Thinking as mental model-building: a Piagetian-cum-mechanistic explanation of the 'engram'

Monograph 12: Brunel University, Institute of Cybernetics, UK.
February 1976
(from a conference paper of April 1975)

Robert R. Traill
email: rrtraill4{at}bigpond▪com *

republished online by: Ondwelle Publications, 29 Charlotte Street, Blackburn South, 3130, Vic., Australia *
April 2007 (a)
and also by: Brunel University Research Archive (BURA).

CONTENTS

The occasional revisions made in this March 2007 edition are either (i) obvious from dates or technology changes etc., notably updating references and links **&/or** (ii) wording within square brackets "[...]", **&/or** (iii) within footnotes marked thus "*text*[B] " using an upper-case letter ("B" in this case). Also (iv) the numbered subsections (see "Contents"↑) have now belatedly acquired subheadings, and (v) there has likewise been some general trivial editing.

* Contact details updated April 2014

Preface to this 2007 online edition

This short conference-paper, plus the much longer account in *Kybernetes* **5** (Traill, 1976), together launched a lengthy project. This project aimed at finding *a believable biophysical basis for "the Mind"* as we know it. That is no mean task, so it is perhaps not surprising if early thoughts appeared as rather vague and macro "block diagrams" like the one below here, in subsection 2.

Taken alone, such diagrams are usually a way of saying *"this does-or-should happen **somehow** in this sort of sequence — but let's not bother yet about just what that 'somehow' entails!"* In other words we are looking at observable phenomenon-patterns without attempting any in-depth explanation. There is indeed a worthy place for such an approach by clinicians, who are usually dealing with people face-to-face at this macro level. (At least that is their role in the short term, where digression into deep theory may be a counterproductive distraction). However, if psychology is to progress overall, someone has to grasp this nettle of "the micro" — and any such investigator is likely to be a theoretician with some understanding of micro-processes in other disciplines.

Towards the micro — and rigorous explanation

It perhaps took a biologist-cum-epistemologist like Piaget to take the main first step in this direction, though the very fact that he borrowed ideas from here-and-there suggests earlier contributions to this process by others. For example he borrowed the terms *"assimilation"* and *"accomodation"* from aspects of digestion! — and his crucial *"scheme" concept* was perhaps partly inspired by ethologists like Konrad Lorenz, with their *"Fixed Action Patterns"* (stereotyped output), and *"Innate Releasing Mechnisms"* (keylike input-patterns).

Chemistry arguably reached this intermediate stage with the atom-promoting work of John Dalton (1766-1844) and Dmitri Mendeleev (1834-1907), etc. — but a proper in-depth understanding had to await the modern wave-orbital model of the Bohr atom and its compounds, and even now there are some unresolved issues about the underlying quantum theory.

Likewise the genetics of Gregor Mendel (1822-1884) was eventually recognized in 1900 as giving us a theoretically-structured basis, involving such *then-hypothetical* entities as the gene and hence the functional chromosome — from which any in-depth understanding required the Watson-Crick DNA-model of 1953. This lifted the chromosome's substructure out of being a mere hypothetical abstraction, and into the status of a tangible physical DNA-string which did have the required coding properties.

How exactly do the Micro-Models help us?

How stable and predictable are the elements we are considering in any given project? *Well-defined items* like the pieces of a standard jigsaw puzzle can be fitted together in a finite time, giving a total "obviously correct" final result. *Practical problems increase exponentially the more these items cease to be well-defined*; so if our jigsaw pieces are well-worn, part-melted, and discoloured we may have great difficulty in solving the problem at all, or we may be left with many unconvincing "possible solutions". Such problems are compounded if we overlook or deny the mismatches — as is all too likely in the social sciences where there is a dearth of truly stable micro-elements![A]

A This shaky aspect of theorizing from *inevitably inadequate social-model elements* has been the undoing of many Social Theorists including most revolutionaries. This fuzziness likewise led *Structuralism* into its now largely-discredited role within the social-sciences. (In contrast, one could logically use the word "structuralism" for the abovementioned atom and DNA cases, and then it clearly *would* signify a legitimate methodology — but of course the word now has a bad name!). Even in the social sciences, one may sometimes have no other adequate method available, and that might be seen a license to proceed anyhow — but *with caution and humility!* In fact Freud argued the jigsaw case for his own fuzzy applications, and he did have half-a-point there, but Freud was not known for his humility! — (See the two quotes in Chapter VII of "Monograph 18": www.ondwelle.com/OSM06.pdf).

Whenever the basic elements of a theory-system are robust enough (and their behaviour-patterns are formulated accurately enough), it then becomes possible to make feasible predictions (i) for specific problems, and (ii) on other generalizations which we might like to consider as new knowledge. Thus in 1924, Louis de Broglie[B] was able to speculate successfully that *all* particles might have wave properties — given that *some do*; and of course the success of Newton's speculative theories is well known.

It was by no means clear that psychology was seriously amenable to such treatment. This certainly seemed unlikely at the macro level of everyday adult life (except perhaps statistically); but why start with such a complex problem? Piaget took the much better course of studying very young children,[C] and this did offer evidence which was simple enough to yield some vision of what the basic elements might be. Hence his abstract concept of the "scheme", as a somehow-preserved *element of action*.

Physical nature of the Piagetian "scheme"? — My impressions by 1975 (this paper)

(1) The scheme had to be some physically-embodied code-pattern.

(2) It seemed that the scheme must be predominantly a stringlike one-dimensional (1D) coding — depicting a linear series of subactions, all *very like a computer-subroutine written in machine language* (and similar to other linear codes such as spoken and written speech, though using a totally different code).[D]
Corollary 2a: This seemed to rule out (for this specific role) the *whole-neuron-switch model* of the 1960s.
Corollary 2b: It also put in question *whether synapses could be lined up meaningfully* for this 1D-code purpose. One would at least need to explain how such a system *might work*. And then: Is there actual evidence that this actually occurs, or could occur?

(3) The scheme needed to have clearcut digital code-states — not vague-or-floppy.
Corollary 3: This seems to rule out both synapses and whole-neurons as scheme-bases, as both seem to be essentially floppy, with no discernible step-jumps between sub-states, apart from complete action-potential discharge. (That supports their use in *other* roles as (i) fine-tuning, and (ii) communication with the peripheral nervous system, for muscle-control etc.).

(4) Other *scheme-candidates* included DNA, RNA, and protein — though there might be others of a larger size, if only we could find them.

Conclusion. There must be a *dual system* within the brain: (a) the textbook action-potential system, and (b) the Piagetian "scheme" system for intelligence; — the two systems being comparatively autonomous but working together in cooperation. This was the *working hypothesis*, which needed further analysis.

Trying to fit the new Physical-Scheme into Piaget's framework: — This paper's agenda

Once one steps beyond the vague-and-abstract, there are usually unforeseen logistical details which need to be addressed. This 1975 paper (page 4, *ff.*) set out to sketch some of these issues — notably how the various Piagetian scheme-dynamics might play out *if* scheme-elements were indeed macromolecules such as RNA or protein.

R. R. Traill

Melbourne, 10 March 2007

[B] L. de Broglie, (1924), *Philosophical Magazine*, **47**, 446.

[C] The study of *insect* behaviour is now also well worth considering in this light.

[D] At that time, I took this pro-1D view mainly because Piaget's accounts seemed to imply it. However there are also more cogent explanations for 1D's special merit: e.g. §4.3–§4.4, "Book B" (2000) www.ondwelle.com/BK1_V28.PDF_

Thinking as mental model-building: a Piagetian-cum-mechanistic explanation of the 'engram'

Monograph No. 12, Institute of Cybernetics, Brunel University, UK.

(Paper presented to the Philosophy of Psychology seminar during the Nottingham Conference of the British Psychological Society, 7 April, 1975)

ABSTRACT

Piaget (like Skinner) appears to deny[E] the relevance or possibility of describing thought in mechanistic[E] terms. Nevertheless, this paper attempts to outline one way in which this might be done for Piaget's concepts.

Three domains, or "worlds" are considered (following both Piaget and Popper): •*Reality and the senses*, • *thought proper*, and • a *symbolic domain* (divided into ***a*** internal, and ***b*** external).
Within the second domain are linear codings (pre-set but changeable) which can comprise "schemes" when activated synchronously in sufficient numbers. Non-linear schemes and schemata are explicable in terms of "sub-programming" and "cross-referencing".

Elementary units for schemata may be scheme-elements (or ensembles of them) which have become more or less permanently stabilized due to their self-stabilizing cross-references. These inhabit the symbolic domain ("world 3a").

1. The Language model: valid for society, false for individuals — and it lacks rigour

Dr. Olson yesterday drew our attention to two types of language used in our community:– written language (in the Platonic tradition) which purports to be precise, as against spoken commonsense language (in the Homeric, poetic tradition). The academic ascendancy of precise written form probably reached its peak in Whitehead and Russell's Principia Mathematica; but in 1931 "precise symbolism" was stripped of any claim it might have had to *infallibility* (*in the hands of mere mortals* in the real world).

In that year, Kurt Gödel demonstrated, for a closed system (with no informationally infallible "helpers" outside the system), that it was impossible to *rigorously define* any symbolic language as the logical positivists would have wished. The implication of this was that — like it or not — we ultimately cannot do without the despised commonsense type of language.

So much for language itself; but what about its relation to knowledge? I suggest that *if* language could be rigorously defined, *then* a verbal description would constitute a *precise model* of whatever it described. (It is hard to see how *any* verbal description could be thoroughly exhaustive and precise, — and this lends commonsense weight to Gödel's finding). In fact verbal descriptions always do have limitations. Nevertheless, limitations or not, I suggest that verbal descriptions are, in some sense, *models of knowledge* at a social level; and this is the level discussed by Popper 1972).

{On reflection,[F] I suspect that Popper may be talking about *social* knowledge as if it were essentially the same thing as *private* knowledge; (or maybe that is simply my interpretation). Anyhow, from an *abstract* epistemological viewpoint, I can see no great objection to this:– one bootstrap-learning-system is *probably* functionally equivalent to another. However, when clothed in structural terminology this is likely to cause considerable confusion if the distinction is not realized.. Thus a written record of knowledge (for instance) is *external to the individual* but may be *socially internal.*}

[E] Here *"deny"* is too strong a word for Piaget's position (though not for Skinner's!). In fact, given the overenthusiastic empiricism of the time, Piaget was probably wise to *avoid* such issues, and simply *keep quiet* about them. Yet in his book *Biology and Knowledge* (1967/1971) he *did* drop some tentative hints on possible mechanism, as discussed in Traill (2006b): www.ondwelle.com/OSM02.pdf "*Mechanistic*": see next paper, footnote D.

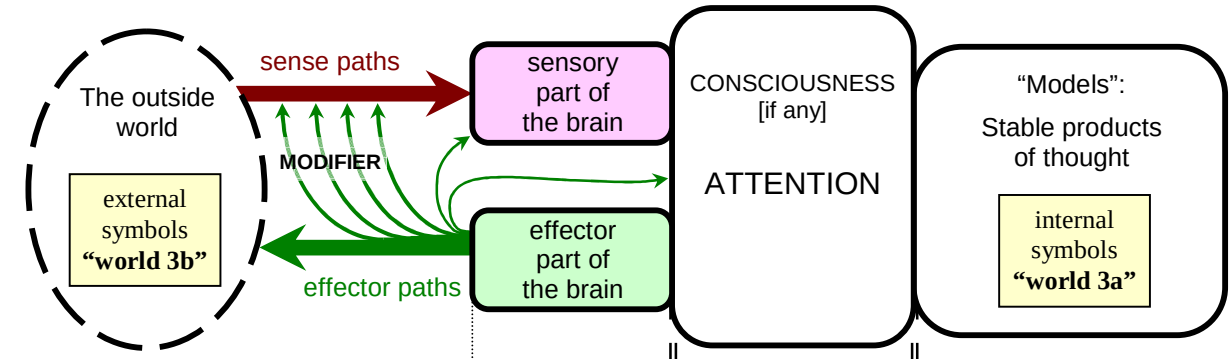
[F] This paragraph was added for the 1976 edition (after the 1975 conference).

2. Self-organized Knowledge within the individual brain

Let us now take a different closed system. Instead of Popper's scientific community, let us follow Piaget and consider an individual person. But is an individual person a closed system in the sense discussed? Well not quite, but I suggest that the defect in closure can be treated as a separate *component* as follows:– Some predisposition and primitive strategies are *inherited*, thus presumably constituting assistance from outside the individual. Nevertheless, such assistance need not be seen as God-given by an intelligent outsider — rather it can be explained in *evolutionary* terms, so it can scarcely be described as infallible or precise. But in any case the extent of such assistance is severely limited.

But surely we as parents or teachers provide the transcendental outside assistance? Well I offer this as a possible debating point for subsequent discussion — at least for the older individual; but I suggest that for the very young infant this is just not so — simply because he has no means of understanding us (or even recognizing our existence as objects[*]), other than those crude hereditary stereotypes which we have just considered.

How then can an individual (or community) build up a model of an environment about which it is so ignorant that it scarcely knows how to begin? Let us look at the situation in diagrammatic form.[G]



Popper:	"WORLD **1**" ["World 1½" ?!]	"WORLD **2**"	"WORLD **3**"
Piaget:		"Thinking", "knowing", "operational"	"symbolic"
Analogy ? (types of memory)	iconic ?	STM ? [i.e. Short Term Memory ?]	LTM ? [i.e. Long Term Memory ?]

What can such an ignorant organism do? It can't make sense of its sensory input if received passively. All it can do is to *add arbitrary activity* (see the effector paths) and try to correlate the ensuing input with what it has just done. Just how this might be done is what I would like to consider next.

---

[*] Dr. G.Butterworth (Southampton), in the subsequent discussion, objected that the infant is less ignorant than one is inclined to suppose. This may be so; but in that case it would seem that we should merely attribute more quantitative importance to the hereditary component (discussed above), without necessarily affecting the argument here. There seems no inherent reason why some of the structures (schemes and schemata) should not have a hereditary (or at least pre-natal) origin.

[G] This diagram builds upon the simple three-block diagram re-introducing Popper's *"three worlds" for mental activity* (1972) — see the original sketch in *Kybernetes*, **5**, page 74 (Traill, 1976); or the 2006 online-edition of it: www.ondwelle.com/Mol-Intel-A.pdf

Block diagrams are rather inadequate representations of reality, and my block diagram is no exception. Even assuming it is correct as far as it goes, it still leaves a lot unsaid concerning dynamics and other detail, and I would now like to go some way towards remedying this deficit. In doing so I will certainly be speculating beyond what the evidence can support using normal scientific criteria (though I have not been unmindful of what is biologically feasible); but I believe this is justifiable on two grounds.

Firstly, Piaget's theories and terminology are often so nebulous as they stand that it is difficult to comprehend them; therefore any reasonable heuristic framework for them is likely to be useful mnemonically, even if it cannot ultimately be accepted as representing the truth. The second justification rests on the very issue we are discussing (though at the Scientific Community level rather than the Individual level). Briefly, the argument is that every invocation of evidence to support a theory has a "common-sense" (non-rigorous) element in its logic, so that there is no absolute criterion; and that under the particular circumstances of this case there is a "common-sense" argument which can be applied with reasonable legitimacy. However, we will be in a better position to discuss this later, if anyone should care to raise it again.

3. The fixed patterns of input and output — "World 1½" between Popper's domains 1 and 2

So now, to elaborations of the model; and for this purpose I will use metaphorical rather than physiological terminology. Suppose then that the box labelled "*effector part of the brain*" contains a large array of *sealed units*. When activated (by a button on a panel within "world 2") this sealed unit will perform a specific invariant elementary task, such as a knee-jerk or eye-blink. That is not to say that there may not be several ways of making an eye blink, but these would involve different sealed units, alone or in combination, sequentially or simultaneously.

The "buttons" would be activated by a transmitted code of some sort. This code could have several components (such as pulse-pattern, phase difference, or general location) but that need not concern us here; we may simply say that the code signal (whatever it is) is selectively received by that "button" which happens to be "labelled" with the corresponding code. Thus some effector is put into execution, either producing a stereotyped *motor* activity, or some *modifier* activity to an input channel, or to the autonomic nervous system, or whatever.

On the sensory side, the mechanism could well be very similar, though in the reverse direction. Indeed there is good evidence (from Hubel and Wiesel) that something very like stereotyped sealed units are in operation; so that it is fair to suggest that there is the equivalent of a display of lights on a panel (similar to the buttons) and that one particular light will come on when (say) a dark boundary is moving across the retina at 5 ± 2 mm/sec. in a certain direction, or the texture grain in the image on the retina (at the fixation point) is diverging at a certain rapid rate — a phenomenon which would normally coincide with an approaching missile. But note that there would be no inherent conscious interpretation for these lights; this would have to be learned — though in some cases there would be unconscious hereditary reflex consequences, like an eyeblink in response to the latter "light" coming on.

When a so-called "light" comes on, it will *emit* a code-signal into the "world 2" domain for the consideration of conscious (or "pre-conscious") thought. In special emergency cases (like the approaching missile) we may suppose that this code-signal is identical to that which activates one of the eye-blink "sealed units", so that a reflex response occurs without delay.

4. The Thought-Domain ("World 2"), and Piagetian "schemes"

So now we may turn to this *thought domain* ("world 2") and discuss *its* contents. For Piaget the main or only functional inhabitant here is what he called the "scheme", though he studiously avoids any suggestion as to what this might correspond to in material terms. Indeed, if I remember correctly, he sometimes talks

about a scheme as *being* an activity, whereas I suggest that we should use the term for whatever it is that *encodes* the activity. Let us say provisionally that a scheme is a list of effector codes, and that when the scheme is read through, by some sort of pick-up head, these codes are then free to "push the relevant buttons" in sequence. This then gives the basis for co-ordinated but *learned* behaviour patterns.

In fact I suspect that a scheme consists of not just one list, but many replications of the same list. I shall not go into my reasons here, but in order to distinguish the scheme as a whole from its supposed replicated elements, I suggest the term "tape" for the latter mini-components.

How then is such a scheme established and used? Taking *use* first: It is easy to envisage that they may be triggered off in some way by a coded signal matching up with a label, in the same way as the output "buttons". It is rather less clear just how this triggering would take effect, but let us say that the code-signal attracts one or more of the appropriate "tapes" to a control-centre (or communication *network*); and that once there, the message on the tape (or tapes) is read out such that the appropriate *buttons — or other tapes —* then become active.

As for the establishment of such schemes, let us suppose that they are easily set up, but with a random coding. In some circumstances they then gain access to the control-network, fortuitously. If the outcome is satisfying, then the corresponding tapes tend to be retained and/or duplicated. — But how? Well one possible explanation is that the satisfaction state (or otherwise) becomes associated to the scheme by classical conditioning (which I shall return to shortly); — and for a fundamental coding like *pain* or *pleasure* it may well be that schemes so labelled are singled out for special treatment (destructive or supportive) on the basis of the extra label or tag so acquired.

Very briefly, classical conditioning could occur thus. Two different stimulus patterns are presented simultaneously or nearly so. Consequently many tapes, with lables appropriate to one or the other, gather around the control-centre (or network). In these circumstances the two types of tape can physically intermingle, and in some cases cross-grafts will occur (like genetic cross-overs). Subsequently then, the label for *one stimulus* (*or* for pain or pleasure) may be associated with the rest of the *other tape*.

5. Coherence as the ultimate-but-fallible test of one's concept-validity

But now let us come back to our original problem of how the individual can develop a model of the structured outside world, without knowing anything much about it in advance, and without God-given help.

The answer seems to lie in an inbuilt homeostatic striving for consistency (or coherence or closure). To the extent that the outside world does have such properties (and solid objects clearly do have *permanence* on the whole, and mathematical *"group"* properties when subjected to rotation), then any procedure which favours sets of schemes which turn out to have collective group-like properties, will have the makings of a model-builder.

The mechanism producing such homeostasis may involve a dynamic "working around" the members of the supposed group, to see if it really does have closure. Maybe sleep has something to do with this. Anyhow such self-closing entities seem to acquire a permanence (or near-permanence) of their own, and it could be this alone which places them in the symbol domain ("world 3"). Maybe *this* is what Long Term Memory is all about.

For issues which have a less clear-cut permanence or group-structure than objects possess, there will be statistically more failures in the sense that the patient's model conforms to only part of reality — and maybe a grossly inappropriate part. On this basis we might well be able to explain many psycho-pathological or deviant behaviours — and it need not be confined to areas traditionally associated with Piaget.

REFERENCES

Bruner, J.S. (1974). *Beyond the Information Given: Studies in the psychology of knowing*. George Allen & Unwin: London.

Bruner, J.S., and B.M.Bruner. (1968). "On voluntary action and its hierarchical structure". *Intern. J. Psychol.*, **3**, 239-255. [Republished in Bruner (1974)].

Bruner, J.S., J.J.Goodnow, and G.A.Austin. (1956/1967). *A Study of Thinking*. Wiley: London and New York.

Furth, H.G. (1969). *Piaget and Knowledge*. Prentice-Hall: New Jersey.

Gödel, K. (1931/1967). "On formally undecidable propositions of 'Principia Mathematica' and related systems". In J. van Heijenoort (1967), *From Frege to Gödel: a source book in mathematical logic, 1879-1931*; Harvard University Press: Cambridge, Massachusetts. Pages 596-616 .

Green, R.T., and V.J.Laxon. (1970). "The conservation of number, mother, water and a fried egg chez l'enfant". *Acta Psychologica*, **32**, 1-30.

Hubel, D.H., and T.N.Wiesel (1961). "Integrative action in the cat's lateral geniculate body". *J. Physiol.. London*, **155**, 385-398.

Hubel, D.H., and T.N.Wiesel (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual center". *J. Physiol., London*, **160**, 106-154.

Olson, D. (1975). "The languages of experience: some comments on human cognition", *Proc. Annual Conf. Br. Psychol. Soc.* (April 1975), Univ. of Nottingham, Page 28 (abstract).

Penfield, W. (1958/1967). *The Excitable Cortex in Conscious Man*. Liverpool University Press: Liverpool.

Piaget, J. (1954/1968) . *The Construction of Reality in the Child*. Routledge & Kegan Paul: London

Piaget, J., and B.Inhelder (1966/1971). *Mental Imagery in the Child*. Routledge & Kegan Paul: London.

Popper, K.R. (1972). *Objective Knowledge: An evolutionary approach*. Oxford University Press.

Traill, R.R. (1976 [*then still "in press"*]). "Acquisition of knowledge without transcendental assistance: an extended Piagetian approach". *Kybernetes*, **5**, 73-82.
— www.ondwelle.com/Mol-Intel-A.pdf (Chapters A1 and A2). # 4/75

