! *If* we manage to guess correctly at any clearcut substructure (as eventually happened in chemistry when it emerged from being alchemy, and again when it adopted quantum ideas), *then* we will have largely transcended the "fuzzy-edge" difficulty and we will be in a much better position to make rapid progress. But meanwhile we will have to flounder round as best we can with the unreliable and fuzzy concepts.

It seems to me that it is precisely this "floundering around" which constitutes *hermeneutics* — the art of making do with unreliable populations of vague concepts expressed in verbal form. And note that it is indeed an important art! This hermeneutic process may be slow and painful. In some cases *it may never get to really clearcut results* (as is probably the case in all the social sciences); yet it may still achieve useful incremental improvements, and do so often-enough to justify the effort. In other cases it may never actually get anywhere; though given our initial understanding at the time, we may still have been sensible to try it in the first place.

[81] See Fenichel (1946/1971, p.32).

So then, what is actually happening during this "floundering about"? It does seem to entail endless debate[82] — with many different people trying to improve the overall "form of the jigsaw picture" and each applying their own biases to the task of re-adjusting the pieces making up that picture. Pieces will be added, others removed; and some will be modified according to this-or-that biased view of how those pieces *should* appear. This is not perfection, but it will often be the best solution available — at least for the time being. Sometimes the real picture will be quite unreachable for all time, or its basis may be in a constant state of flux so that we can never hope to be quite up-to-date with our model anyhow. These limitations may vex us, but we may just have to live with them is many cases.

Occasionally we will be able to hit on an exact fit, or something like it; but that may well be the exception rather than the rule. Meanwhile we have a sort of statistical *"best-fit for the moment"* strategy, and we may be well-advised to accept it along with our reservations about it. Otherwise we are likely to be saddled with some dogmatic solution which will almost certainly be more seriously wrong factually, and possibly very painful for us in the long run!

A partnership with "coherence"

Hesse had mentioned coherence and hermeneutics as two possible strategies; but you may have noticed by now that this hermeneutics procedure is effectively the other side of the coherence-seeking strategy! — Coherence testing will not avail us much unless we are free to try different combinations of elements to see whether they might cohere better — and that surely is a hermeneutic process. Likewise if we are trying to optimise our "jigsaw picture", then our criterion for the "goodness" of our picture must surely be some measure of its coherence as a system of elements. (So it seems that the alternative to pure empiricism is "hermeneutics *and* internal-coherence optimization, *both together*", rather than suggesting them as two separate alternatives.[83]

Finally, whichever way you choose to view this coherence/hermeneutic strategy, notice the key role played by *trial-and-error*. This is a theme which will continue to haunt us at various parts of the following discussion, and of course it is no mere coincidence that it also occurs separately in the two rather different systems suggested by Darwin in 1858-1859 (for species), and by Ross Ashby in 1952 (for the brain).

[82] Hermeneutics will often entail public debate, though it seems to me that *one-person debates* would also qualify as long as they use language and competing candidate-ideas. One might argue as to which is the more efficient, but that is perhaps a separate issue.

[83] Of course some will not see coherence as variable in this way, but rather as an *all-or-nothing* affair — mathematically pure with no room for imperfection, and naturally that would not mix at all well with the inherently messy business of hermeneutics. It should however be clear by now that "coherence" as I use the term, must include imperfect cases with lesser degrees of coherence. Only then does it make sense to consider strategies for *optimizing* this graded variable.)

 — ISBN 0 9577737 1 4

9. MIND: THE "HARD" PROBLEM

9.1 The "hardware argument" favouring mind-molecules

As medical scientists find out more and more about the fine detail of cell-membranes and synaptic junctions, their accounts make more and more references to *chemical* phenomena.

In a sense this is inevitable because clearly membranes and suchlike are made up of chemical complexes, and many of their dealings have an obvious chemical component: allowing certain ions to pass through certain pores, transducing chemical energy, or releasing and responding to specific *transmitter* chemicals like adrenaline or serotonin.[84] So, given this inevitable chemical basis, we are left with some hope that the underlying dependable substructure might offer the "precision and structure" we were looking for.

Certainly some molecular structures are comparatively rigid and dependable in their configurations, and therefore capable of storing memory codes without undue corruption, in principle at least. Indeed even in the immediately bigger-than-molecule domain there can be an encouraging degree of regular order reminiscent of crystal structure. Thus note the way units are organized in serried ranks on a cell-membrane surface (Livingston et al., 1973), though this picture of regularity is somewhat sullied by the knowledge that there are frequent protein complexes poking through such membranes, and these seem to be placed in a rather haphazard way.

So now we have established that it might at least make sense to look for the missing "mind" domain within the realm of molecules; and that these molecules might well be substructures associated with the already-studied membrane-and-synapse structures, as reliable components within an unreliable floppy whole — perhaps interpenetrating it *on a very different scale of magnitude* so that in some ways we can think of them as two different spatial dimensions.

(It is also likely that the "reliable" molecules might be physically situated *external* to the floppy units themselves. However they probably would need to be in close communication — a question which we will return to later.)

So far, so good; but this does not give us much to go on. We are left with a vague notion that molecules themselves (rather than their more visible assemblages) may be the key action centres for the human mind. But *which* molecules? *How* could they store memory, and in such a way that it can be operated upon? And *just how* could such molecules communicate with the orthodox millisecond blip-coding which drives our muscles and carries reports from our sense organs? (See section 4.3).

These are pretty searching questions and we should not expect easy answers; but that is not a valid excuse against further investigation. Scanning the existing research reports might offer some further clues, like:

- Findings about the protein "memory-molecules" in our immune system, or
- Reports on the increased *RNA* concentration associated with learning[85], or
- Reports from brain-surgeons about the triggering of *well-ordered sequences* of memories.[86]

"Fine" you might say — "But what then?"

Ideally we might want to plunge into an experimental program to test these ideas; but could we really monitor reliably the memory-activity and state-changes in likely molecules? — That might just be possible, but it is likely to be very demanding and expensive research. So is that truly the most cost-effective way of pursuing the matter? And even if we did have a detailed account of each possibly-relevant molecule, we would still be left with the formidable task of working out what the system-as-a-whole might be doing and why that activity might be helpful to the organism concerned.

In other words: Here we are talking about *a bottom-up approach*, in which we start trying to identify the elementary units of the system, and then try to devise their significance within the wider scheme of things.

There are formidable problems at both these stages. Firstly the very demanding task of investigating at an ultra-micro scale of size; and secondly we are trying to build up a model of macro-reality on this basis, without really knowing where we are going or why, nor whether we have really hit upon the right basic elements.

However there is an alternative to endlessly perservering with this *bottom-up* approach based on hardware elements; an alternative which does not necessarily discard whatever insights we might have already reached. Obviously this alternative is the *top-down* technique, centring on *over-views*, on *"purposes"*, and on the *"software"* of the overall system:

[84] There is by now a vast literature on these issues. Bloom (1994), for instance, offers an anthology in this area. Katz (1966) is a fairly readable introduction to the subject; but of course being a relatively early book, it lacks any mention of some of the interesting later insights like details about the "gating" of ions through pores.

[85] There is the evidence that RNA-depositions correlate well with learning; and also various experiments using *injections* of RNA from other individual animals, and hence apparently transfering mental encodings of learnt information. — Hydén (1967a, 1967b), Glassman and Wilson (1969) — and further references cited by Piaget himself in 1967: Penfield (1958), Babich *et al.* (1965 Aug, Nov), Fjerdingstad *et al.* (1965), with a dissenting voice from Gross and Carey (1965).

[86] prompted by stimuli to parts of the brain during surgery (Penfield, 1958; Penfield and Roberts, 1959; cited briefly by Shaffer, 1967; and by Piaget 1967).

 — ISBN 0 9577737 1 4

9.2 *Psychology, and the "software argument" favouring mind-molecules*

Psychology takes a top-down approach to studying the mind. Like most other sciences it gathers an impressive array of descriptive material, and then usually tries to make sense of this material — with greater or lesser degrees of success. But what does it mean to "make sense" of data in this context?

Some psychologist like the behaviourists J.B.Watson and then B.F.Skinner have held that it *does not make sense to try to "make sense"* of our experimental findings, and that it is quite meaningless to speculate on what we cannot actually observe. In other words they are telling us to simply take our observational findings at face value, and go no further than compile statistical laws to give us a modicum of predictive power. But meanwhile we should not expect any insight into what is "really going on deep down" because, in their view, there is no "deep down" for us to even contemplate!

Fortunately there is now no need to take such a scientifically-pessimistic view, and indeed an alternative has long been implicit within Piaget's theories, as we shall see in *Book D*. Here, of course, we are considering *mind*-processes, but we have already briefly considered the equivalent change in our understanding of scientific method — the "mind-strategy" *of society*; (see occasional references in *sections 2.1 and 5.1, etc.*).

9.3 *The case for stringlike encoding, whether molecular or not*

If you were given the job of designing a code-system for representing your environment, what sort of "map-system" would you choose? At one extreme you might opt for a three-dimensional working model of reality, which would of course be splendid in many respects but might be unduly difficult to build, amend, maintain, and store. At the other extreme you might somehow record each isolated interaction you had with the environment, and (metaphorically speaking) note it on a scrap of paper which you then simply drop into a big unorganized box or "bottom drawer" (see section 4.4). Roughly speaking then, we have here both a sophisticated model in three-or-four dimensions, and a disorganized system in "zero-dimensions" with disconnected pointlike elements. The latter would of course be much easier to implement, but would it be any use in practice?

It might perhaps be fair to say that *robotics* adopts the former multidimensional approach by algebraic means, using such devices as *4×4* matrices. It is less clear as to which approach appeals most to orthodox brain-theorists, or whether they would even agree with this categorization at all. In this context they would probably differ greatly amongst themselves. Some would certainly see their models as three-dimensional (or more), though maybe involving some specialized technique like holography[87] or Fourier-transform principles. Others, perhaps including some synaptic-modification specialists, might either favour the pointlike "zero-dimensional" option, or simply not concern themselves with the issue at all. However it is my impression that very few have shown any conscious interest in two- or one-dimensional encodings.[88] That is perhaps their mistake, as we shall soon see.

In making our choice between the various coding possibilities, we have to reconcile two conflicting factors. We firstly want something which is flexible in its use and not too cumbersome to look after — and yet we also want something which is adequate for the job of capturing the essential connectedness which occurs in space-and-time within the real world, and that effectively means the *structure* of the real world. In short then, we are looking for a compromise; so let us start by eliminating the extremes:

Long-term *three-dimensional* models would be a logistical nightmare; and on the other hand, the isolated points of a *zero-dimensional* system would be a pretty hopeless way to encode structure! That seems to leave us with a short list of two possibilities: two-dimensional drawings or *"maps"*, and one-dimensional code-strings or *"sentences"*.

Precedents

Before we go any further, let us look for precedents in systems which we already understand. Firstly the *genetic code* for controlling our development is clearly one-dimensional — stored in stringlike strands of DNA.[89] Of course this coding does often "unpack" to give us three-dimensional protein molecules, but surely that is exactly what we would wish from a concise encoding: the ability to conjure up more complicated structures whenever these are called for,

---

[87] This holography idea was promoted by Pribram, Nuwer and Baron (1974 and earlier). It was also severely criticised by Willshaw, Longuet-Higgins and Buneman (1970) on the assumption that the underlying laser-like waves would be closely related to the traditional millisecond action-potential pulse — i.e. with a frequency of about *1 KHz* — and they claimed, justifiably, that such sources would lack the necessary stability. (As far as I can see, Pribram seems not to have acknowledged this criticism in his main publications (1971, 1991). Indeed it may have become somewhat irrelevant to him as he moved more into artificial intelligence).
It may not have been possible to take that argument any further at the time, but it is not too late to do so now: — I suggest that the lowish-frequency assumption need not hold; and if the underlying waves were supposed instead to be in (say) the infra-red region, then the Pribram model becomes plausible once more. (Traill, 1988).

[88] Of course much of our communicational environment is already organized in 1D, whether we are conscious of it or not — notably our language, our logic (linear reasoning), *and our conventional computer programs*. We likewise use 2D images without due credit.

[89] Strictly speaking, of course, the DNA strands are normally stored as *complementary pairs*, which makes them collectively into something of a two-or-three-dimensional structure. From a coding point of view though, this may still be seen as one-dimensional. Meanwhile the extra structure means (amongst other things) that the coding is largely protected from being corrupted — a feature which is dropped when the code is subsequently read off onto an RNA strand, which *really is* one-dimensional from our coding point of view.

 — ISBN 0 9577737 1 4

but to keep "out of sight" as a tidy storable stringlike code-strip at other times.

As a special case of this, let us consider the related coding for the manufacture of immunological antibody molecules. In the usual way, these codes are originally held on DNA, transferred to RNA and thence used to generate a nominally stringlike protein sequence. Protein however does not usually "lie down" in a neat straight line. Instead it has a great propensity to fold, crosslink, and generally knot itself up into three-dimensional shapes — but in a predictable way for each code-pattern. In this way the immune system is furnished with highly specific "lock and key" shapes which are vital to the way it functions. These are certainly three-dimensional, but they are just as certainly stored as one-dimensional codes.

As a second general example, let us consider how *society* encodes its public information. Here we certainly find abundant examples of both two- and one-dimensional encodings. Pictures, maps, and even medieval stained-glass windows are ever-present, but so too are those strings-of-words which we call speech (spoken or written). Thus we might say that both are of comparable importance, and perhaps we should leave it at that. However there are some further less-obvious considerations which cast an extra vote or two in favour of the the one-dimensional approach as being ultimately more important within society.

Pictures-as-communication tools have become very common on computer screens whereas a few decades ago they were a very rare form of computer output; and that certainly says something about their commercial and psychological value. However note first that these images or icons are frequently used as word-like symbols (somewhat like Chinese characters, or ancient Egyptian hieroglyphs). As such they no doubt help us to "unpack" the intended message, but often the main content of that message is essentially stringlike or conversational. Secondly note how these pictures are produced. Usually they are reconstructed from code — a linear stringlike code, which is either a list of pixel-properties, or else a list of instructions on how the "pen" should move when it is re-drawing the picture (and of course that is how we usually do freehand drawings). Certainly some of these linear activities may sometimes be arranged to occur in parallel, but perhaps we might just see that as a minor variation on the same theme.

1D versus 2D coding

So then, what is the verdict within these non-brain contexts? I would say perhaps two or three votes for string-like coding against one vote for the maplike version. Other things being equal then, we should be provisionally prepared to cater for both types of encoding in other similar contexts like the brain, but nevertheless expect the one-dimensional encoding to be the more important of the two. Piaget deals with both. If I understand him correctly, he reserves the term "schema" (plural *schemata*) for the embodiments used in the "figurative" two-dimensional case, whereas his term for normal linear encoding is "scheme". (Note that translators have often not been aware of this distinction) (Furth, 1969).

I thus commend the figurative aspects of mental processing as a legitimate and important area of study. I note that visual and tactile fields have been shown to map in an orderly way onto local fields within the brain; and I appreciate the need for special recognition networks which are presumably best handled in two dimensions, in part at least. However in this current project I choose to concentrate instead on the one-dimensional approach to coding in the belief that it is ultimately the more important — and the more amenable to simple explanation.

9.4 "Darwinian" trial-and-error processes — in a digital context

The severe demands for any knowledge-acquisition system

Let us look again at the "heretical" idea introduced earlier (in an analog context) in *sections 4.2 and 5.3*, and remind ourselves of the issues. As was suggested there, perhaps the biggest technical mystery about the brain is this:

"How do we record meaningful impressions into memory — and without losing track of them once they are stored?"

Given any *already existing* code and where to find it, it is not too difficult to speculate sensibly about how it might be *read off* to produce systematic behaviour. That is the comparatively easy bit, and is well understood in the different context of *protein*-production (rather than the *behaviour*-production considered here) — Miller (1970).

But next there is the moderately difficult question of how any non-trivial impressions could be *laid down as "recordings"* in the first place, onto what medium, and according to what formula?

This is no simple question once one seeks to go beyond such vague and unspecific suggestions as:

"supportive feedback of success reinforces the latest synaptic re-arrangement".

Not that there is necessarily anything wrong with that claim as far as it goes, but it does not actually go very far when trying to explain human capabilities.

As in section 5.3, we might try thinking through *the actual logistical requirements*:

(1) Perhaps look carefully at the design-specifications for a commercial tape-recorder, or disk-cutter, or computer memory — and see if any of these designs might have any feasible applicability to "recording" within a biological context, or whether they inspire us to think of some other approach which might be more suitable. That is no light task, but we might come up with some good ideas. Yet even if we do have success here, that is still not the end of the road. Thus:

(2) Next make sure that we can explain how our chosen system can *self-assemble* within the individual, with no direct guidance or help beyond normal trial-and-error processes (via species evol-

 — ISBN 0 9577737 1 4

ution or otherwise). Our account would not lose credibility if we were to explain this development as occurring *in collaboration* with other "sister" structures also being unpacked from their DNA-coded blueprints; indeed such co-development and interdependence would be the rule rather than the exception. But our explanation would be seen as dodging the issue if it depended on deliberate guidance (as within a tape-recorder factory), since this would not explain the seemingly-divine purposefulness underlying this guidance or design.

(3) Note that recording messages and reading messages are always likely to entail quite different technical procedures. The *reading* process is not just *"anti-recording"* as if all we had to do was put a minus-sign in front of something! That means firstly that we also have to explain a separate subsystem to do the reading of the messages when they are required; *plus* how they are to be found and activated at the right times; — *plus* how they are to self-construct along lines similar to the recording mechanisms. — And that is still not all!

(4) Because reading and recording are separate processes, there is no clear guarantee that they will be using the same "language" or coding conventions — nor indeed is it clear how the codings of input and output messages would relate to each other anyhow. Such matters will therefore need to be cleared up before we can reach a convincing overall explanation of advanced bio-memory.

One is likely to throw up one's hand in horror when first confronted with such a demanding list of pre-specifications; but surprisingly such constraints are actually *very helpful* — or at least they *can* be, if we handle them correctly. Remember that we are not just trying to pander to some visionary's impossible dream, for *we do already know that brains can-and-do operate* according to the specifications. So the constraint-list is effectively a set of valuable *clues*[90] upon which we can build our detective work; and moreover these clues are more likely to be quite searching and very-much to the point — possibly much more so than any traditional experiments we might be able to attempt.

The Darwinian precedent

But then there is that extra "Darwinian-model" clue from outside the topic[91] — a clue which almost gives us an answer ready made, as if we had cheated and looked at the "answers page" at the back of the book! Let us briefly look at the likely parallel between the two learning situations (genetics and brain) once we consider Piagetian *schemes* rather than neurosynaptic connections as the elemental units of information-store:

Traits coded on → *DNA*	Selection between → *phenotypes*	best fitted, best chance
Action-codes on → *"Schemes"*	Selection between → *action-patterns*	best fitted, best chance

One vast difference is in the time-scales, though the same basic principles could still apply regardless. Thus inherited DNA-codings can only be varied at each new generation at the earliest, whereas "action-schemes" mutation and selection (and the concepts presumably built up from them) could feasibly mutate, and be selected within minutes. Indeed perhaps the most crucial aspects[92] could even occur within fractions of a second as we evolve our concepts during a witty conversation or a tennis match.

In other respects, the two models look very similar in their formal approach. Moreover there is also ample scope for some types of interaction between the two. In particular, some of the genome code (stored on the DNA) could be generating standard "benchmark" scheme codes as a basis for one's behavioural repertoire. No doubt many of these could then be modified, mutated, or overridden, but they could offer a stable starting point for each individual — though as with other personal traits, genetics can sometimes endow us poorly.

Could this apply to the brain?

It is tempting then to guess that at least some aspects of our brain-based memory will operate in the same sort of way since the formal problems are very similar. If we were to adopt this as a working hypothesis, our next duty would seem to be to see how it would measure up against our above list of requirements. Thus:

---

[90] This point could be developed formally in terms of *the law of requisite variety*, as discussed by Ashby (1956/1964). However an informal mention will suffice here.

[91] It is important here to remember the original genetic debate was between Lamarckian and Darwinian views (see sections 5.3 and 4.2 — and *Book A*). Lamarck thought that characteristics acquired by an individual were *somehow* recorded internally and hence passed down to the next generation, thus accounting for evolutionary changes over the generations. Of course that would leave us with huge problems in explaining, even in principle, just *how* this transcription could take place — *and then* be effective in the development of the offspring.
In contrast, Darwin saw the changes in the developmental blueprint as occurring spontaneously and randomly — a blind and rudderless system of variation which only achieved direction through the poorer life-expectancy of the relative "failures" and the better reproduction-rate of the relative "successes". — Note that the successful codes are *not* seen as being copied down from direct lessons as Lamarck would have it, nor are they deliberately pre-set by some designer; instead they are just simply produced like a dice-throw, and then subject to trial-and-error elimination processes.

[92] This implies a bio-strategy similar to "Critical Path Analysis" of the commercial world. The overall process can be greatly facilitated by ensuring that any slow sub-processes are performed *in advance* or *in parallel* wherever possible, so that they do not hold up the key activity at the crucial moment. Thus for example, in this case any code-replication or perhaps mutation could take place well in advance. (The immune system might also offer some useful analogies here).

(1') How is the brain to record elements of memory? — The simple surprising answer would then be that it does not actually record them at all! — Instead it must produce quite large numbers of elements which might conceivably apply to the situation, and then discard most of them as unsuitable. (That would fit in well with the molecular-coding hypothesis, especially if we consider the largish numbers and the logistics of wasted failures; but such ideas would be hard to envisage within a system of adaptable synaptic connections alone since they would probably not be prolific enough even if they did carry sufficient precision).

(2') Self-assembly? Almost no difficulty regarding the actual elements. It is not likely that they would be DNA themselves, since DNA is normally too rigid and stable; but their initial coding could certainly come from DNA codes directly if they are RNA, or indirectly in the less likely offchance that they are proteins. Then again, they might be PNA[93], a more primitive linear molecule-type.

This leaves us to explain how the related "machinery" could self-construct; but that is a problem which is surely common to all accounts of metabolism, so it has no obviously special significance for our current hypothesis.

(3') Reading this coding? What does the genetic-code case tell us by way of analogy? In that case there is a stage when a molecular-contraption (a ribosome) travels along the RNA strand, like a small portable tape-player travelling along a tape hung between two trees. Meanwhile this ribosome generates a strip of protein with the help of the RNA code-strip feeding through, whilst drawing upon the prevailing supply of suitable raw materials. (Miller, 1973).

In our present brain-memory case, we could envisage something similar if the coding does happen to be on RNA; but there is some reason to believe that it could actually be a little simpler in operation than for a ribosome. Note that we would now be dealing with signal-generation rather than protein-building, and signals do not need much in the way of raw-materials, nor any special procedures for dealing with cumbersome solid products such as protein-structures. It is therefore possible to imagine a similar read-off process, but with something rather less formal than a ribosome to act as the zip-hold: perhaps some sort of energy-disturbance[94] travelling down the RNA string, or maybe just a lightweight pseudo-ribosome of some sort. In either case, there would probably not be any generation of new molecules, but rather there might be a patterned emission of photons, each of quite definite light-frequency and generated from specific quantum jumps at predictable sites — perhaps also having predictable phase-relationships with other photons so that specific optical-interference patterns might arise.

Of course this all sounds uncomfortably speculative, especially to those of us who have been endlessly drilled in the need for clearcut experimental evidence. That evidence might come later, though first it remains to be seen what funds might be found for such investigations! But meanwhile you might like to weigh such ideas in terms of coherence criteria, perhaps making use of Thagard's (1992) computer algorithm for such assessments. Moreover we will see below[95] that this hypothesis has some unexpected dividends which also contribute to the overall coherence of the approach.

(4') Coordinating the languages of reading and recording mechanisms. This problem simply disappears because there is not supposed to be any direct recording at all. The laying down of memory is thus actually a selection process, depending on reading rather than writing; and of course memory-retrieval will presumably also use the same reading process — so there seem to be no loose ends left to reconcile. Moreover both processes could presumably use the same addressing-or-locating technique, whatever that might be.

9.5 *Signal-interface considerations — and molecular implications*

Descartes (1649/1989) saw the mind as operating in an *other-worldly domain*, divorced from the mundane material phenomena of the everyday physical world. Now apart from anything else, this then left him with the vexing question of how this *otherworldly* mind could find any way at all to interact with its profane *worldly* body. In other words he had an interface problem — trying to explain how signals in one domain can cross some troublesome boundary so that they can then continue their journey within a different domain.[96] (For what it's worth, his suggestion was that this link-up occurred within the pineal gland, situated within the brain).

Coming back now to the present, what are the actual or possible domains which we should consider, and what can we say about mechanisms for transferring signals between them?[96] To start with at least, let us try to find an exhaustive list of possible mind-related domains:

[93] See section 5.4, *and/or* Böhler, Nielsen, and Orgel (1995)

[94] a *phonon* maybe — that is to say, a photon trapped within a solid medium; see Cope (1973).

[95] and in *Book C*, (and also the later chapters of *Book A*).

[96] See *Book A* (Traill, 1999) for further discussion of these issues.

 — ISBN 0 9577737 1 4

(a) The outside environment. — That, after all, is what our minds are trying to model.

(b) The traditional synaptic nerve-signal system. — This is clearly the carrier for all-or-most sensory-input, *and* the motor-output sent to muscles or glands. It is also involved in figural processing such as pattern-recognition.

(c) The stringlike molecules as discussed in this book — supposedly having arrays of independent codable sites *within* each of these molecules.

(d) The supernatural domain assumed by Descartes and others. Note that this view still has a wide following within the community, even though biologists now see little merit in it. It is important for our theoretical discussion at least.

(e) Other. — This of course is just a "catch-all" category to cover any domains I may have overlooked, or dismissed as neither likely nor theoretically interesting. (For example, here we might include any systems supposedly using microwave frequencies, perhaps via whole-molecule switching of some sort, and maybe thought to be vulnerable to the new "hi-tech" radiations?) I have no plans to explore such possibilities myself, though others might well wish to do so.

It is instructive to draw up a *5×5* table (or bigger if you wish to include other possible domains), depicting how each of these five domains might transmit signals into each of the other four, *and also* within itself ((b)-to-(b), (c)-to-(c), etc.). You can then go through all or some of the *25* possibilities, and think about what mechanisms might possibly be involved at each of these postulated interfaces. My own remarks here will be mostly confined to the cases I see as most relevant or feasible, but there is a case for looking carefully at all combinations and their likely implications.

To start with I shall eliminate all (d) and (e) cases as I have nothing constructive to say about them here[96] apart from my above references to Descartes and to popular belief. This leaves us with table 9: a *3×3* table with nine possible entries, and I am not equally interested in all of those. The *muscle*, *senses*, and *synapse* entries obviously all relate to the orthodox synaptic system and its interrelationship with the environment (as *output*, *input*, and *internal communication* respectively). There is nothing new in that, so let us move on.

The entries for *"read"* and *selection* do refer to new concepts, but we have already just discussed them in section 9.4 above. As for the "—" entry; this item is irrelevant here because it does not involve the mind at all.

Table 9

MIND/BRAIN INTERFACE CHART

	from Envir. (a)	from Axon (b)	from RNA? (c)
to Environ. (a)	—	muscle	×
to Axon (b)	senses	synapse	"read"
to RNA? (c)	×	selection[97]	**quanta**
where "×" = *apparently-forbidden interface*			

That leaves us with two untreated talking-points: (i) The ***"quanta"*** entry which raises the issue of how the molecular-based signals could travel about within their own domain. This turns out to be a far-reaching question, and its consequences will actually occupy us for most of the rest of this book. (ii) The two "×"-entries which suggest a *forbidden* communication channel — a two-way interface type (between *a* and *c*) which may be inadmissible. Hence any *a↔c* signal-traffic would have to take a roundabout route (via the traditional nerve-mode, *b*) if the direct *a↔c* paths really are unworkable. Let us look at these two issues in turn:

(i) **Quantum patterns as the "language" for communicating within the molecular coding domain**.

In the previous section — 9.4 (3') — there was discussion about phonons possibly travelling along RNA-like molecules. But more importantly there was the further notion that the code-string "sentences" read-off from such a molecule would be likely to appear physically as a pattern of quantum jumps. These quantum jumps would doubtless have energy levels within the biophysical range, and that would probably mean photons with frequencies in the near infra-red. These might either travel more-or-less freely through isotropic media, or as *"phonons"* if they were held captive within some sort of crystal lattice (and that might include linear molecules like RNA or PNA).

(ii) **A forbidden interface — or ESP instead?**[98]

As depicted in table 9, and just discussed, the molecular-memory domain can only contact the outside world through an intermediary, and never directly. The suggestion is that it is *wholly dependent on the traditional synaptic system* to serve as this channel; hence the void "×" entries at (c)-to-(a) and (a)-to-(c).

---

[97] not strictly "write" in any conventional sense, as explained in sections 4.2 and 9.4, above.

[98] ESP = Extra-Sensory Perception, as discussed within psychology, though there is substantial doubt as to whether it actually exists.

Despite this, it might still be possible in principle for the molecular system to bypass the orthodox nerve mechanism, and in this case we would have something very like those "Extra-Sensory Perception" phenomena (ESP) which some psychologists and others espouse.

For the sake of argument, let us provisionally assume: firstly that ESP-like phenomena do genuinely exist in some form (even if they can only be *seen experimentally* as faint statistical effects); and secondly that the aspect of mind which they relate to is the RNA-like molecular system (and not some supernatural aspect of mind, nor the synaptic system). The best task for ESP-apologists would then seem to be to explain the physical mechanisms[99] whereby the molecular-system might communicate directly with the environment, without any intermediary aid from the orthodox synaptic "sensory" domain (hence arguably justifying the term "extra-sensory").

Some of the signals could plausibly make the necessary direct journey by fair means or foul, but it would actually be an uphill battle — mainly because the kind of signals probably involved (*infra-red* as we have just seen) would travel very poorly indeed through wet media like human tissue. In fact a half-life distance of *20µm* is about all we can expect through water, (see the chapters in *Part II*, below).

With less watery animals like *insects,* this particular situation may well be very different — an issue which has been hotly debated in some quarters.[100] But that is another story; and of course insects would hardly use such a link for the same purposes.

On these theoretical grounds then, my feeling is that ESP and/or its "motor" equivalents just might occur occasionally, but any such effects would usually be highly unreliable and uncertain, and it is very questionable whether the informational bit-content (or "baud rate") could ever be anything but trivial — perhaps enough to convey "present-or-absent" messages, but not much more. And perhaps indeed that is consistent with the never-wholly-convincing experimental findings about ESP to date!

In any case such alleged phenomena seem to be, at most, no more than a minor anomaly for our present purposes, so I shall say no more about them here.

9.6 Initial thoughts about mind and consciousness

Consciousness is the key mystery of the mind — that intractable unknown at the centre of the "*hard* part" of brain theory which concerned us at the start of chapter 1. So is there any conceivable chance of explaining this consciousness in any sort of material terms?

[99] Some might prefer to invoke the supernatural. However this present discussion offers a more natural physical explanation, *if indeed* there is really any effect here to be explained!

[100] This refers to the later writings of P.S.Callahan and his critics.

Let us first bear in mind the possibility that *freewill* can coexist with the more mechanistic aspects of reality — as discussed in chapter 6. Then let us provisionally assume, on this basis, that the task of explaining consciousness is scientifically feasible, at least in principle[101]. But even if this explanation is feasible, it has surely not been forthcoming as yet; so what are the prospects for fresh developments within this present project? I have no miracle answers either, but perhaps I can offer some helpful comments:

As with other related issues, I feel it is best to divide such problems into "software" and "hardware" issues. In this case I suspect that the most informative explanations will evolve within the *software* domain, and I will therefore defer any deep discussion of the matter until *Book 3*. Of course the new software thoughts will be somewhat dependent on revised thoughts on *hardware*, and these are being generally discussed in the present volume. These new general *structural* ideas are mainly aimed at more mundane problems like how the brain might come to hold the encoding of an object, a *set* of objects, or a hierarchy; however it is at least possible that a rigorous treatment of these lesser problems may pave the way for the more vexed question of consciousness. In other words, we may provisionally hope that the tools developed here in *Book B* will serve just as well for explaining some new aspects of consciousness as for other brain issues. If further hardware concepts then need to be developed, then there will be time enough to do that later.

Meanwhile let us raise some exploratory thoughts about the software "specifications" for *consciousness*. To start with, we may note that some philosophers treat consciousness as meaning *the ability of some parts of the mind to 'spy upon' other parts of the mind*. As I see it, that is merely a start. In effect it is giving the key requirement for any ability to manage hierarchical organization within the mind. Surely this is important, and crucial to developing concepts of objects and their sets (as will be evident in Piaget's writings, or in *Books 0 and 3*). But it barely scratches the surface of what is needed, and it certainly does not capture the spirit of that subjective consciousness which makes us aware human beings.

To this we surely need to add an extensive list of further requirements; and I shall start with a few suggestions here.

[101] In this present context of human explication, my use of the words "in principle" usually implies something like *"that would be possible if only we had enough time or money (or some other resource) to complete the job"*. This also carries the expectation that at least typical *parts* of such a project could be explored in full; (though it might be a moot point as to whether such partial analyses would be of much help in any particular case).

I prefer to avoid using the *in principle* term for any hypothetical tasks which I believe to be fundamentally unrealizable, like any detailed prediction of our own future (even where it may be strictly-speaking predetermined, see chapter 6).

(In the somewhat different context of evaluating a supposedly comprehensive set of postulates, the expression "it seems adequate *in principle*" implies that the ideas are coherent given the information currently to hand.)

 — ISBN 0 9577737 1 4

For one thing we probably have to acquire an awareness (whatever that means) of time and temporal-sequence; space and spatial sequence; causality and causal sequences; and other such basics including some elementary concepts of probability.[102]

These concepts too are fairly fundamental, and are probably required even by some animals which do not share consciousness in the human sense. In other words, such conceptual abilities are presumably *necessary but not sufficient conditions* for consciousness in the subjective-human sense — and no doubt we could say the same about such things as *emotional states*.

Perhaps a more telling requirement might be *self awareness* — the concept of self as being one sort of object amongst other objects, and yet an awareness that this 'object' is extraordinarily responsive to one's own thought processes in a two-way causal interchange. Onto this we might add concepts of wish, wish-fulfilment, and concepts of society.

No doubt we could find many more crucial criteria. So what should we be looking for if we are trying to add to this list of 'specifications'? For one thing, we might well contemplate our own conscious state, and ask ourselves what is so special about it? — Why should my friends also be just as confident that they too are conscious? — And what would be needed to *really* convince us that the structure within an artificial system was also capable of subjective consciousness?

With these unresolved questions, I shall shelve this vexed issue for the present.

[102] — all concepts whose development within the infant were studied by Piaget in defiance of the traditional Kantian view that such mind-constructs came fully inborn.

 — ISBN 0 9577737 1 4

10. IMPLICATIONS OF THE NEW APPROACH

10.1 *Molecular communication via infra-red quanta? — a summary*

It is all very well to postulate a storage mechanism for memory; but for such proposals to be credible, we also need to have plausible suggestions for all the important logistical details of how they could operate. At this postulate stage, it is too early to expect hard evidence (especially if that is likely to be very expensive), but we should at least expect suggestions about the substructure and its transactions to make sense in terms of physics and information theory.

Ideas about molecular storage of memory are not new, as we have already seen (in footnotes 85 and 86) in connection with the work of Hydén and others. However these all seem to have been purely correlational findings which did not dare to suggest any conceivable mechanisms — and indeed in those pre-Kuhnian days of the 1960s, such postulates would not have been welcome.

But we no longer live in the 1960s, and it is high time we asked for some in-principle details for the workings of all proposed methods of memory storage. In this case, I have yet to discover any way in which the traditional action potentials (the millisecond voltage blips sent along nerve-axon membranes) could *efficiently* communicate with the supposed molecular storage; (see section 4.3). Initially that seems perplexing, since it is hard to believe in the rival *"reverberating circuits"* as plausible mechanisms for long-term memory (LTM); and the correlational evidence for at least some molecular storage did seem fairly strong.

The trouble was that any precision in the communication with storage molecules (with dimensions measured in *nanometres*, that is $10^{-9}m$) would hardly match well with millisecond action-potentials whose pulse-lengths would be measured more in *metres!* Nor does the situation look any better if one re-casts the problem in quantum theory terms; (briefly discussed in *Book A*).

If molecules as such *are* to be repositories of memory code, then surely there must be some more efficient and precise mode of communication for them — at least when they are intercommunicating amongst themselves. Conventionally, molecules are mostly thought of as communicating only through close contact, either directly or through intermediate molecular messengers, but photo-chemical effects have been well known for more than a century, so why not consider that here? Indeed any code-switching would be likely to be associated with the emission or absorption of a photon quantum of energy, as suggested in section 9.4 above — or perhaps an orderly *pattern* of such quanta (together with the opportunity for optical-interference patterns which this could entail). On the face of it then, this looked much more promising; and if such transactions were happening, then the size of these biological quantum jumps would be expected to involve frequencies in the *infra-red* range — Szent-Györgyi (1968), and as discussed in the following chapters.

We have already encountered some of these issues, from a different perspective, in chapter 9, and especially section 9.5 and its table of interrelations. As pointed out there, it seems there are thus at least *two* modes of communication: the traditional action potentials with their patterns of millisecond spikes, and now the postulated quantum-exchange patterns as well. Each would have its own preferred domain, but there must also be provision for interfacing between them whenever this becomes essential — just as the miniature logic circuits of a computer chip in a robot must ultimately translate its coding into those cruder power circuits which drive its gross mechanical actions.

However the main point here is that there is now a plausible case for expecting infra-red signals to be part of the distance-communication system of the body, so we need to be on the lookout for:

- possible empirical *evidence* for it;
- possible further *applications* for it if it is actually there; and —
- *likely theoretical implications* and whether these helped or hindered the overall coherence of the idea.

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

PART II

POSSIBLE INFRA-RED NEURO-SIGNALLING AND ITS CONSEQUENCES

11. FIBRE OPTICS — IN NERVES?

11.1 An outline of the idea of infra-red neurotransmission

In *Part I* we saw how our memory and higher thought could well be carried on at the *molecular* level, and not just at the synaptic-junction level. We also saw that any such molecular encoding would be likely to "speak" to its neighbours by using some pattern of infra-red quantum emissions — sometimes contributing to those action-potential voltage spikes involved in synapses — but also sometimes free to take part in the much more individual and specific activities which higher thought would seem to require.

Likewise, such molecules would need to be able to "listen for" highly specific messages — either by recognizing their own digital "callsign" codes within those messages, or specific because the messages have already been steered their way by a similar recognition process elsewhere.[103]

In other words some molecules, somewhere, must be able to decode and act upon any selective "telephone numbers" included within the messages; and if they are to have sufficient precision, such patterns would need to be made up from infra-red emission patterns. —

Or so it seems. True there will be many *other* bio-signals present, well known systems as described in the text-books: • The slowest and crudest (though still important) would be those hormones and other endocrine secretions which drift unselectively throughout the whole body as general chemical messengers — like unaddressed public notices littering our newspapers, TV-screens, and letterboxes. • Next there would be similar chemicals delivered to specific targets, in the form of chemical "neurotransmitters" released onto particular synaptic membranes — a faster and more selective version of the above, but now analogous to advertising mail directed at some narrow market-sector such as *dog-owners*.

These synaptic neurotransmitters are activated in turn by • a third signal mode: those millisecond *action-potential spikes* which travel as a renewable chain reaction along nerve membranes, or leap capacitatively along segments of myelin sheath. These are probably ideally suited to muscle-control and for relaying much of our sensory input; that is, for interacting with big-and-clumsy interfaces linked to the outside world. However I really cannot see how such millisecond pulses could possibly communicate *efficiently* with any well-developed system of molecular memory. There would simply be too great a mismatch between the wavelengths, or pulse-widths, or "baud-rates" appropriate for the two different systems.

Of course there would surely be *some* intercommunication between the millisecond pulses and the molecules, especially on early input and final output, even if the effect is only slight or poor. • But any intervening *higher thought processes* which directly involve molecular codings must surely use a signalling system directly tailored to that internal role — and I strongly suggest that local *infra-red* is the most promising candidate for that task, at least on the basis of this preliminary enquiry.

Any new proposal like this is likely to evoke criticisms which seem damning. These criticisms may indeed *be* damning or indeed fatal, but then again they may not. Initially it is difficult to tell. For instance, we might try to put ourselves in the shoes of those who objected to that "outrageous" suggestion that the Earth was rotating on an axis: "Wouldn't we all spin off? — Wouldn't the vibration spill the sea everywhere?" — and so on.

There are similar critical questions about our present topic of infra-red communication, and we shall look at these in the next section.[104] Such misgivings *might* eventually turn out to be justified, but then again they might not: — They might later be seen as *misunderstandings* (as in the "sea-spillage" case), or as *overemphasized quantitative* effects, like the supposed centrifugal danger from the Earth's rotation.

[103] The distinction here is between *"intensively" defined* signals (identified by their own characteristics), and *"extensively" defined* signals (identified by location or connection). — See *section 12.3* for a fuller account.

[104] For instance, "Water absorbs infra-red very strongly, so how could the signals survive long enough to be useful?"

 — ISBN 0 9577737 1 4

But meanwhile there are also other further arguments in favour of IR explanations. In fact what set this idea going, apart from the signal-mismatch problem discussed above, was a certain scepticism about the textbook explanation of *saltatory conduction* — the way nerve signals jump more rapidly along those nerve-fibres which have an insulating coating of the fatty tissue — the *myelin*.

Never mind the details at this stage, though we will come back to it later. Suffice it to say that, from a physics point-of-view, there seemed to have been a theoretical omission in the standard explanation:

Geometrically the myelinated nerve-fibre looked very like a coaxial cable (as often used in TV aerials, with a cross-section like this ⊙); and one important variable in calculating the transmission properties of such a cable is its *inductance* "*L*" which can be calculated quite well from the cable's geometry. Yet the standard medico-biological accounts have totally ignored this inductance, though of course they do consider resistivities "*R*", and capacitance "*C*".

Now this simplifying assumption (that *L* does not matter) still gives us good practical forecasts in many circumstances, *though not in others.*[105] It has always accounted well for those familiar voltage-spikes described in the textbooks; but what if the pulses or wavelengths were to become so very short that *L* could no longer be ignored? (This would take us into optical regions — say *1μm* in length, or with a duration measured in *picoseconds*[106] or less — about a million times briefer than the traditional millisecond spike.) If we continued to ignore the existence of *L*, then we would never notice that there is another theoretical possibility.

However, if we correct our equations to allow for *L* in all circumstances (whether its effect is significant or not), then we will discover that there are *two types of solution* to the signal-transmission problem. There still remains the milli-second spike solution of the textbooks — but we now also have a second type of solution[107] arising from the same equations, allowing for a different mode of signal transmission down the cable-or-fibre. And what range of frequencies seems best fitted to this new mode? — None other than the *infra-red* band of frequencies — that very band which seemed most promising for molecular com-munication. (Traill, 1988).

Of course, as evidence goes, this concordance of predict-ions is just a start. More evidence and detail will come later, followed by attempts to gauge the overall coherence of the whole ensemble of ideas. But meanwhile I suggest that we have enough of a *prima facie* case for the topic to be approached methodically; so now let us take stock of the more obvious *criticisms* which need to be heeded:

11.2 Arguments against infra-red neurotransmission

The main criticisms of this infra-red hypothesis seem to be these:
(i) Water is severely opaque to infra-red, so wouldn't that kill the signals promptly?
(ii) Noise contamination. At body temperature *in any material environment* there will always be a significant amount of ambient infra-red radiation, as "black-body rad-iation" — an apparently-random background *noise* which would surely swamp and corrupt any would-be messages coded as infra-red signals?
(iii) Doubts about the optical suitability of the myelin-dielectric and its boundary surfaces. After all, in the com-mercial world, a dented coaxial cable is likely to produce signal distortion. So what about geometrical imperfections in the myelin?

These imply *technical* objections which cannot simply be dismissed at will, and the hypothesis must remain suspect until we can find at least some sort of *plausible answer* to these difficulties. Of course mere plausibility does not provide a guarantee, but it does enable us to proceed in a credible way to test the ensemble of ideas — whether by experiment or by other coherence-criteria.[108]

However there are also some other criticisms which stem more from interpersonal issues, and here the solutions will probably lie more in rhetoric and polemics:

[105] This is discussed more fully in *Book A*, along with the historical background to this topic. (There is also a short account in footnote 109 , on page 57.)

[106] 1 *second* = 1000 *milliseconds* = 1,000,000 *microseconds* (i.e. 10^6 *μs*) = 10^9 *nanoseconds* = 10^{12} *picoseconds* = 10^{15} *femtoseconds*. — And similarly for *metres — mm — μm — nm — pm — fm.* In fact as we shall see, both picosecond and femtosecond time-scales are now of considerable importance in chemistry — as *"femtochemistry"* (Wiersma, 1994; Chergui, 1996). Note that in one femtosecond, light travels only *0.3 μm* (= 300 nm = 3000 Åu) — about half the wavelength of visible light, and comparable to the size of synaptic mechanisms.

[107] Of course there is nothing new in this idea of multiple solutions to a problem. Nature often offers alternative at all levels: — Right-or-Left hands — Lungs-or-Gills — This-or-that energy source — etc. And the computer which I am now using allows me to choose keyboard-or-mouse for many activities. More relevant here though is the well known fact that $x^2 = 9$ has two solutions. The *obvious* answer is "$x = 3$" , but then "$x = -3$" is equally valid *even though we sometimes overlook it*.

[108] Remember that, in this postmodern deconstructionist world, there is no legitimate hope of any absolute proof about anything — nor of any completely rigorous *dis*proof either (despite Popper!). (Moreover many of our supposedly fundamental ideas must now be regarded as suspect, and should be *deconstructed*, at least provisionally.) As we saw in *Part I*, the best we can do is find *the most coherent* account available at the time — often in competition with other more-or-less coherent accounts. (That would seem to be a *post*-postmodern approach — perhaps better described as *re*constructionist). Anyhow direct experimental results will often enter into this overall reconstructionist coherence assessment — but their unquestioned dominance should no longer be allowed. In particular, a mere absence of "adequate" experimental evidence need not necessarily block all further research on a topic.

 — ISBN 0 9577737 1 4

Table 11
Local maximum and minimum absorption rates in water, for *infra-red* light.
— after Zolotarev *et al.* (1969).

Wavelength in vacuo *(μm)*		Distance infra-red travels in water before attenuating to $^1/_e$ (36.8%)	
		local min. (μm)	*local max. (μm)*
"Near" IR*	1.6		1819
	2.93	0.75	
	3.8		88.94
	4.72	23.77	
	5.3		43.04
	6.1	3.79	
	7.7		18.57
	15.0	2.75	
	38.0		8.35
"Far" IR*	48.0	7.83	

** The terminology varies somewhat for this longer-wavelength end of the IR scale. Hecht (1990, p.69) is typical in taking the IR band to extend up to 1mm (1000μm), the start of the microwave band; and rather implies that "far IR" means λ>50μm). On the other hand, those more concerned with practical laboratory work on spectra are more likely to take "far" to mean 6 to 20μm; and any IR beyond that as being "extreme".*

(iv) Criticism *from the rigorous empiricists* who have not yet understood the postmodern (deconstructionist) attack upon their fundamental assumptions. We might perhaps expect them to make unreasonable demands for rigorous experimental evidence before it is politically possible for anyone to comply — and that could lead to a *"Catch 22"* block against any progress. (See footnote 108 again, and *Book A*).
(v) Criticism of the role of *inductance, L,* or its optical interpretations — by those who can't quite see how these connect with the circuit electricity which they still (misleadingly) liken to water flowing in a pipe.[109]
(vi) And no doubt there will be other misunderstandings also. But time will tell!

So now leaving aside these latter polemical issues, let us go back and look more closely at the three technical criticisms we started with. Of these, *(i):blackness* (IR absorption by water) will be dealt with within the remainder of this present chapter; then *(ii):noise* and *(iii):cable-suitability* in locations as set out at the end of this chapter.

11.3 The blackness of water — from an IR viewpoint

(i) Water is exceptionally opaque to infra-red[110] as we have seen, so how could the signals get through this very *black* medium?

The actual figures for this effect are summarized in *table 11*, calculated from the values given by Zolotarev, Mikhailov, Alperovich, and Popov (1969):

Of course *visible light* has no great trouble penetrating water, and here we are talking about wavelengths ranging from about *.4 μm* (violet) to *.78 μm* (red). Understandably then, in the immediately adjacent "near infra-red" region (wavelength of *1.6 μm*), water is still *mildly* transparent, allowing a reasonable amount of this light to penetrate *1819 μm* (about *2mm*); see the first row of the table.

After that, the prospect looks much bleaker according to our normal standards of transparency — with the typical 36.8%-penetration being only about *20μm* deep[111] (*.02 mm*) — not much more than the size of a typical medium-sized cell.[112] But then maybe that cell-size range is quite adequate for some purposes! Indeed maybe the common cell-sizes have evolved to fit in with this fact-of-nature?

But in any case, not all bodily tissues are watery; and it just so happens that fatty tissues are *particularly transparent to infra-red*. So in this sense, we may now be able to picture any such lipid tissue as a potential fibre-optic channel (*and that includes the myelin sheath* around nerve-fibres). Meanwhile the watery tissues can be imagined as an

[109] There is actually a well documented precedent here. The case I am putting may be new to neurophysiology, but it turns out to have had a long history in a slightly different context — the theory of under-ocean telegraph cables. (The size there is rather different, and the optimal wavelengths for the "surprise" solution are correspondingly somewhat removed from infra-red, being RF or radio-frequency instead; but the main principles are identical). The real pioneer of the inductance effect in cables was Oliver Heaviside (1850-1925), but he encountered tremendous opposition from Sir William Preece, the then head of the British Post Office — a bureaucrat whose understanding of electricity never went far beyond the simple "water-in-a-pipe" model. (Nahin, 1987).
Oddly enough though, this water-pipe model would actually be reasonably close-to-nature ***IF*** we allowed for an extra natural feature: *acoustic vibrations* passing through the water in the tube in addition to the more obvious "flow". See *Book A*, especially regarding the two signal modes for solid rods — total movement *versus* acoustic vibration.

[110] *Why* is water opaque to IR? — "black" when viewed under IR. This is largely due to the resonant rotational frequency for the water molecules. The frequencies of *visible* light are too high for these polar-charged water molecules to keep up with the imposed field; but when the frequency drops to IR levels, the molecules can then move in response — thereby actively absorbing energy from the radiation; (Hecht, 1990, ch.3). Of course *other* organic molecules (and/or the side-chains on them) are also likely resonant "sinks"-or-sources for such radiation (Sorrell, 1988) — and indeed that is one reason why the current theory is so preoccupied with this IR-and-molecule combination.

[111] This penetration ranges from *.075 μm* to *88.94 μm*, as shown in the table.

[112] E.g. see Alberts *et al.*(1983). Cells *can* be considerably bigger than the usual *10 μm*; though it is perhaps significant that some of the largest cells, *120 μm*, are the adipose fat-cells in which the abovementioned water-penetration limitations would clearly not apply (*ibid.* p.24).

 — ISBN 0 9577737 1 4

intensely black shielding[113] which generally protects each channel from unwelcome interference from neighbouring channels — but still allowing just enough range for the message to be assembled or received within single target cells.

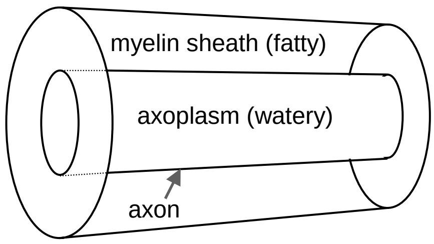

fig. 11:1
Myelin and its relation to its axon

The extended significance of myelin

"Fine. —— But wait a minute! Aren't we talking about electrical conduction through the central watery axoplasm core (and then back through the surrounding watery medium)? And isn't the myelin coating supposed to be an *insulator*? — Yet it seems you are saying the opposite: that the signal travels through the fatty myelin, and it is insulated by the watery media!"

Well, yes actually! This is the dilemma that arises when we think we can model electrodynamics as if it were just simple plumbing — as a "flow of water through pipes" which seems a plausible explanation for how my electric torch operates. But electricity is not really like that, except as an approximation in such especially simple torch-like cases. In fact electricity is still embarassingly wierd, though we try to hide our embarassment by clothing our discussion in elaborate maths. The resulting formulae usually do give us the right answers, but they do not really *explain* the wierd fundamental behaviour, and there is no guarantee that they ever will!

> "Formerly a current was regarded as something travelling along a conductor … But the existence of … electromagnetic actions at a distance … has led us, under the guidance of Faraday and Maxwell, to look upon the medium surrounding the conductor as playing a very important part …." *and*

> "On interpreting the expression it is found … that the energy flows … perpendicularly to the plane containing the lines of electric and magnetic force …"
>
> Poynting (1884).

> "Had we not better give up the idea that energy is transmitted through the wire altogether? That is the plain course. The energy from the battery neither goes through the wire one way nor the other. Nor is it standing still. The transmission takes place entirely through the dielectric. What, then, is the wire? It is the sink into which the energy is poured from the dielectric and there wasted, passing from the electrical system altogether."
>
> Heaviside (1885, sec.2)[114].

> "…electromagnetic energy is transported directly through space[115]. Therefore special devices such as lead wires, transmission lines, etc., are needed not for transporting the energy, but rather for properly shaping the fields in the surrounding space…"
>
> Jefimenko (1966, p.509)
> — one modern textbook.

Familiarity has bred complacency, so we think we understand those occult-seeming things called *"fields" — electric, magnetic*, and so on. We may know, in principle, that circles of magnetic field surround a current-carrying wire, but many of us will still find it hard to believe Heaviside[116] where he insists that nearly all the energy of the current travels *through the space around the wire*, and not through the wire itself! This may be more credible for cases where high frequencies are involved, but Heaviside insisted that it also applied even for those direct current (DC) situations where the old "plumbing" model *seems* to work.

Today we have the benefit of hindsight. After all, radio or *"wireless"* transmission (which is a direct consequence of this approach) can now hardly be denied. But we can understand how bureaucrats and mechanical-engineers of the pre-radio era would find these ideas very difficult to swallow, as Heaviside learnt to his cost! — But what can that tell us about nerve signals and the role of myelin? Let us look first at their commercial analogues in coaxial cables (fig.11:2):

---

[113] "black" in the sense that it would *look very black* when seen through an *infra-red* viewer (as noted in footnote 110). That is just another way of repeating what *Table 11* tells us — that water vigorously absorbs infra-red.

[114] "Electromagnetic induction and its propagation: II. On the transmission of energy through wires by the electric current"; p.437, v.1 when republished 1892/1970.

[115] or, of course, some other insulator (at the relevant frequency), such as air, glass, plastic, etc.; [RRT]. See p.505 (*ibid.*) for more on this frequency-dependence.

[116] Heaviside and Poynting were simply taking Maxwell's (1865) equations to their logical conclusion.

 — ISBN 0 9577737 1 4

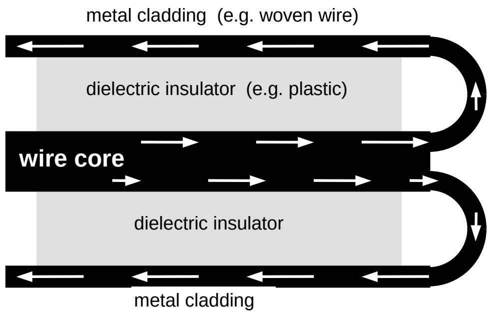

fig. 11:2a. The *MISLEADING* "plumber's" circuit model of transmission in a coaxial cable

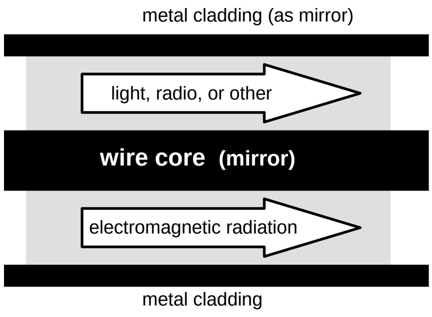

fig. 11:2b. The *CORRECTED* view of energy-transmission *through the dielectric* in a coaxial cable

The fibre-optic interpretation

Note that the "corrected" version looks very like the diagram for a *fibre-optic* transmission. That is no mere coincidence because, in a sense, a coaxial cable is indeed a fibre-optic pathway — though its cross-sectional size and other details will make it more suitable for some frequency-ranges than others.

Anyhow in answer to the question raised several paragraphs back: Yes, despite our habit of thinking of "current flows in conductors", it is better instead to think of the signals as passing optically through the dielectric[117] — that is, via the dynamics of electric and magnetic fields within the fatty myelin! Of course, in practice we will probably not bother to think that way when dealing with the traditional millisecond effects; — but if we are to make sense of the infra-red hypotheses, we really do need to take the more general case seriously, along with its *inductance L*, and the consequent *optic effects* "shining through the dielectric"!

There is a further minor complication which I should at least mention before we move on to the *second criticism*. The commercial coaxial cable we have just considered depended on inner and outer *metal* surfaces to serve as mirrors to keep the light signal from leaking out of the dielectric. That task of containment-by-reflection can also be achieved by *non-metallic* means. In commercial fibre-optics, this is usually achieved easily by having the refractive index of the fibre higher than the medium outside — and then the light can be totally reflected if its angle of incidence is within the right range (*e.g.* Hecht, p104) .

So what about myelin and nerve fibres? On the face of it, we would expect to find the total reflection mechanism just now suggested; but actually there is some reason to doubt whether that is the whole story. That is the main basis for *criticism (iii)*, which we will examine in the latter part of *section 12.3* — from page 63.

Criticism (ii) was the view that ordinary thermal "black-body" radiation would swamp any supposed IR signals. That is a rather more complicated issue, and we will look into it in *chapter 14*.

Meanwhile, in *chapter 12*, we will first make the simplifying assumption that this signal-interference poses no problem. That leaves us free to look more closely at just what the postulated IR signalling could involve — *especially its considerable scope for high selectivity*. After all, this strategy might then perhaps be robust enough to circumvent any swamping-tendency, and hence operate *despite* any intense or unhelpful thermal interference considered in *chapter 14* or elsewhere.

[117] Optical transmission through a dielectric — "fibre optics" — requires some ability to reflect any sideways-directed light back into the system. Note that the concept of electric current is then only secondary, with currents playing the subsidiary role of *causing the reflection* as if they were in relay-aerial systems, or part of a transponder system — analogies which are actually fairly close to the truth.

 — ISBN 0 9577737 1 4

12. Selectivity and Variety — Needed in any Signal Network

How do we distinguish one thing from another? Partly by its own *attributes or "labels"* — and partly by which *"box or enclosure"* it happens to be in.

12.1 Signals with distinctive "labels" — one type of selective communication

Attribute-labels include features like colour, shape, or orientation. In this present context, each attribute serves as a *label* to identify whether each thing is relevant to this-or-that activity — *i.e.* whether it belongs to the particular "set" which that activity calls for.

(In fact, usually the label will depend on *several attributes simultaneously*. Thus a *"go"* traffic light will surely be green; but we may not accept it as a valid traffic light if has the wrong position, shape, orientation, or intensity.)

If the body does use internal IR signalling, it will almost certainly have evolved means for discriminating between various emissions — and that will probably mean using some sort of "label" properties of these emissions. Indeed we already know that nature has used this strategy elsewhere, as in embryology and immunology.[118] There the label is expressed largely by shape-in-space, but there are other possibilities such as shape-in-time. As with the traffic-light example, the most obvious feature in any IR system, would be the *"colour"* or *frequency-combination*:

Carrier frequencies

As we know from tuning-in on a radio set, it is possible to be highly selective amongst a whole plethora of competing broadcasters with different carrier frequencies. Of course there is a limit to how far this selectivity can be taken (and that mercifully prevents us from having an infinite number of radio stations!) but with improved design-or-evolution, the bandwidths can be made quite narrow. This suggests that even given an inconvenient amount of rival signal traffic or unstructured noise-interference, the body may often have been able to circumvent this "simply" by evolving a finer tuning-mechanism.

Frequency combinations — in time-sequence

We need not be limited to the body's ability to discriminate individual frequencies. We can transcend the simple radio-dial scale and use instead an ordered sequence of different frequencies of that sort, like the sequence of values in your telephone number or combination lock, and achieving the same sort of selectivity. (Of course that presupposes that the body can offer mechanisms capable of processing such coding adequately, and we shall return to such molecular-level issues in *Book 3*).

Moreover such modulations of the frequency are indeed known to photochemists, notably those who study what goes on at the time-scale of *femtoseconds* and *picoseconds*, (10^{-15} and 10^{-12} sec. respectively, see section 11.1 above). Mind you, the typical modulation considered so far within femtochemistry appears to be of a limited type: a *"glissando"* to put it in the language of music — a run of monotonically increasing-or-decreasing frequencies rather than an arbitrary mix of all values in the range. Anyhow *"chirping"* is the technical term used here by the femtochemists; (Grishanin *et al.*, 1994; Duppen *et al.*, 1994; Kohler *et al.*, 1994).

Frequency combinations — direct mix

Then of course frequencies could be combined in other ways, like the *simultaneous* mix which gives us most of our real-life colour perceptions. For instance, the "yellow" we see on a TV screen is actually a mixture of red and green light, and not the monochromatic yellow light we might have expected. Other real colours will usually be mixes of this type, with an extended spectrum of components which *could* carry a signal *if* our eyes or instruments were set up to do the necessary analysis. Here the eye has missed an opportunity to discern further data (perhaps because it already gathers all the information it can usefully handle). But maybe *molecular IR-receptors* in a similar situation *could* have grabbed at this opportunity of the extra **spectrum-analysis** information — thus effectively serving as miniscule eyes.

Polarization

Another opportunity which is missed by the human eye lies in the information carried by the polarization of light — unless, of course, we are wearing Polaroid sun-glasses. Some other animals are free from this limitation, and the same might well be true of any IR detectors within the body, whether at the molecular level or not.

Direction-of incidence

The direction-of incidence is also important (as in the traffic-light example where a north-facing green light is hardly the same as one facing west); and anyone who has tried to adjust a temperamental TV aerial will have some further insight into this sensitivity. Such angulation clearly applies to reflection and refraction processes as well, especially where *total internal reflection*[119] is used to restrain light from crossing an interface. This is particularly important in the theory of optic fibres, and hence of any optic-

---

[118] The body's ability to recognize keylike molecular-shape patterns in *3D space* is now common knowledge. Within *embryology* the technique is used for recognizing target sites for growth-cones; while within *immunology* it is used for recognizing unwelcome antigens.
In either case, the antibodies *or* the growth-cones will latch onto special highly-specific *stereochemical shapes*, in the "lock-and-key" matching process. The difference here is that we are now considering *time-based patterns* as well; but the principle is essentially the same.

[119] E.g. see Hecht (1990), page 104.

fibre role that myelin might have; so this is a topic we will return to shortly.[120]

Optical interference patterns

Let's come back again to the single frequency case, but now consider beam-splitting and the static optical interference patterns between the different subsidiary beams of the coherent light. Holograms are a spectacular example of this effect, made possible commercially by the development of laser technology. (Remember that a laser beam ideally has its emitted light all neatly timed so as to be in the *same phase* across a well-defined wavefront). In principle then, a hologram depends on our splitting off a part of the source-beam for use as a benchmark for the phase; and this allows us to reconstruct the virtual 3D image in empty-space — of something which is not actually there.

Holograms as such may perhaps be dismissed as just an elaborate party-trick. But at a somewhat more modest level, such interference effects can still be very powerful, and they have many far-reaching effects — for instance in navigational aids developed during World War II (Crowther and Whiddington, 1947) and in the subsequent commercial world, (Decca, 1976; Thom, 1985), and ultimately the world-wide satellite-navigation system, *GPS*.

But even without such commercial-or-military precedents, just consider the potential power of these optical interference patterns within biology.

Firstly they enable narrowly focussed beams to emerge from an apparently messy mix of electromagnetic activity when the conditions are right. That happens in the space around an RF (radio-frequency) transmission aerial, and it could be just as applicable for IR-frequency activity at the molecular level.

Secondly such stable interference patterns can-and-do often form a network of "gridlines" which offer a powerful means for repeatably finding *exact locations*, or any other use we might expect from gridlines on a map.[121] (True such gridlines will normally be hyperbolic rather than rectangular, but that difference is of no great importance). This technique has been successfully used in aircraft navigation for more than 50 years, and could be useful in real-time memory access within the brain. Moreover if it *could* be thus used, then evolutionary pressures are likely to ensure that sooner or later it *will* have been used for location-finding processes within the body — in the brain or elsewhere, and possibly both.

Redundancy

There is another feature which greatly increases reliability, including the reliability of labelling. This is the use of redundancy — using the same coding several times as a safeguard against localized mistakes or information-loss. If a particular memory is recorded in just one place by whatever means, then there is considerable scope for mistake or irretrievable loss. But if the encoding has many copies, and if their communication links are also highly coordinated (by laser-like signals or by other means), then there is little scope for stray noise-interference signals to disrupt the system.[122]

12.2 Selectivity and insulation

But let us not get carried away with the special possibilities raised by these various attributes. Here my sole immediate purpose is to show that any IR-signal receptors are likely to be highly selective, and so most unlikely to be activated in a counter-productive way by *any ambient random stimuli*. Thus if *receptor x* can only be triggered by a signal of *frequency* ν_1 which modulates to *frequency* ν_2 within a time-lapse of Δt, and if that signal must arrive at a site with coordinates within ($\Delta x, \Delta y, \Delta x$), and travel in a direction within the bounds of $\theta \pm \delta\theta$ and $\phi \pm \delta\phi$; then it is most unlikely that any stray photon-combination arising *at random* from black-body radiation could "break the code" — no, *not even if the black-body radiation were fairly intense!*

And we have not finished yet either. The receptor might also require the *polarization* of the incoming signal to be within an angle of $\alpha \pm \delta\alpha$ and it might well also be necessary for a majority of (say) *60% out of 20* such receptors to all get the same message "simultaneously" (within a time-tolerance of τ *femtoseconds*) before any significant action would be taken.

Of course this is all hypothetical, and the threshold values are merely algebraic and unspecified; but my point here is simply to establish a *prima facie* case — to show that any claims *"that black-body radiation within the IR band would necessarily swamp any potential IR-signals"* would be overstepping the mark at this stage.

Obviously we cannot foretell what future evidence will disclose. Later research *may eventually* show us that the selectivity of all such potential-receptors *really is* too lax to allow for selective IR signals after all. But until such evidence arrives, it is perhaps fair to say that any such critical claim is currently premature and even implausible.

Meanwhile there seems to be a persuasive enough case for pursuing the *IR-signalling-system* hypothesis further — even to the extent of investing in some expensive experimentation!

---

[120] Also see works such as Nöckel and Stone (1997) and the references cited therein concerning multiple reflections within a cavity, and the optical patterns which result.

[121] See *Book A* or *Book C* for more detailed expositions. In this current book the topic of *cell navigation* is only of peripheral relevance.

[122] Note that any *molecular based* memory system would be well placed to offer this sort of redundancy — on a massive scale if need be; and note that both the genetic system and the immune system illustrate this technique at just such a molecular level. In contrast, a *synapse-based* system (taken on its own) could afford only limited amounts of repetition; and given the relative floppiness of its elements, such replication would be of dubious reliability anyhow — not to mention the logistical difficulties in establishing, coordinating and operating the replicated-elements.

 — ISBN 0 9577737 1 4

One *could* rest the case right there, without even looking at the claims of black-body theory (apart from our obvious use of the fact that it is seen to be *random* in its occurrence). However let us press on nevertheless, as follows:

First we may look at *"fences-or-reflective-walls"* as a further class of constraint which might help to keep disruptive external radiation at bay — in case that turns out to be necessary after all.

Secondly in *section 12.4* we shall look at the view that the real external danger is from *rival* meaningful coded signals ("crossed wires"); and that the *absorption* actually does a useful job even if it also produces intrusive noise. That useful job would be to scramble those rival vagabond signals before they intrude — leaving the relatively harmless randomness of thermal radiation.

(Chapter 13 digresses a little to consider the positive value of • safeguarding, conserving and facilitating the signals that are supposed to be there, as well as the negative value of obstructing the unwanted ones. It also suggests • that such devices as reflecting walls and tuned-receptors might well serve *both* roles.)

Only then, in *chapter 14*, do we really get down to looking properly at black-body radiation itself:
• The *assumptions* which underlie this notion of black-body radiation — and their shortcomings in any real-world situation (which inconveniently contains "lumpy" molecules whose properties actually depend on frequency in a somewhat messy idiosyncratic way).
• The actual *total* radiation-intensities predicted for a given temperature.
• The radiation-intensities expected within a specified attribute range.

And to this we may add some thoughts about whether it is more meaningful to count the flow of thermally-produced *photons* passing through an "active site" — or the more conventional counting of the associated *energy*[123] passing through that site.

12.3 *Segregating IR messages by "mirror fences" — or by labels as before*

What techniques do we use to *define* a set of objects or events? — "Mammals" — "Red-things" — "My-things"? —— And, of more interest to us here: *"Signals-relevant-to-receptor-x"?*

Intensive and extensive definitions (for members of a set)

Membership definitions are often classified[124] into two types: ***Intensive* definitions** which identify members of a set by telling us their attributes; thus *"Mammals are living beings which suckle their young and have hair"*, and such definitions can be applied even to unruly members which we will never encounter nor even be aware that they ever existed. Likewise, in the last section, the postulated IR-receptors were seen as accepting only those signals which fitted some rather demanding selection criteria; so in effect these receptors were applying an *intensive definition* to identify which of the incoming transmissions were "relevant".

***Extensive* definitions** effectively *fence-off* or *tether* the chosen items as the way to separate them from non-members. For instance "everything in this room" — or "items on this shelf" — or "everyone connected to this private phone-line".

There is no point here in going into quibbles over exact meaning, or demarcation disputes.[125] For the moment, a crude commonsense distinction will suffice. Anyhow let us notice two things about this divide. Firstly, living things spend a lot of time identifying aspects of their environment by using intensive criteria, *i.e.* recognizing things like "potential-food" by some implicit list of descriptions; — and then setting about "fencing in" (perhaps swallowing!) those things which are of most interest, and that amounts to extending the already-*extensive* definition (via intensive criteria)!

> (A similar process is common in mental activity too. We have our own private definitions of "friend-in-general" and we then decide whether to put John-or-Mary's name on that *list* — a new extensive definition.
>
> Or as a society, we may first painstakingly use an intensive definition to decide, for instance, that a whale or a platypus is a mammal — despite some initial misgivings. We *then* become dogmatic and create an extensive definition by placing its symbol-or-name on our list of mammals; where that list is the symbolic equivalent of a fenced-off region.)

Secondly, and rather more important here, our life can often be made easier if items or signals are kept sorted into various "pens" (extensive categories) right from the start of their existence, and held there throughout their usefulness. That is a brief theoretical justification for that everyday phenomenon of the information-channel, which is largely isolated from the world-at-large — and from *other* information channels. *Telephone lines* and *nerve-fibres* are two clear examples:

If I want to send a message to you on the other side of town, I *could* use a very powerful loudspeaker and yell out the message prefixed with your name and perhaps other

---

[123] in *electron-volts* or *joules*, or whatever.

[124] classified by mathematicians and logicians; and also theoretical psychologists like J.Piaget who draw upon this tradition. Of course this is the dichotomy implied informally at the start of this chapter, with talk about "attributes" and "bins" respectively.

[125] I.e. demarcation disputes like • whether, during the discussion of *extensive* definitions, the word "item" implies an embarassing *intensive* definition (a list of descriptive criteria which would, for instance, distinguish *objects* from "mere" shadows); — or • disputes whether the abovementioned "$\Delta x, \Delta y, \Delta x$" criterion (for assessing the relevance of a signal event) is really an extensive or a descriptive-intensive condition, or both.

 — ISBN 0 9577737 1 4

identifiers. You might even receive the message, despite the wrath of my neighbours! But if everyone used the same technique routinely, we would all have increasing difficulty in discriminating whatever was relevant to us. And on a world-wide scale (now using unselective radio *broad*casts), the system would obviously be totally unworkable, especially if we had to repeat the prefix with every word we sent!

What we usually do, of course, is either meet face-to-face in what amounts to a loosely fenced-off local environment, or else we use telephony in one form or another: I use an intensive definition (your phone-number or e-mail address) which then supposedly allocates a stable communication channel "fenced-off" from everyone else — an easy-to-apply *extensive* definition.[126] Or indeed I might use a dedicated private line in which no dialling is needed, and that really is an extensively defined domain for my messages, with a minimum of doubt about their reaching the intended target.

Application to mind-theory — the suitability of myelin

If we stop to think about it, it is obvious that any effective communication channel will (*a*) allow signals to travel within it, but also (*b*) it will hinder signals from entering or leaving its domain except at relatively few designated input or output sites. In other words it provides a *fenced-in* domain with few gates. And that raises the question as to just what that *fence or tethering-system* is to be made of in any particular case.

In an *un*myelinated nerve-fibre, for instance, the action potential (voltage-"spike") is transmitted along the cell membrane; so clearly the signal is *tethered to* that membrane, eventually activating synaptic mechanisms (its designated output-points), and possibly other structures relating to growth and maintenance (which could also be seen as designated outputs). The signal might also *incidentally* affect other mechanisms — including your laboratory's measurement-probes.

We can return now to the more familiar *myelinated* fibres[127] when they are dealing with the same conventional action-potential spikes. Here the signal keeps swapping between two different types of domain:

• the active membrane[128] as above, *and*

• a capacitative transmission along myelin segments. — And here the transmission will be tethered to the myelin, essentially by electrostatic attraction between positive and negative charges separated by the dielectric barrier of myelin. (At least that is the orthodox view; and that will serve here quite well).

As for the designated input/output points: they will now often be the interfaces between the alternating subdomains of these two types:

membrane at a node → adjacent myelin →
membrane at a node → adjacent myelin →
membrane at a node → adjacent myelin →
… *(and so on).*

But what about the extra transmission-mode postulated in this book? Any *IR* signals sent into the myelin (in addition to the above, much coarser, orthodox "spikes") would amount to *fibre-optic transmission*, and would have to satisfy many of the technical requirements currently considered within that fibre-optic industry; though one significant difference would arise because IR-waves ($\lambda \approx 1$ to $20\ \mu m$) would be a fairly tight fit in myelin cables of about that same range of diameters, whereas commercial fibre-optics (which usually uses IR too) is free to use rather larger fibres. This may not sound important, but it does make quite a difference to possible optical interference patterns and modes within the cable. However this narrowness is no actual bar (as long as the fibres are not too much narrower than λ)[129]; though it does affect some of the expected properties of the system.

So what could serve as the fence in this case? This question arose at the end of the last chapter in relation to fig.11:2b depicting the myelin as a possible fibre-optic channel. The simple answer is that the boundary should be a *mirror* of some sort — and like the mirror surfaces of a thermos-flask, it should ideally keep all relevant internal radiation right there on the inside, thus playing the important role of conserving the signal and the energy which underlies it. But then this mirror should also play the "opposite" role of *keeping out* all the external radiation (random or otherwise) which might interfere with its own internal signals.[130] In effect then, we might perhaps expect to find

---

[126] Whether this is really extensively defined in today's world is rather less certain, but to the naive user it *seems* to be a dedicated channel, and let us leave it at that. After all, if necessary we *could* set up such a pure system just to illustrate our point.

[127] See fig.11:1 — and the related discussion explaining the less obvious aspects of wire-like conductors (such as axons) and their surrounding space (such as myelin insulation).

[128] membrane free of myelin at the *Nodes of Ranvier* between myelin segments along the axon, and at the unmyelinated non-axon extremities, (including dendrites, and the cell body).

[129] The absolute minimum-width for *square* waveguides is $\lambda/2$; and for a simple *uncored round* cylindrical tube the limit is similar (though an exact solution involves Bessel functions). For a cored *coaxial* cable, (i.e. the shape of an ideally-regular myelinated axon), the situation is a bit more flexible because this configuration allows and encourages the so-called *TEM* mode, as used in TV aerials. This allows the transmission of longer wavelengths which would not "fit" into a simple uncored fibre.

[130] It is not entirely clear just what sort of mirror surface might operate here; and that uncertainty is one basis for *criticism (iii)* introduced in section 11.2. For normal commercial fibre-optics, the mirror effect is achieved by *total internal reflection* (Hecht, *p.104*). On the other hand, with metal-clad coaxial cables, the boundary reflection is due to electron-mobility — and logically enough this is called *"metallic reflection"* (Hecht, *p.112*), though it is not actually essential for metal to be present if adequate electron-mobility can be engineered in some other way (such as the use of conjugated double bonds?). However, for the moment let us overlook these difficulties and simply consider the implications of such reflection *assuming it exists* — and insofar as it exists.

the signal-related IR radiation herded into separate fenced-off pens — one for each segment of myelin — and meanwhile the "noise" of other thermal emissions would roam around in the "wilds" outside.

Technical problems

Such perfect segregation would be unlikely in practice, if only because the reflection would be less than perfect for several reasons which we shall look at shortly. But even if it only partly applies, we could expect it to offer some degree of protection. Let us look at random and non-random interference separately:

First consider that the reflective boundaries would offer some extra protection from the *random noise outside* each fenced-off domain.[131] However I have already argued that, thanks to the likely selectivity of receptors, random noise is probably not a serious threat to the integrity of signals anyhow — not even if this noise is fairly intense (and I shall later suggest that it is actually rather mild in any case). So any such *extra protection-from-randomness* may now be seen as possibly superfluous.

But secondly there is the question of *non-random* interference in the signal — the "crossed wires" disruption we must surely expect from neighbouring communication channels in a crowded network *unless* some careful design-or-evolution has gone into preventing it![132] At first sight, the supposed mirror-like boundaries would seem to suffice to prevent this disruption for our supposed IR system, but mirrors have their limitations, especially when their active surfaces are very close to each other — i.e. at distances comparable to the wavelength in question, or less:

For one thing, "total internal reflection" (*TIR*) can only be *"total"* within a certain range of angles-of-incidence *within* the optic fibre. To prevent escape, it works best when the light is nearly parallel to the surface (the general direction we would expect our transmission to be travelling in anyhow); and reflection is minimal when the light falls directly onto the surface at right angles to it. Moreover TIR, being "internal", cannot work in reverse to actually prevent invasion from outside, though any such incursions would be angled differently from the rays trapped inside, so that would possibly set them apart. Indeed these intruders would necessarily fail to fall within the required angle-range for TIR, so there would be a good chance that they would simply *exit again* when they reached the opposite boundary, provided the boundaries remain reasonably parallel.

Then, in the case of so-called *metallic* reflection, the efficiency of the process will depend on just how free the mobile electrons are. In other words, the *conductivity* of the metal-like surface is a factor to consider. (Born and Wolf, 1970, *ch.13*).

A rather more surprising and significant feature is that mirrors of both types do not actually transact all their business at the exact geometrical surface, but rather within the reflecting medium itself, with this activity usually attenuating exponentially such that it eventually becomes negligible at a depth of a few wavelengths. This means that "tunnelling" phenomena can occur through supposedly-insurmountable walls (as happens also with quantum barriers) such that the light can-and-does penetrate the mirror surface if it is thin enough. (Hall, 1902; Hecht, 1990, *p.85*; Lavin, 1971, *ch.2*; Born and Wolf, 1970, *sec.13.3*).

This all seems to suggest that while such boundaries must be adequate to prevent undue loss of signals during their transmission, there nevertheless remains considerable doubt about whether these boundaries would suffice to deflect *unwanted signals from neighbouring channels*, especially if their coding-conventions were similar enough to be mistaken as "legitimate for *this* channel". So if the mirror boundaries are not a sufficient protection on their own, what other device or procedure could there be?

12.4 Further protection by scrambling rival signals — (optical absorption)

According to the argument previewed above in section 12.2, the real danger to neuro-signal traffic comes from rival encodings from other channels and *not* from random noise. If that is indeed true, then one solution becomes obvious:

Convert the stray signals into the relatively harmless random mix — "digesting" the invading signal-structure, and breaking it down into the bland primitivity of black-body radiation, or something approximating it.

It is now also obvious what the mechanism is likely to be. We have already seen in section 11.2 that water is *intensely "black"* when seen from an IR-radiation viewpoint. Thus any IR signal which escapes from its proper channel is likely to be degraded into "heat-vibrations" within about *1 to 90 microns* of travel, depending on the frequency. (See *table 11*; or Ray (1972) for a broader view). And traditional abstract physics tells us that this "heat" will sooner or later manifest itself as black-body radiation with its random-based characteristics and spread of frequencies.

---

[131] That would not save it from any black-body emissions *within* the domain, though given that the chemical composition of that domain is likely to be fairly standardized, its radiated emissions will probably be quite unlike those random outpourings of the ideal "black body", and therefore much more amenable to control.

[132] Has this problem been considered regarding the *orthodox synaptic-circuitry* aspects of neural communication? Perhaps; though I am not aware of any such work. Maybe the chain-reaction transmission of action potentials along the cell membrane is sufficiently robust to resist disruption (though that is not entirely obvious), but surely the electro-capacitive aspects of saltatory conduction associated with myelin would be vulnerable to interference from neighbouring myelinated fibres?

 — ISBN 0 9577737 1 4

Admittedly this will only approximate the ideal black-body radiation of theoretical physics, which predicts a smooth *single-peaked graph* for absorption-or-emission[133] at any given temperature, as shown several times in the chapter 14. But in reality, the IR-blackness of water has about *five* main peaks in absorption-rate (each being a local minimum of penetrability) as shown in *table 11* (page 57); so this clearly departs somewhat from the theoretical ideal.

Nevertheless this should still be quite sufficient to mop up and scramble most of the vagabond messages which happen to wander loose. After all, three of the "less black" regions are still quite opaque: with penetration distances of only about *43, 19, and 8 micron* (at wavelengths of *5.3, 7.7, and 38 micron* respectively). This leaves only two substantial gaps[134] in the barrier, with penetration-distances of about *2 mm, and 89 micron* (at wavelengths of about *1.6 and 3.8 micron* respectively). Moreover evolution, with its opportunistic ways, may well have found a useful application for these mild gaps in the defences.

[133] According to Prévost's theory of exchanges (of 1791), the emission-rate will equal the absorption rate whenever the relevant body is at the same temperature as its surroundings. In practice then, the ability of a surface to *emit* is taken to be the same as its ability to *absorb* radiation of the same sort — so the two terms tend to be used interchangeably.

[134] Here we are talking in terms of a relative coarseness in the spectral graph — "broadband" as Ray (1972) puts it. We would of course find more exceptions if we were to go in for finer and finer resolution, but that would probably merely effect the detail, and not the general comment being made here.

13. Consequences of these 'Labels' and 'Fences'

13.1 The need for both fences and labels — which then allow "multiplexing"

What does all this tell us about the likely situation in the nervous system? For one thing, any complex information-system will probably need to use *both* intensive *and* extensive definitions to control its signals such that they normally reach appropriate destinations. We have already considered how intensively defined label-items can come to have *extensive boundary-definitions* as well; but the trend can also run the other way, starting with boundary definitions alone. In the commercial world, this became evident quite early in the development of the telegraph:

Initially of course, the situation was simple and each wire carried only one morse-message at a time — each neatly "tethered" spatially to its wire, *extensively*.[135]

But Heaviside and others soon realized that it was possible to *multiplex* different signals on the same ("fenced-in") wire simultaneously. That is to say, several different carrier-frequencies were fed to the wire at the same time — frequencies which were all significantly higher than the square-toothed signal that the expected series of morse-dots would give us — carriers *which could later be separated again* by tuned circuits at the far end of the wire (because the carrier-frequency served as a "label"). It was then possible to pattern each carrier separately: modulating its amplitude with a "dots-and-dashes" message, which could then eventually be separated out again, along with its carrier.

In other words, a commercial communication channel which had been defined in purely extensive *boundary* terms originally (when the system was primitive and simple), later developed some intensive *label* definitions to ensure an appropriate signal separation[136] in the face of growing complexity. So what was initially just extensive or "fenced-in", *came to adopt intensive-"descriptive-label" properties as well*. This now leads us to the important but oft-neglected question of:

How does a signal "know" which path to take when it is offered a choice?

It seems we commit ourselves to a "fenced in" domain like a *path* for one of two reasons[137] (or perhaps a combination of them):

• Assessment of attributes, an evaluation of intensive definitions — and that does seem to be what we usually mean by the term "decision-making".
• Because of some pre-existing extensive definition — *i.e.* because we are already on a path which leads into the one we are supposedly "choosing". Arguably this is no choice at all, or at best it is a "forced choice", and it does not accord well with the notion of *decision making*.

In fact extensive boundary definitions usually present as dogmatic and *unthinking* rules handed down from the past, no matter how valid they may or may not be: *"This drug is on the banned list, so you cannot have it no matter what arguments your doctor may offer about its attributes!"*

In brief then, if we are serious about studying such mental processes as decision-making and creative abstract thought, it is not enough to concentrate on "wiring" and implied boundaries (extensive definitions), important though they may be. We really must also have some ideas about how the nervous system copes with intensive definitions and "labels".

13.2 Labels for use in the computer — and elsewhere

The computer

How do computers manage? Note that much of their present power is actually due to their ability to handle labels: *address-labels* (including generating or calculating them when appropriate) and *search-labels* — keywords, attribute codes, and suchlike. Path choice is usually made on the basis of some comparison of attributes which obviously implies the use of intensive definitions.[138] There may well be a lesson for us in this.

---

[135] Even then, the matter was not quite so simple if we consider *time boundaries* because there was always a need for some sort of "goodbye" signal as a meta-linguistic message within the text — effectively an intensively defined submessage which served the purpose of a "time-based fence" to separate each message from what followed!

[136] Of course the *human telegraph-operators* had always applied some sort of intensive-definition-procedure to decide which messages should be directed to which channel and when. At first sight we might dismiss this as a separate issue. But of course, the minds of such human operators should be very much on our agenda in this book about the human mind! Such human agency is easily overlooked because of its very familiarity.

[137] There might be a case for including a third category: • *Random influences over which we have no control*. However that would take us back into the philosophical minefield of chapter 6, without doing anything much to clarify the present discussion. Suffice it to say that there is a case for supposing such cases could be apportioned to the other two categories — though we might have trouble agreeing on just how to do that.

[138] This was not always so to the same extent. A common programming technique of the 1960s was to use *switches* within the programming code-sequence itself (like setting railway-points) which meant that, later on, the course of events was determined blindly by this forced "choice" imposed on the tethering system. However this technique soon fell out of favour, and is now virtually unheard of (and even the word "switch" now has different computer connotations).

 — ISBN 0 9577737 1 4

Table 13 — COMPARING NET-LIKE SIGNAL-SYSTEMS

SIGNAL HANDLING SYSTEMS

	CONTROL-*AND*-PROTECTION for SIGNALS: *intensive* definition by **"labels"**	*extensive* definition by **"fences"**
Telephone system	Phone numbers	Wire or cable; &/or focussed microwave-beam
Computer	Memory address; keyword; etc.	Dedicated wires, and shared *"bus"* wires
Postulated molecular-IR system	Tuned IR incl. *chirping?*	Reflection at myelinboundary? Optical beaming?
Traditional SYNAPTIC-neural system	**???**	① "spike" tethered to cell membrane; then ② +/- volt-attraction across myelin sheath. ③ ? Transform and lens-like focussing ?
Lens systems	none	Focussed beams; Reflection; etc., in Euclidean media
Fourier and other transform systems, including *Recognition networks*	none?? (see text)	"Neural Network" of elements usually linked in a non-Euclidean way; so *"Huyghens' constructions"* here yield special effects.

Of course for computers to achieve genuine creative thought, they would need more than just an ability to cope with labels. As suggested in *Part I*, the biological brain seems to require a number of special "digitally orientated" features — and similar requirements would quite likely apply also to any artificial substitute if such a thing is possible. The actual details[139] need not concern us right here, but the crucial point is that any features of this sort will surely need some kind of "labelling-procedure" as a minimum requirement — in *any* system which purports to produce creative thought.

Comparison with IR and synaptic systems

Computers do use such labels within signals; but how feasible would this be for the brain? If *the IR system* does exist then such label patterns are fairly easy to envisage, and indeed that is one of the significant attractions for the IR-signal concept.

On the other hand it is far from clear that the traditional *action-potential signalling* could manage this feat unaided. There are questions of sequencing which we will look at shortly; but there is also the problem of precision:

The comparatively large body of the neuron-cell, and the seemingly ramshackle multitude of casually-placed synaptic connections on its surface do not inspire confidence in its accuracy and reliability — and such skills would seem to be needed for efficient label-handling. No doubt this whole-neuron has other merits, but perhaps they are the analogue virtues of a Picasso portrait (as compared with the alphabetical virtues of a Shakespeare text).

In principle we could have a copious replication of such synaptic circuits and so acquire our precision statistically; but the relative largeness of such cell structures would not help the prospects for the sort of replication that would be required; and in any case, how would such replication be established or coordinated during its operation? Seemingly it would need some unseen additional telegraphic-or-telepathic system to guide it!

(These difficulties would seem to be comparatively trivial in the rival system, the here-proposed case where the labels are encoded on RNA-like sequences and/or patterns of IR emission. These would have the strength of precision at the ultramicro level, which could then bestow a more reliable basis even for their statistical effects).

These thoughts are summarized in the centre rows of table 13, (see the bold type and shading). It is illuminating to compare these two candidate nerve-signal types with the other four systems from the better-understood world of technology:

At the top, we have the two types of system which we have already discussed, computers and telephones — systems which clearly do make use of labels within their signal-text, and which would thus probably have much in common with the molecular-IR system if it actually exists.

[139] The features include: • a large capacity for rapid trial-and-error mechanisms at the molecular level — • a redundancy of "digital" coding at the same level, with a signalling system capable of communicating with it, and coordinating it — • the ability to organize its elements into structure-like ensembles, and establish hierarchies amongst them — and so on. See above, plus *Books A and D*.

 — ISBN 0 9577737 1 4

At the bottom we have two lens-like[140] systems which seem to have no clearcut label-identification strategy to help in guiding signals into their supposedly-correct paths, though they do sometimes have impressive means for signal-manipulation. Indeed they often do serve to *recognize signal patterns* of certain types (such as in character-recognition), so why not accept this as label-identification?

In *section 4.4* we discussed the advantages of *1D* encoding for labels; e.g. that such linear codes are by far the best way to ensure a thorough scan of the available code, allowing for faithful replication and other precision-based properties. In that light, let us now insist that any fully respectable *"label"* must ultimately hold its coded potency in *one-dimensional digitally encoded form*.

Any other identifiers would then be *less respectable*, but in what way? Here we are concerned with neural nets, and lenses:

The capability of neural nets is well recognized, and they have been copied commercially, so where are their weaknesses? Well, for one thing, we seem to be dealing with analogue systems, which are not good at capturing certain types of precision even though they may have other virtues. Then again, each exemplar will contain a wealth of linkages which capture some aspects of experienced reality — but these are linkages which would usually be uneconomic to record and reproduce blindly in other systems. In short, for any new setup, it will often be more efficient to simply start again from scratch with the network settings. We might say that their patterns are usually too untidy — too unpredictable.

Lenses may be seen as very special cases of neural nets; but in contrast they are *much too tidy!* These cases are so special that there are one-to-one correspondences between input and output. Moreover if points A-and-B were immediate neighbours, then their images A'-and-B' will also be immediate neighbours, so any image mappings will conserve connectivity. Such a special system does not to have its linkages established, as they are already inherent in the system. (But then a lens system, taken alone, seems ill-suited to *recognize* any of its images. That would seem to need a more complex connectivity which converts a mostly analogue input into a loosely digital selection-decision — applied to a finite list of things which might be recognized.)

It might seem then that any identifiers which depend on *networks-or-lenses* will generally be inadequate for sophisticated manipulations — especially for such tasks as forming hierarchies of concepts. Perhaps then we might use the term *pseudo-label* for any identification-feature which a network can use. Or, getting away from the term *"label"* altogether, we might choose another such as *"badge"* which has *3D* connotations[141] sufficiently remote from the *1D* case.

Anyhow these two systems, listed at the bottom of the table, would seem to have much more affinity to the nerve-system as traditionally understood, *i.e.* the synaptic system with its voltage-spikes. Indeed it is, of course, a well-established practice to refer to all non-lens networks of this type as being "neural nets" — whether literally or metaphorically.

Summary

The argument within this section has not been rigorous, but it does suggest
(1) that there are important roles for both *fence* and *label* control of signals;
(2) that the postulated IR-system (like computers and telephones) would tend to be best at the label role, *while:*
(3) the traditional synapse-and-action-potential system would be best at the *direct delivery* of standard message-patterns — (and in this, it would be like lens systems, and other sorts of "neural net").
(4) There is probably ample scope for a symbiotic cooperation between the two neural systems, if they both exist. And note that if they *don't* both exist, there will remain many questions left unanswered until someone suggests a better explanation.

---

[140] It may not be immediately apparent that there is a formal similarity between these two types of signal-conduit. Here it is helpful to try applying *Huyghens' construction* to each at the micro-level. A clear medium like glass is effectively a very regular network which transmits a wavefront of light in a very regular way — a way which is readily explained by the straightforward application of Huyghens' construction (e.g. see Hecht, 1990, pp.80-81). Logically though, we should be able to extend this approach to any medium which constitutes a network of communication links *even if their linkages are seldom directed to near neighbours!* By normal optical standards, such processing might seem bizarre; but if it is stable enough in the long term, (and yet mildly amenable to fine-tuning), then it has significant potential for performing such feats as character-recognition or reversible Fourier-transforms, etc. — comparable in some ways to what a lens does when it forms a clear image on a screen. Indeed, as I have implied, a lens might be seen as a very special case of a "neural net".

[141] It is perhaps unfortunate that the word "label" itself has *2D* connotations. We should, of course, focus on the *1D written-text* on our luggage-labels — rather than on their physical *2D* area!

 — ISBN 0 9577737 1 4

14. "BLACK-BODY" RADIATION — WOULD IT HELP OR HINDER *IR*-SIGNALLING?

14.1 *Viewing cells through "IR eyes"*

We often see pictures of cells as seen through microscopes of various sorts — 2D images which usually serve their purpose well, but which edit out some of the possible ambience of the cell's local situation. Accordingly it might help here to imagine ourselves as gifted with a special type of vision.

***FIRSTLY* suppose we could see our ordinary full-sized environment by infra-red,** and *not* by normal visible light. Most metal surfaces would look silvery, even if they had looked dull and rough in normal light. This is because the less-demanding slower frequencies of IR enable more of the metal's electrons to resonate effectively, and hence re-emit the signal as a better-than-usual reflected beam. Likewise any organic compounds which offered electron mobility would also be likely to act as silvery reflectors — and that could conceivably apply to some tissues of the body.

Meanwhile what would water look like? *Table 11* (page 57) tells us that water will be very opaque. In fact it will be a very effective absorber of any of our IR which might otherwise penetrate it; and in terms of traditional *black-body theory*, we would thus have to call it "black". But then in the special conditions of thermal equilibrium envisaged by that theory, this "black body" will also emit an equivalent amount of radiation — though not necessarily from exactly the same site or with any locally-predictable frequency.[142] So in some circumstances, water will indeed seem black in the ordinary sense of the word, often it might just look very muddy, but at other times it might look more like glowing molten iron from a blast-furnace.

It is actually a major task of this chapter to judge whether water will normally look like that molten metal (as seen through IR eyes), and if so, whether it would be likely to upset IR signalling. In fact as a first defence against that signal-upset proposition, we might think of an analogy in ordinary light:

Inside an iron foundry, we would normally still see the rest of our environment quite well even if there were quite a lot of glowing molten iron around. Indeed it might even be the light-source which illuminates an otherwise dark work-place at night.

In other words, the randomized signal-free "black" emissions would yield energy to other items in the environment which could *then* add significant signals into the radiation before it reached the eye. And isn't that what *normally* happens when we switch on an electric light-bulb?

Admittedly, in some circumstances, we *could* be dazzled and thus incapacitated; and if ever the whole system is in thermal equilibrium (as pure black-body theory assumes), then the radiation will be coming uniformly from all directions — as in an Antarctic "white-out" blizzard — and then we might indeed lose our way.

SECONDLY now, **suppose we could shrink ourselves** (and our IR vision) down to the scale of the cell's environment, **so that a micron (*µm*) would *look like a metre***. Water would now not look quite so opaque, but rather more like a *fog*, (or a cloud of black dust — depending on relative temperatures).

Table 11 tells us that for the near-visible wavelength of 1.6 µm, we could see for about *"2 Km"*, while for 2.93 µm, our vision would be limited to about *"75 cm"*. At other likely wavelengths, visibility would range between *"24 m* and *89 m"* — a workable distance for many local activities.

Meanwhile any surfaces with metal-like conductivity would still appear like silvered mirrors, possibly shielding various domains from any dazzling white-out effects which the water-"fog" might produce. Likewise, if Cope (1973) is right, each mitochondrion will probably have such silvery walls (like a thermos flask) to store a standing wave within its interior resonance cavity, and nerve-fibres too will have a silvery surface.

If we could cut into a myelinated axon (normally *"1.4 m to about 10 metres"* in diameter[143]), we could peer into the comparatively clear optical-pathway of the myelin dielectric. We would perhaps be blinded by the light-beam of signal traffic from the far end of the segment (perhaps *"2 Km"* away). If we could see them clearly, the sides would look silvery due to internal reflection and possibly aided by any metal-like shielding At Nodes of Ranvier, the uncut wedge-shaped ends of the myelin "pipe" *would perhaps be seen (at certain angles)* to glow with the beamed signal-emissions. — But that is taking us away from the black-body problem, to which we should now return.

14.2 *Bad noise? — good noise? — and signal control*

Until this chapter, we had assumed that the radiation emitted thermally within the body is *bad*. After all it looks as though it would tend to corrupt any attempted IR signalling — and that, at best, it would simply be a *less*-harmful byproduct from the extinction of unwanted stray signals, as

[142] In this special sense then, "blackness" has a rather odd meaning — black as far as incoming radiation is concerned, but also a sort of mild "white-hot" in regard to emission). Accordingly Planck speaks of both *"black radiation"* and *"white radiation"*, while Rayleigh (1900; paper 260) refers to *"complete radiation"* .

[143] These diameters *include* the myelin. It is well documented that axons smaller than about *"1 metre"* (i.e. *one micron*) are unlikely to myelinated. But why? This is discussed briefly in *Book A*, and in more detail elsewhere.

 — ISBN 0 9577737 1 4

discussed above in section 12.4. But can we be so certain of this badness? That rather depends on our assumptions about three things:

- Badness itself;
- The detailed nature of this supposedly random thermal radiation, and;
- The detailed nature of the signals which are supposedly disrupted.

"Bad-for-me" will usually refer to something which hinders my survival in some sense. Anything which disrupts my internal signalling-system would seem to be destructive or *bad* in that way. But even if it is, I might nevertheless find a way to harness it by some non-obvious strategy — through lateral thinking or through lengthy evolutionary experience. So we might well ask "Would that make it *good also*, despite its badness?"

If the postulated IR signalling-system does really exist, could it make use of such a strategy? Its energy must certainly come from somewhere; but this energy would be needed for two different purposes, and logically these two purposes could be supplied separately. First there is the "brute-force" energy *underlying the IR wave-or-quantum itself*, regardless of its distinguishing properties like direction, phase, and polarization. Is it essential that this energy should come from any special source? — And if we are not unduly choosey, could it be that the naturally-occurring thermal radiation could serve this purpose?

Secondly of course, something extra would be needed to impose the required patterns of polarization or phase etc. onto such quantum emissions.

This could occur (i) as an integral part of the IR photon-*creation* along with the main energy, and that is what I had tacitly implied until now — a *single-source* with both power and "steering" (if any) produced simultaneously.

But in principle there are several other possibilities which might also apply — either separately, or in any combination:

(ii) Could it be that some mechanism *modulates* pre-existing energy-flows (thermal or otherwise), imposing the required properties on them belatedly? TV and radio transmission-circuitry does this all the time, but could we realistically expect such finesse at the bio-quantum level — presumably consuming some minimal amount of energy in this steering role? — A *double-source*: power + steering.

(iii) Perhaps there is a mechanism to *select* amongst the pre-existing emissions, choosing only those which are appropriate at that particular time and place. Once again some minimal energy would be needed to exercise that choice; — another double-source strategy.

(iv) The imposed properties could come automatically from some static *geometrical* feature of the source molecule, which would then be functioning essentially like a radar-or-TV aerial imparting the required extra polarization and direction automatically. In that case no extra energy need be expended — unless, of course, the geometry is contrived to be *variable*, and that would seem to take us back to case (ii) again. — So yet another double-source, even if one contributor is static.

This is not the occasion to dwell on the relative merits or failings of the four possibilities and their mixtures or hybrids. The only point which concerns us here is that any one of the three double-source strategies *(ii)-(iv)* could be in operation for at least some IR signalling — and in any of these cases, one of the components could be a "wild" natural force, and in particular: *The main signal-energy* COULD *be coming from the supposedly-random thermal radiation* within the body tissues.

If so, of course, we would have to re-think the "noise" issue. We could no longer assume that all of this internal heat radiation is necessarily disruptive to any IR-signalling system. It still *might* be, especially if case (i) still applies extensively — but then again *it might not*. At this stage it is difficult to say.

This does mean though, that we cannot automatically rule out the postulated IR-signalling system merely because the thermal IR (its supposed rival) is known to reach such-and-such a level of intensity: This intensity could well be no rival at all (at least within the relevant micro-domains), but rather an integral part of the signalling system itself. — Or this could at least be partly true.

With these thoughts in mind, let us now look a bit more deeply at the known-or-likely nature of thermal radiation — including its idealized form as so-called "black-body" radiation:

14.3 The text-book view of the thermal radiation

There is a clearcut formula for this thermal radiation — the energy emission rate from any "ideally black" body at a particular temperature — *where* C_1 *and* C_2 are known constants):

$$W_\lambda = \text{Energy-rate (per } wavelength \text{ unit of spectrum range)} = \frac{C_1}{\lambda^5} \times \frac{1}{e^{C_2/(\lambda.T)} - 1}$$

↑ ↑

Note how the two terms behave.
Each will tend to zero once, ↑ ↑
($\rightarrow$ *0*), ie. *when* λ *is:*··· LARGE, ≈ 0,
ensuring a hump or "bell" shape in between.

 — ISBN 0 9577737 1 4

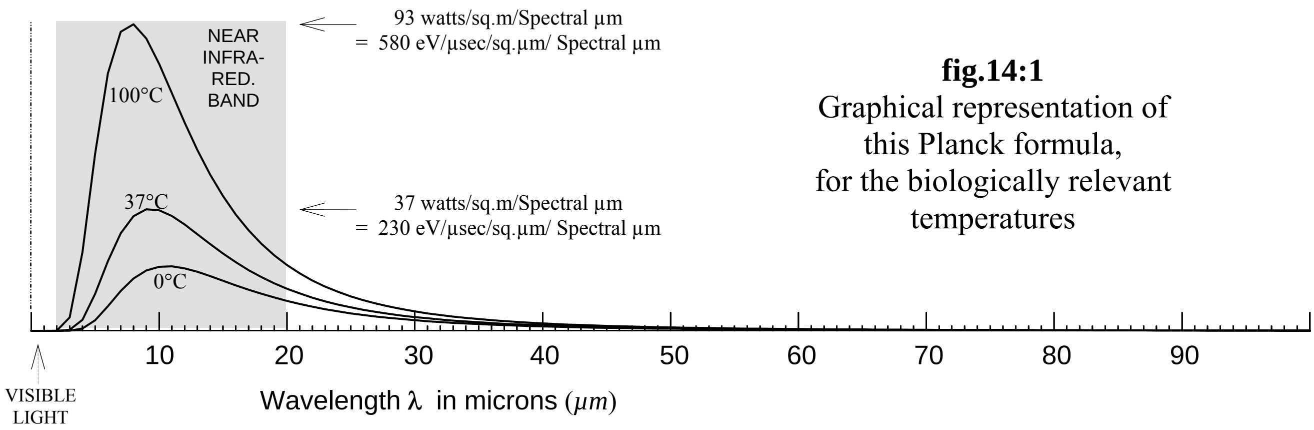

fig.14:1
Graphical representation of this Planck formula, for the biologically relevant temperatures

Put briefly, and accepting any biologically-sensible temperature, this famous Planck-formula[144] tells us that:

1• Nearly all the emissions will be ***within the infra-red*** wavelength range of *2µm* to *50µm*.[145]

2• The distribution-graph between these two wavelengths will be a ***single smooth hump***, peaking at about *10µm* — implying that the underlying matter is fundamentally homogeneous, or all endowed with properties which are randomly distributed on a continuum.

3• As the temperature rises, emission increases for all these wavelengths.

There is little doubt that this relationship is at least broadly true, its detail seems plausible, and the usual textbook account leads us to believe it has sound experimental backing. Moreover the constants C_1 and C_2 can be completely expressed in terms of the standard physical constants[146]: *c*, *h*, and *k* — expressions which offer weighty theoretical prestige when they are justified by theoretical and mathematical argument. In fact we have here an impressive degree of conceptual coherence (with the theory made respectable by mathematics), in line with our earlier discussion in *Part I*.

But this formula is not holy writ. For all its respectability, it is not above investigation; so we should at least have some idea of where it came from, and what assumptions underlie it. It will help therefore if we have a brief look at its history[147] and background:

14.4 Critique of the black-body radiation theory

There are two very seductive aspects to black-body theory. Firstly, we now know that it led to the quantum theory, starting when Planck (1900) suggested the above formula as improving on the Wien (1896) formula — $C_1\lambda^{-5}$*exp*(-C_2/T).dλ — and then set out to offer a rationale for this empirical success. But of course this outcome was totally unforeseen during the 1800s.

The great attraction then and later, was surely that it was a shining example of classical detective-work in research, perhaps starting with:

(1) early work establishing the general nature of heat and its mechanical equivalence — involving names like Count Rumford, Joule, and Prévost. Then:

(2) Kirchhoff (1860)[148] clarified and simplified the issue of what factors affected radiation by narrowing the list down to just two items *when the conditions were ideal in a particular way* — and of course it is this same idealism which we will need to question later on. In his §11 *ff.* he argued that "**I**", the intensity of emitted radiation, depended

[144] Planck's formula of 1900 in one of its various forms (Planck, 1900, 1914/1959; Einstein, 1917; Slater, 1955; Hecht, 1990; etc.).

[145] Beyond that range the emissions will be negligible; with one or other of the two terms being close to zero, as we have just seen.

[146] Thus: $C_1 = 2\pi hc^2$ (= 4.9921×10^{-24} Joule.metre)
$C_2 = hc/k$ (= 1.43879×10^{-2} metre.(°K)),
where
h = 6.6256×10^{-34} Joule.sec/cycle* = Planck's constant
k = 1.38063×10^{-23} Joule/°K/particle*
= Boltzmann's constant
c = 2.99792458×10^{8} metre/sec = speed of light
Tennent (1971), Hecht (1990), and other sources

(ν cycles*/sec = frequency of the light
λ metres/cycle* =its wavelength; c = λ.ν metres/sec)

* Words like "per cycle" and "per particle" are usually omitted in such unit-*expressions*; but that tradition is frankly misleading despite any merit it may have.
Note also that "ν" is the Greek letter *nu*, conventionally used for this role — and it is *not* "v"-for-Victor.

[147] Other accounts can be found in philosophy of science texts like Kuhn (1978); and physics text-books such as Slater (1955) chapters 1 and 3, Halliday and Resnick (1966) chapter 47, or Hercus (1950) — or briefly in paper 260 of Lord Rayleigh's collected works (1964), originally published in 1900, but now with a 1902 footnote added to take account of Planck's correction.

[148] Balfour Stewart also developed similar ideas independently, though his work is less well known.

 — ISBN 0 9577737 1 4

in some unknown way on the wavelength λ, and the temperature **T**. We might express this list and its "unknown way" vagueness by using an algebraic-function statement: $I = f(\lambda, T)$ and of course we can see that this is indeed consistent with the now-known distributions of fig. 14:1.[149]

(3) At about the same time there was an apparently unrelated development which was to play a crucial role some time later. This was the kinetic theory of gases, starting with Clausius (1858/1860) and Maxwell (1860); and its importance lay in its eventual use (about 1868) of Maxwell-Boltzmann statistics which turned out to be applicable to radiation as well.

(4) Another necessary side-issue development was Maxwell's (1865) electromagnetic theory which argued strongly that light (including IR and UV) were *electromagnetic waves, like radio emissions but with much shorter wavelengths*. However this idea did not make much impact until after Maxwell had published a book on the subject in the 1870s, and Herz had demonstrated practical radio transmissions in 1888.

Having reached this stage, the argument was now at a fork in the road. The artificial "black-body/full-radiator" setup offered a great simplification, making the mathematics tractable; but it *was still a simplification* which put it a step or two away from most real-life situations. Of course the researchers of that time were right to take this path, but the danger then was that success might obscure the limiting assumptions. Meanwhile this path did lead to the discovery of the quantum theory, perhaps the most important technological advance of the 1900s — so we shouldn't complain too loudly!

In those years, the purity, tidiness and practical success of black-body theory deservedly achieved victor status. But victory can be just partial or ephemeral, and with hindsight we can now afford to look instead at any "silver-medallists" with *other less-tidy* stories to tell — accounts which might be more applicable in other conditions. Let us then forgo the main victory parade, or diplomatically relegate it to cursory footnotes.[150] That frees us to take a brief look down the relatively untidy track which seems more in line with biological systems:

The total absorption device used in practical black-body studies was a cavity in a block of metal at uniform temperature, with only a small hole to allow emissions and absorptions to pass through. The idea was that any radiation *entering* the cavity would either be absorbed when it first hit a wall, or else it would keep reflecting indefinitely within the cavity until it did become absorbed. (Virtually no escape, and hence "perfectly black"). Meanwhile all the walls will have been *emitting* radiation according to their own temperature and other characteristics. This population of radiation would be randomized or "shaken up" within the cavity, and a small representative sample would escape through the hole as "black-body radiation".

Now this works fine mathematically; but even though we can now legitimately enlist quantum theory to help us, the detailed micro-physics is far from clear when we try to get behind the statistics. What exactly happens to the energy when a photon is absorbed? How it the frequency determined for any particular emitted photon? Presumably it is determined either by a quantum-jump or by the direct oscillation of a charge — implying the presence of orbitals or oscillators. If so, how is it that their emitted frequencies fit so neatly into the Planck distribution curve?

Fortunately we do not need here to go into such matters for this extreme full-radiator case. Even if such ideal-black bodies exist biologically, we can simply rely on the Planck formula and leave it at that. But we may well need to ask similar questions about other matter-based emitters and receivers.

Breaking away from the "full radiator ideal"

Once we get away from that great mixing-bowl homogenizer of the "black"-cavity, we must expect some usable irregularities in the spectrum — not the single featureless artificial hump. Each different feature is likely to have an identifiable cause, and each such identification gives us some opportunity for directed communication.

Kirchhoff himself was well aware of spectral lines which did not fit the black-body paradigm[151] — indeed the line-

[149] Note the poor consistency of notation. Over time and according to whim, the left-hand side of such equations has seen several different letters used to express *the intensity-or-whatever*. (Thus here Kirchhoff uses "I", while Kuhn (1978) renders it as "K_λ" , and "W_λ" was used in the Planck formula as given above.)

[150] **(5)** As one part of the problem, Stefan (1879) found experimentally that the area under Kirchhoff's graph was proportional to T^4 — and Boltzmann (1884) explained it theoretically.
(6) Likewise Wien (1893 Feb) found that the wavelength of the graph's maximum was proportional to T; and then accounted for it himself (Wien, 1894 Apr)
(7) As for the overall formula, its algebra was becoming a bit more explicit: $C_1\lambda^{-\alpha}exp(-C_2/T)$, where α is an unknown constant (Paschen, 1893). This was found experimentally to be ≈*5* — and then confirmed as *5* by Wien (1896), giving us $C_1\lambda^{-5}exp(-C_2/T)$ which was nearly right, but not quite for longer waves, see the graph in Halliday and Resnick (1966, page 1177).
(8) Planck(1900 Oct) offered his improvement, as an empirical curve-fitting task.
(9) Rayleigh (1905, May, Jul) and Jeans (1905 Jul) argued from first principles. Their version accounted for the long-wave end of the spectrum, but failed completely for short waves.
(10) Quantum limitations were postulated to account for this discrepancy.

[151] In §13-14 Kirchhoff (1860) asserts that *"I is a continuous function of the wavelength"*. But then, as if contradicting himself, he also discusses some interesting exceptions "independent of temperature" — including Fraunhofer lines, and spectral lines from metals such as sodium and lithium. (Of course we *now* know these as the spectral lines characteristic of various electron orbital transitions at the atomic level — in line with Bohr's later model of the atom.)
So, although he does not quite say so explicitly, he is effectively suggesting that there are two mechanisms operating concurrently.

 — ISBN 0 9577737 1 4

spectrum was a major interest of his, while the black-body distribution was evidently a side-issue; (Rosenfeld, 1973; Asimov, 1966). Moreover he was aware of some other assumptions and exclusions:

"…when a body is sufficiently charged with electricity, or when it is phosphorescent or fluorescent. Such cases are, however, here excluded." (Introduction).

"…that they emit no polarized rays whatever" — §4.

"…the effect of magnetic force must be excluded" — footnote to §8.

These exceptions could well serve as the *signal* component — and it is against them that the *noise* of black-body radiation would need to be compared. If so, they really deserve a book of their own and I will not attempt to pursue their details further here. Suffice it to say that these "irregularities" are often quite well understood; and they seem close to what was needed for macromolecular communications, as discussed earlier in this book.

However there still remain some intermediate cases — ordinary "hot-body" radiation which could be disruptive even though it has no claim to the 100% efficiency sought in the classical studies. Halliday and Resnick (1966, page 1174) show a typical graph of the emission-distribution for the outside of a full radiator compared with its idealized inside. In this case the "outside" distribution is the same *shape*, but its magnitude is only about 25% of the ideal; and in general the figures can range from about 2% to 98%. Within the body we can hardly escape water and its heavy absorption (and corresponding emission-noise) for several IR wavelength bands — so such scaled-down approximations to the Planck graph will need to be considered at least.

Could the residual noise still be too much for meaningful signalling to survive? Perhaps it is too early to say for certain. At any rate it would seem to be premature to abandon the IR-communication model due to any despair over likely noise. Apart from anything else, even noise can actually be helpful if it is handled appropriately; as we shall now see:

14.5 Tapping the monster's strength by steering it?

Getting back to the Planck formulation itself, let us remind ourselves of the situation. If we take the equation on its own, as depicted above and as usually given in the basic textbooks, we will miss out on some further relevant information which the mere equation does not make explicit:

- In its usual form, this formula tacitly assumes that the radiation is ***randomly propagated in all directions, with random polarization***. However we *can* be more selective and reduce C_1 accordingly — typically dividing by 4π or 2, or both. These distinctions are expressed (if at all) by different names for the symbol on the left of the equation.

- "***Black-body***" implies that the full quota of supposedly-available radiation does actually emerge unhindered through the surface of the body without any reflection; i.e. with an emissivity of unity ($\varepsilon = 1$). In practice though, this ideal can never be exactly achieved, and in fact the value of ε can be quite small[152].

These two factors erode the equation's claim to universality. In short, we might say that black-body theory is too shaky for it to be taken very seriously if it is used as a criticism of any other theory.

However if we concentrate on its well-founded parts, it does have something positive to tell us. It is surely a good guide as to *approximately* where in the spectrum the main signal-energy-flows are likely to be found, even if it cannot forecast the individual anomalous "high spots" caused by emission bands. In other words we have a coarse-resolution guide suggesting where the main action will be, but unable to account for the fine detail. It might then be open to some other influence to impose the detail — to masterfully "switch the railway points" to direct a train which it cannot otherwise control.

In this sense then, it really might not greatly matter *how much* energy is flowing nor in what broad wavelength categories.[153] What would matter, from a signalling viewpoint, would be the "steering, shaping, or modulation" of the emission to determine its *polarization, direction, fine-tuning,* or *collective patterning, etc.,* — those effortlessly imposed signal characteristics as discussed above in chapters 12 and 13 (and applied comparatively effortlessly to each emission at the submolecular level).

Moreover this "steering, shaping, and modulation" would mostly arise from the *irregular* peculiarities of molecular dynamics and geometry — irregularities in the very substrate that the black-body formulation was originally based on.

14.6 The good aspects of thermal radiation.

Seen in the way just considered, the "black body" criticism of the IR-signalling might just evaporate. Indeed we might plausibly see the thermal radiation as no rival but rather as an integral part of the process — almost as the provider of the "carrier wave" of radio theory, a medium upon which the signal-carrying "modulation" is then imposed.

However it will now still be of interest to pursue the quantitative aspects of broadband thermal emission — though arguably this will now be of more interest to the study of signal traffic rather than of the random noise which might disrupt it.

---

[152] Within normal temperatures, ε ranges from *0.98* ("Parsons black" paint) down to *0.02* (Polished silver at 27°C). (Thomas, 1980, p.669).

[153] Providing, of course, it does not reach destructive levels; and that would scarcely apply to the bio-thermal emissions in question here.

 — ISBN 0 9577737 1 4

Not that we have disposed of random noise altogether. No doubt it will still be significant, and indeed in sections 12.3 and 12.4 it was argued that such random noise served the useful purpose of scrambling unwanted disruptive signals. But we are now in the position of not being sure just how much of the thermal emission will be genuine random noise, and how much will have been steered into a signal mode. Moreover if there is a flexibility in the dividing line here, evolutionary forces might well have adjusted it to some optimal position — building up the "steering" mechanisms to just the right level for the signalling task in hand, and allowing whatever level of random noise which might best suit the system.

Finally we should note that Moss and Wiesenfeld (1995) have argued that a certain level of noise is *actually beneficial* in electronic signal processing — offering a useful random element which aids sensitivity by helping certain detection activities to cross their threshold levels. We would clearly not want a signalling system to be flooded with excessive noise, but the mere existence of noise in the system is not necessarily a disaster. The question is *how much should there be*, and *how much is there actually?* At the moment it is difficult to say clearly, and it would be well to investigate the matter further; but it would seem that the Planck formula will not tell us.

So what *do* such straightforward formulae tell us? They effectively say *how much average chatter there will be in the room under given global circumstances*. But they will tell us nothing about the *content* of that chatter, nor *who* the chief participant will be, nor what *role* they might adopt. If these activities are meaningful and constructive then this part of the "black-body radiation" may be seen as *good* communicationally, and not actually random at all.

Outside observers may miss any such subtle details. To them the activity will doubtless appear to be random, and the conversation will seem to be the mere babble of brainless nonentities.

14.7 Anyway, how much radiation can we expect, and at what wavelengths?

The now-conventional Planck formula on *page 70* tells us how to calculate predictions of energy-flow for any given temperature, at least for ideal average conditions; and some of the derived graphs are shown also. However these forecasts can be expressed in other ways which might suit us better. For instance, the algebra changes significantly if we want to put *frequency* or *a logarithmic scale* (rather than simple wavelength) along the *x*-axis. Moreover we might also like to play around with the units until we get the information in a form which makes biological sense at cell level. (Physicists will talk about *watts/sq.metre*, whereas we might find *eV/microsec/sq.micron* more appropriate).

However I offer Table 14 which features a *photon-**count*** instead of the *corresponding* energy flow —
(applying: *Energy/photon* = $h\nu$):

Of course it is mainly the shaded parts of this table which concern us here — the wavelength-range from about *1 to 20 microns*, and a human-body temperature of about *37°C*, but the other entries help to set the context.

The *37°C* column does show up an interesting dichotomy. Within the near-IR range (of about *1 to 4 microns*), the photon emission will be negligible as long as we are depending on traditional black-body mechanisms alone. In that case, any significant traffic in IR signals will have to arise from activities independent of the traditional black-body theory. Instead they will have to come from direct chemical quantum jumps or suchlike.

The situation is different for IR wavelengths beyond about *5 microns*: For most of the separate *1-micron wavelength-intervals* within this range range, the thermal photon-flow will be more than 1000 photons per microsecond for each square micron aperture. That seems likely to be a traffic to be reckoned with — either as a "bad" obstructive noise, or else as a significant source of "potentially-good" signal-vehicles which the nervous system needs to have some way of steering or gating. — Or indeed this radiation might fill both roles to some degree: good *and* bad!

14.8 Conclusion

In physics we can often afford the luxury of assuming that a radiation system has a single set of laws governing absorption and emission for any given set of circumstances. Thus we might *exclusively* invoke: — reflection/refraction — or Bohr orbital-jumps — or black-body — or photo-chemical — or special conduit for existing captured heat-energy.

However biological networks are often far from homogeneous, so we must expect to find *any or all of these mechanisms operating together within the same system*. At the time of writing, It is difficult to assess the relative bio-importance of these five-or-more ways of coping with radiation. I suspect that they all have important roles to play, but it is perhaps too early to identify all these roles with any certainty.

This issue arose from a concern that black-body radiation might swamp any other infra-red activity, and thus make the postulated second neuro-signal mode unworkable. Such fatal disruption now seems at least questionable for several reasons — chiefly (1) the likely presence of reflective surfaces which may isolate subsystems from unwelcome interference; (2) the likely wide-spread use of "tuned" transmissions which would tend to contradict the random-emission assumptions behind black-body theory.

Nevertheless black-body radiation is still to be taken seriously, especially considering that water is effectively black within many bands of the infra-red range. Hence the importance of asking about the extreme scenario: *Just how much radiation would black-body theory predict for a perfectly black emitter at body temperature?* The answer offered is in table 14, and it seems to say that (3) there would actually be negligible black emissions within much of the *near* infra-red range (NIR).

Table 14: *PHOTON* emission-rate:

	Log.λ	λ (µm)	0°C	37°C body	100°C	500°C	1000°C red-hot	6000°C Sun	Step-size along the spectrum, (µm) (weighting factor)
	-1	.1	0	0	0	0	0	2060	.1
	-.9	.13	0	0	0	0	0	91755	.025
UV	-.8	.16	0	0	0	0	0	1548753	.032
	-.7	.2	0	0	0	0	0	12095870	.041
	-.6	.25	0	0	0	0	0	51223360	.051
	-.5	.32	0	0	0	0	0	133455600	.065
Violet	-.4	.4	0	0	0	0	0	236732700	.081
Bluegreen	-.3	.5	0	0	0	0	5	310453000	.103
Red	-.2	.63	0	0	0	0	198‡	322022100	.129
	-.1	.79	0	0	0	0	3134	279171700	.163
	0	1	0	0	0	16	23276	211389900	.205
	.1	1.26	0	0	0	285	94715	144668600	.258
	.2	1.58	0	0	0	2374	239085	91826780	.325
	.3	2	0	0	0	10577	413653	55106890	.41
	.4	2.51	0	0	10	28687	531998	31711400	.516
	.5	3.16	1	8	95	52529	543603	17680780	.65
	.6	3.98	13	65	466	70619	465932	9624600	.818
	.7	5.01	81	285	1361	74666	349806	5144225	1.03
	.8	6.31	281	762	2642	65678	237879	2711150	1.297
	.9	7.94	625	1380	3717	50277	150278	1413413	1.633
IR	1	10	976	1839	4071	34689	89866	730670	2.056
	1.1	12.59	1160	1930	3678	22152	51577	375246	2.589
	1.2	15.85	1116	1689	2873	13354	28699	191723	3.259
	1.3	19.95	913	1288	2012	7711	15599	97561	4.103
	1.4	25.12	662	886	1299	4310	8328	49488	5.166
	1.5	31.62	439	565	790	2351	4385	25040	6.503
	1.6	39.81	272	340	459	1258	2284	12645	8.187
	1.7	50.12	160	196	258	664	1180	6375	10.308
	1.8	63.1	91	109	141	346	606	3210	12.977
	1.9	79.43	50	60	76	179	309	1615	16.337
	2	100	27	32	40	92	157	812	20.567
	Weighted TableTotals	→	26601	39271	69026	624230	2796727	335205200	← Total photon-count /microsec/sq.µm

Body of table shows radiation in: *Photon-count/microsec/sq.micron/Spectral-micron* — rounded to nearest integer, or 7 significant figures.

N.B. Emission rate is given *per micron* of the spectral scale — *not* per spectral-metre, and *not* per (variable) log-interval shown here, as listed in the right-hand column. (This varying step-interval hardly matters unless we are trying to tally all photons across all wavelength-intervals, in which case we have to use the last column's figures as weighting factors).

λ = wavelength in *microns* **(µm).**

‡ This figure "*198*" depicts the incipient *redness* of the red-hot object. Compare it with the other columns.

Such emissions could begin to cause interference to any signal-wavelengths longer than about 4µm; but then such signals would be less likely to "fit" into the more available fibre-optic channels anyhow, so the issue would perhaps not arise. Moreover there would be a further obstacle for such wavelengths in the "blackness" of water.

Then again, if any of these combinations turn out to be on the borderline of possibility, that need not necessarily surprise us. Life itself is full of dangerous-looking balancing acts between the possible and the impossible. Too unstable and you fly to pieces. But going too far the other way and overdoing the stability, may ultimately turn you into an inert pillar of salt — or a lifeless golden statue.

REFERENCES

Alberts, B., D.Bray, J.Lewis, M.Raff, K.Roberts, and J.D.Watson (1983). *Molecular biology of the cell*. Garland: New York.

Ashby, W.Ross. (1952/1960). *Design for a Brain*. Chapman & Hall: London.

Ashby, W.Ross. (1956/1964). *An Introduction to Cybernetics*. Methuen: London.

Asimov, I. (1963). *The Genetic Code*. Signet: New York.

Asimov, I. (1964/1975). *Biographical Encyclopedia of Science and Technology*. Pan: London.

Ayer, A.J. (1959) (Editor). *Logical Positivism*. Free Press: Glencoe, Illinois.

Ayer, A.J. (1978). "On the cleansing of the metaphysical stables". *The Listener*, **99**(2549), 266-270. [Originally broadcast under the title "Logical Positivism and its Legacy", BBC-2 (23/2/78)].

Babich, F.R., A.L.Jacobson, S.Bubash, and Ann Jacobson (1965 Aug). "Transfer of response to naive rats by injection of ribonucleic acid extracted from trained rats". *Science*, **149**, 656-657.

Babich, F.R., A.L.Jacobson, and S.Bubash (1965 Nov). "Cross-species transfer of learning: effect of ribonucleic acid from hamsters on rat behavior". *Proceedings of the National Academy of Science*. **54**(5), 1299-1302.

Beer, A.Stafford. (1972/1981). *Brain of the Firm*. Wiley: Chichester, W.Sussex.

Beer, A.Stafford. (1974). *Designing Freedom*. (Massey Lectures, 13th series). Wiley: Chichester, W.Sussex.

Beller, Mara. (1992). "The birth of Bohr's complementarity: the context and the dialogue". *Studies in History and Philosophy of Science*, **23**(2), 147-180.

Bergmann, G. (1954). *The Metaphysics of Logical Positivism*. Longmans, Green & Co.: New York.

Bloom, F.E. (Editor) (1994). *Neuroscience: from the molecular to the cognitive*. Elsevier: Amsterdam.

Blum, M. (1962). "properties of a neuron with many inputs". In *Principles of Self-organization* (Eds. H. von Foerster and G.W.Zopf), Pergamon Press, Oxford, pp. 95-120.

Böhler, C., P.E.Nielsen, and L.E.Orgel. (1995). "Template switching between PNA and RNA oligonucleotides". *Nature*, **376**, 578-581.

Boltzmann, L. (1884 May). "Ableitung des Stefan'schen Gesetzes, betreffend die Abhängigkeit der Wärmstrahlung von der Temperatur aus der electromagnetischen Lichttheorie". *Wiedemanns Annalen der Physik und Chemie*.[†] **22**(6), 291-294.

Born, M. and E.Wolf (1970: 4th.Edn). *Principles of Optics: Electromagnetic theory of propagation, interference and diffraction of light*. Pergamon Press: Oxford. [sec. 2.4.2, and p.71].

Carnap, R. (1930/1959). "The Elimination of Metaphysics through Logical Analysis of Language"; in A.J.Ayer (Ed.) *Logical Positivism*. Free Press: Glencoe, Illinois.

Cavalli-Sforza, L.L., and M.W.Feldman (1981). *Cultural Transmission and Evolution: a quantitative approach*. Princeton University Press: Princeton, NJ.

Chalmers, D.J. (1995 Dec). "The puzzle of conscious experience". *Scientific American*, **273**(6), 62-68.

Chergui, M. (ed.) (1996). *Femtochemistry: Ultrafast chemical and physical processes in molecular systems.* World Scientific: Singapore.

Clausius, R. (1860). *Philosophical Magazine (series 4)*, **19**, 434-436. [Translated from (1858) "Ueber die mittlere Länge der Wege, welche bei der Molecularbewegung gasförmiger Körper von den einzelnen Molecülen zurückgelegt werden; nebst einigen anderen Bemerkungen über die mechanische Wärmetheorie". *Poggendorffs Annalen*,[†] **105**, 239-258.]

Clausius, R. (1865). "Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie". *Annalen der Physik*, **125**, 353-400.

Cope, F.W. (1973). "Electron-phonon (trapped photon) coupling and infrared coaxial transmission line theory of energy transport in mitochondria and nerve". *Bull. Math. Biol*. **35**, 627-644.

Crick, F., and C.Koch (1995 Dec). "Why neuroscience may be able to explain consciousness". *Scientific American*, **273**(6), 66-67.

Crowther, J.G., and R.Whiddington (1947). *Science at War*. HMSO: London.

Darwin, C. (1859/1968). *The Origin of Species*.

Daub, E.A. (1971). "Clausius, Rudolf". In Gillispie (chief editor) (1970-1976) *Dictionary of Scientific Biography*.

Davies, P.C.W., and J.R.Brown (Editors) (1986). *The Ghost in the Atom: a discussion of the mysteries of quantum physics*. Cambridge University Press.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

[†] Alias *Annalen der Physik*, or (before 1900) *Annalen der Physik und Chemie*. Poggendorff was editor from 1824 to 1877, Wiedemann 1877 to 1899, and Drude 1900 to 1906.

Dawkins, R. (1982/1983). *The Extended Phenotype*. Oxford University Press.
Dawkins, R. (1986). *The Blind Watchmaker*. Longman: Harlow, Essex.
Decca. (1976). *The Decca Navigator System: Principles and Performance*. The Decca Navigator Co.Ltd: New Malden, Surrey; [since taken-over by Racal].
De Morgan, A. (1872/1915/1954). *A Budget of Paradoxes*. Dover: New York. [Edited posthumously by Sophia De Morgan (1st.ed.), and D.E.Smith (1915)]
Descartes, R. (c1630/1973). *Le Monde et le traité de l'homme*. — Also in *Œuvres philosophiques; **I** (1618-1637)*. Garnier Frères: Paris.
Descartes, R. (1644/1973). *Principia Philosophiae* — as "les principes de la philosophie" in *Œuvres philosophiques; **III** (1643-1650)*. Garnier Frères: Paris.
Descartes, R. (1649/1989). *The Passions of the Soul*. Hackett: Indianapolis.
Dhombres, J. (1988). "Science and anti-science: an old story". *Impact of science in society*, **38**(2/151), 209-217.
Dixon, N.F. (1976). *On the Psychology of Military Incompetence*. Futura Publications, Macdonald & Co: London.
Donaldson, H.H., and G.W.Hoke (1905). "The areas of the axis cylinder and medullary sheath as seen in cross sections of the spinal nerves of vertebrates". *J.Comp.Neurol*. **15**, 1.
Duppen K., E.T.J.Nibbering, and D.AWiersma (1994).
"Solvent dynamics probed by photon echo". In D.A.Wiersma, (*loc.cit.*); pp.197-208.
Dutton, J.M., and W.H.Starbuck (1971). "Computer simulation models of human behavior: a history of an intellectual technology". *IEEE Transactions on Systems, Man, and Cybernetics*. **SMC-1**(2), 128-171.
Eccles, J.C. (1986). "Do mental events cause neural events analogously to the probability fields of quantum mechanics". *Proc. Roy. Soc. Lond.*,B **227**, 411-428.
Eccles, J.C. (1989). "A unitary hypothesis of mind-brain interaction in the cerebral cortex". *Proc. Roy. Soc. Lond.*,B, **240**, 433-451.
Edelman, G.M. (1987). *Neural Darwinism*. Basic Books**:** NY.
Einstein, A. (1917). "Zur Quantentheorie der Strahlung". *Physikalische Zeitschrift*, **18**(6), 121-128. .
Eysenck, H.J. (1965). *Fact and Fiction in Psychology*. Penguin: Harmondsworth, Middx.
Faye, J. and H.J.Folse (eds.) (1994). *Niels Bohr and Contemporary Philosophy*. Kluwer: Dordrecht.
Fenichel, O. (1946/1971). *The Psychoanalytical Theory of Neurosis*. Routledge & Kegan Paul: London.
Fjerdingstad, E.F., Th.Nissen, and H.H.Roigeerd Petersen (1965). *Scand. J. Physiology*, **6**, 1-
Forman, P. (1971). "Weimar culture, causality, and quantum theory, 1918-1927: adaptation by German physicists and mathematicians to a hostile intellectual environment". In R.McCormmach (ed.) *Historical Studies in the Physical Sciences,* **3**. University of Pennsylvania Press: Philadelphia.
Friede, R.L., and R.Bischhausen. (1982). "How are sheath dimensions affected by axon caliber and internode length?". *Brain Research*, **235**, 335-350.
Furth, H.J. (1969). *Piaget and Knowledge*. Prentice Hall: Eaglewood Cliffs, NJ.
Furth, H.J. (1973). *Deafness and Learning: a psychosocial approach*. Wadsworth: Belmont, California.
Gaito, J. (1963). "DNA and RNA as memory molecules". *Psychological Review*. 70(5), 471-480.
Glassman, E. and J.E.Wilson. (1969) *Pulse labelling of RNA during a short learning experience*, Paper presented at the annual conference of the British Psychological Society, Edinburgh.
Grishanin, B.A., A.Yu.Chikishev, N.I.Koroteev, V.D.Vachev, and V.N.Zadkov (1994). "Fast (Pico- and Femtosecond) reaction dynamics in the excited states of large molecules: fluorescence studies and computer simulations". In D.A.Wiersma, (*loc.cit.*); pp.169-195.
Gross, C.G and F.M. Carey, (1965). "Transfer of learned response by RNA injection: failure of attempts to replicate". *Science*, **150**, 1749.
Hall, E.E. (1902). "The penetration of totally reflected light into the rarer medium" *Physical Review*. **15**, 73-106.
Halliday, D., and R.Resnick. (1966). *Physics.* Wiley: New York.
Hanfling, O. (1981). *Logical Positivism*. Basil Blackwell: Oxford.
Heaviside, O. (1885, Jan 10). "Electromagnetic induction and its propagation: II. On the transmission of energy through wires by the electric current" *The Electrician*, 178- [Republished in his *Electrical Papers* (1892/1970, *q.v.*), **1**, 434-441.]
Heaviside, O. (1892/1970). *Electrical Papers* [2 vols]. Chelsea Publishing Co.: Bronx, New York.
Hebb, D.O. (1949/1964). The Organization of Behaviour. Wiley: New York and London.
Hecht, E. (1990:2ndEdn / (1974 orig.)). *Optics.* Addison-Wesley: Reading, Massachusetts.
Hercus, E.O. (1950). *Elements of thermodynamics and statistical mechanics*. Melbourne University Press.
Hesse, M.B. (1961). *Forces and Fields: the concept of Action at a Distance in the history of physics*. Thomas Nelson & Sons: London.
Hesse, M.B. (1980). *Revolutions and Reconstructions in the Philosophy of Science*. Harvester Press: Brighton, Sussex.

Honner, J. (1994). “Description and deconstruction: Niels Bohr and modern philosophy”; in Faye and Folse (op.cit.), pp 141-153.

Hydén, H. and E.Egyházi (1962). “Nuclear RNA changes of nerve cells during a learning experiment in rats”. *Proceedings of the National Academy of Sciences*. **48**, 1366-1373.

Hydén, H. (1967a). “RNA in brain cells”. In Quarton *et al.* (1967), *q.v.*; pp.248-266, 857-859.

Hydén, H. (1967b). “Biochemical changes accompanying learning”. In Quarton *et al.* (1967), *q.v.*; pp.765-771, 913-914.

Jeans, J.H. (1905 Jul). “On the partition of energy between matter and æther”. *Philosophical Magazine (series 6)*, **10**(55), 91-98.

Jefimenko, O.D. (1966). *Electricity and Magnetism: an introduction to the theory of electric and magnetic fields*. Appleton-Century-Crofts: New York.

Katz, B. (1966). *Nerve, muscle, and synapse*. McGraw-Hill: New York and London.

Kenny, A. (1992). *The Metaphysics of Mind*. Oxford University Press.

Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan: London.

Kirchhoff, G. (1860). “On the relation between the radiating and absorbing powers of different bodies for light and heat”. *Philosophical Magazine (series 4)*, **20**(130), 1-21. [Translated from (1860) *Poggendorffs Annalen*,† **109**, 275-301.]

Klir, G.J. (1969). “On the relation between cybernetics and general systems theory”. In J.Rose, (Ed.), *Progress of Cybernetics*, **vol.1**. pp.155-165. Gordon & Breach: London.

Kohler, B., J.L.Krause, F.Raksi, Ch.Rose-Petruck, R.M.Whitnell, K.R.Wilson, V.V.Yakovlev, and Y.J.Yan (1994). “Femtosecond pulse shaping for molecular control”. In D.A.Wiersma, (*loc.cit.*); pp.209-217.

Krips, H. (1996). “Quantum mechanics and the postmodern in one country”. *Cultural Studies (London).* **10**(1), 78-114.

Kuhn, T.S. (1978). *Black-body theory and the Quantum discontinuity*. Oxford University Press.

Lamarck, J.B.P.A deM. (1809). *Zoological Philosophy*.

Landsberg, P.T. and D.A. Evans. (1970). “Free will in a mechanistic universe?” *British Journal for the Philosophy of Science*, **21**, 343-358.

Lavin, E.P. (1971). *Specular Reflection*. Adam Hilger: London.

Livingston, R.B., K.Pfenninger, H.Moor, and K.Akert. (1973). “Specialized paranodal and interparanodal glial-axonal junctions in the peripheral and central nervous system: a freeze-etching study”. *Brain Research*, **58**, 1-24.

Lyne. J. (1996). “Quantum mechanics, consistency, and the art of rhetoric: response to H. Krips”. *Cultural Studies (London).* **10**(1), 115-132.

Mach, E. (1905/1976). *Knowledge and Error*. Elsevier: Dordrecht.

Maxwell, J.C. (1860 Jan). “Illustrations of the dynamical theory of gases — Part I. On the motions and collision of perfectly elastic spheres”. *Philosophical Magazine (series 4)*, **19**, 19-32.

Maxwell, J.C. (1860 Jul). “Illustrations of the dynamical theory of gases — Part II. On the process of diffusion of two or more kinds of moving particles among one another”. *Philosophical Magazine (series 4)*, **20**(130), 31-33.

Maxwell, J.C. (1865). “A dynamical theory of the elecromagnetic field” *Philosophical Transactions of the Royal Society, London.* **155**, 459-512.

Maxwell, J.C. (1873?/3rd ed.1904). *Treatise on Electricity and Magnetism*. Oxford University Press.

McCollum, G. (2000). “Social barriers to a theoretical neuroscience”. *Trends in Neuroscience*, **23**, 334-336.

Miller, O.L, jr. (1973 Mar). “The visualization of genes in action” *Scientific American*, **228**(3), 34-42 and 128.

Moss, F., and K.Wiesenfeld (1995 Aug). “The benefits of background noise” *Scientific American*, **273**(2), 50-53.

Nahin, P.J. (1987). *Oliver Heaviside, sage in solitude*. IEEE: New York.

Newton, I. (1686/1962). *Principia* [:*Philosophiæ naturalis Principia mathematica*], **vol.1**. University of California Press: Berkeley and London.

Nöckel, J.U., and A.D.Stone (1997) “Ray and wave chaos in asymmetric resonant optical cavities”. *Nature* **385**(6611), 45-47.

Papanicolaou, A.C., and P.A.Y.Gunter (eds). (1987). *Bergson and modern thought: towards a unified science*. Harwood academic publishers: Chur, NY, London, Paris, Melbourne.

Paschen, F. (1893). “Ueber die Emission erhitzter Gase”. *Wiedemanns Annalen*,† **50**, 409-443.

Pask, G. and B.C.E.Scott. (1972). “Learning strategies and individual competence”. *Int. J. Man-Machine Studies.* **4**, 217-253.

Pauling, L. (1935/1960). *The Nature of the Chemical Bond and the structure of molecules and crystals*. Cornell University Press: Ithaca, NY.

Pauling, L., and E.B.Wilson (1935). *Introduction to Quantum Mechanics*. McGraw-Hill.

Penfield, W. (1958/1967). *The Excitable Cortex in Conscious Man*. Liverpool University Press: Liverpool.

Penfield, W. and Roberts, (1959). *Speech and Brain-mechanisms*. Princeton University Press: Princeton, NJ.

Piaget, J. (1949). *Traité de logique*. Armand Collin: Paris.

Piaget, J. (1952). “La logistique axiomatique ou ‘pure’, la logistique operatoire ou psychologique, et les réalités auxelles elles correspondent”. *Methodos*, **4**(13), 72-85.

Piaget, J. (1967/1971). *Biology and Knowledge*. Chicago University Press; and Edinburgh University Press.

Planck, M. (1900 Oct). “Ueber eine Verbesserung der Wien’schen Spectralgleichung”. *Verhandlungen der Deutschen Phys. Gesellschaft, Nr.13* — with *Annalen der Physik*, **3**, 202-210.

Planck, M. (1901 Mar). “Ueber das Gesetz der Energieverteilung im Normalspectrum”. *Annalen der Physik*, **4**, 553-563.

Planck, M. (1906/1914/1988). *Theory of Heat Radiation.* American Institute of Physics; (The History of Modern Physics 1800-1950, volume 11). [Both English and German texts].

Pockley, P. (1988). “Communicating science to the sceptics”. *Impact of science in society*, **38**(2), 219-229.

Popper, K.R. (1934/1959/1968). *The Logic of Scientific Discovery* [*Logik der Forschung*]. Hutchinson: London.

Poynting, J.H. (1884). “On the transfer of energy in the electromagnetic field” *Philosophical Transactions of the Royal Society (London)*, **175**, 343-361.

Pribram, K.H., (1971). *Languages of the Brain: experimental paradoxes and principles in neuropsychology*. Prentice-Hall: Eaglewood Cliffs, NJ.

Pribram, K.H., (1991). *Brain and Perception: holonomy and structure in figural processing*. Lawrence Erlbaum: Hillsdale, NJ.

Pribram, K.H., M.Nuwer, and R.J.Baron. (1974). “The holistic hypothesis of memory structure in brain function and perception”; in Atkinson, R.C., S.H.Krantz, R.C.Luce, and P.Suppes (Editors), *Contemporary Developments in Mathematical Psychology*, W.H.Freeman & Co., San Francisco.

Quarton, G.C., T.Melnechuk, and F.O.Schmitt (Editors) (1967). *The Neurosciences: a study program*. Rockefeller University Press: New York.

Rasmussen, N. (1996). *Lived experience and the electron microscopic world*. Paper delivered at the AAHPSSS conference; HPS Dept., Melbourne University.

Ray, P.S. (1972) “Broadband complex refraction indices of ice and water”. *Applied Optics*. **11**(8), 1836-1844.

Rayleigh (1900). “Remarks upon the law of complete radiation”. *Philosophical Magazine*, **49**, 539-540. [Paper 260]

Rayleigh (1905 May). *Nature*, **72**, 54-55. [Paper 305b in his collected works].

Rayleigh (1905 Jul). *Nature*, **72**, 243-244. [Paper 305c in his collected works].

Rayleigh (1964). *Scientific Papers by (John William Strutt) Lord Rayleigh*. Dover Publications: New York.

Röseberg, U. (1994). “Hidden historicity: the challenge of Bohr’s philosophical thought”; in Faye and Folse (op.cit.), pp 325-343.

Rosenfeld, L. (1973). “Kirchhoff, Gustav Robert”. In *Dictionary of Scientific Biography* (chief editor Gillispie; 1970-1976).

Ryle, G. (1949/1963). *The Concept of Mind*. Penguin Books.

Sayre, Anne (1975). *Rosalind Franklin and DNA.* Norton: New York.

Schrödinger (1928). *Collected Papers on Wave Mechanics.* Blackie & Son: London and Glasgow. [All originally published in 1926].

Shaffer, J. (1967). “Mind-Body Problem” in *Encyclopedia of Philosophy*, (Paul Edwards, editor).

Shannon, C.E., and W.Weaver. (1949). *The Mathematical Theory of Communication*. University of Illinois Press: Urbana, Illinois.

Slater, J.C. (1955). *Modern Physics*. McGraw-Hill: New York.

Sorrell, T.N. (1988). *Interpreting Spectra of Organic Molecules*. University Science Books: Mill Valley, California.

Stefan, J. (1879). *Wiene Berichtungen*. **79**, 391-423.

Szent-Györgyi, A. (1968). *Bioelectronics*. Academic Press: New York and London.

Tennent, R.M. (ed.) (1971). *Science Data Book*. Oliver & Boyd: Edinburgh.

Thagard, P. (1992/1993). *Conceptual Revolutions*. Princeton University Press, Princeton, NJ.

Thom, T. (1985). *Private Pilot Navigation*, **vol.1**. Aviation Theory Centre: Williamstown 3016, Australia / London.

Thomas, L.C. (1980). *Fundamentals of Heat Transfer*.Prentice Hall: Eaglewood Cliffs, New Jersey.

Thomson, W. [Lord Kelvin] (1867). “On vortex atoms”. *Proceedings of the Royal Society of Edinburgh*, **6**, 94-105; reprinted in *Phil.Mag*., **34**, 15-24.

Traill, R.R. (1988). “The case that mammalian intelligence is based on sub-molecular memory coding and fibre-optic capabilities of myelinated nerve axons”. *Speculations in Science and Technology*. **11**(3), 173-181.

Traill, R.R. (1999). *Mind and Micro-Mechanism: a hunt for the missing theory*. Ondwelle Publications: Melbourne. [*“Book A”* in the present series]

van Gelder, T. (1996). *Dynamical versus computational models of cognitive processes*. Paper presented at neuroscience seminar, Melbourne University, 13 March.

Wallace, A. (1996). *X-rays and N-rays: Discovery and expectation in physics at the end of the 19th century*. Paper delivered at the AAHPSSS conference; HPS Dept., Melbourne University.

Watson, J.D. (1968/1991). *The Double Helix*. Scribner; and Mass Market Paperbacks.
Wien, W. (1893 Feb). *Sitzbericht der Akad. Wissenschaft*, **9**, 55-.
Wien, W. (1894 Apr). "Temperatur und Entropie der Strahlung". *Wiedemanns Annalen*,† **52**(5), 132-165.
Wien, W. (1896 Jul). "Ueber dir Energievertheilung im Emissionsspectrum eines schwarzen Körpers". *Wiedemanns Annalen*,† **58**(8), 662-669.
Willshaw, D.J., H.C.Longuet-Higgins, and O.P.Buneman. (1970). *Nature*. **225**, 178.
Whitehead, A.N., and B.Russell. (1910-1913). *Principia Mathematica*. Cambridge University Press.
Wiener, N. (1948/1961). *Cybernetics, or Control and communication in the animal and the machine*. M.I.T. Press: Cambridge, Mass.
Wiersma, D.A. (ed.) (1994). *Femtosecond Reaction Dynamics* North-Holland: Amsterdam.
Zolotarev, V.M., B.A.Mikhailov, L.I.Alperovich, and S.I.Popov. (1969). "Dispersion and absorption of liquid water in the infrared and radio regions of the spectrum". *Optics and Spectroscopy*. **27**, 430-432.

 — ISBN 0 9577737 1 4

INDEX

D

E

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

L

M

N

 — ISBN 0 9577737 1 4

 — ISBN 0 9577737 1 4

Q

R

S

 — ISBN 0 9577737 1 4

T

 — ISBN 0 9577737 1 4

U

V

W

Z

 — ISBN 0 9577737 1 4