FASHIONS OF MIND

Margaret. A. Boden

For meeting of British Association for the Advancment of Science, Liverpool 1982.)

Cognitive Science Research Paper

Serial no: CSRP 019

The University of Sussex
Cognitive Studies Programme
School of Social Sciences
Falmer
Brighton BN1 9QN

FASHIONS OF MIND

Margaret A. Boden
University of Sussex

(For meeting of
British Association for the Advancement of Science, Liverpool 1982.)

Only Princess Diana's wedding-dress was awaited more impatiently, greeted more enthusiastically, and copied more slavishly than are new ideas in psychology. Psychology is especially prone to changing fashions, because there is no theoretical uniform to fit all figures and please all tastes. That is, there is no generally accepted paradigm defining a psychological science. Not only do we not know the right answers, we are not even sure of the right questions. Individual psychologists may be confident that they are posing the central problems, but others will surely disagree. So if one is told that someone is "a psychologist", one is able to infer very little about that person's professional beliefs, or even interests.

Given this disagreement on what style of theorizing best suits the mind, any new approach is likely to be hailed as the missing paradigm, the link carrying psychology from myth to science. The computational approach – in which minds are compared to computer programs -- is the most recent psychological fashion, and I shall say more about it presently. But it is not the first intellectual style to be welcomed as the saviour of psychology, nor is it the first to be mocked by those preferring different modes.

Distinct psychological fashions have been designed by such theorists as Freud, Pavlov, Piaget, Skinner or Laing. None of these has achieved the enduring status of a Chanel suit. Their popularity has waxed and waned over the years, and varies according to social groupings. Workaday styles in psychology – such as intelligence tests and personality profiles – have been widely adopted for practical purposes, but are seen by many as disguising the true nature of what they are intended to display. And if we move out to the accessory fringe, we find such figures of the psychological demi-monde as Reich of the orgone box, Erhardt of EST, and Janov of the primal scream. No psychologist, it seems, can be so maverick as to lack a body of faithful followers, while none is so authoritative as to persuade all comers of their theoretical infallibility. Psychology is not a unified church.

But, unified or not, church it is. The different styles of psychology resemble religious sects, arousing emotional commitment and antagonism to a degree rarely seen in other branches of scientific enquiry. This is not surprising, for any psychology – whether advertize© as the science of mind, brain, or behaviour – has implications bearing on deep issues concerning self and society. So psychological theories typically arouse not only intellectual disagreement and rejection, but passionate denouncement and scathing ridicule.

As in more theological forms of sectarianism, psychologists to great lengths, or sink to surprising depths, in opposing theoretical fashions they find unattractive. Even in the gent days of the nineteenth century, William James (no less civi figure than his brother Henry) remarked of the new experi statistical psychology that it "could hardly have arisen in a whose natives could be bored. Such Germans as Weber, Fechner, an obviously cannot."

By the 1920's the invective had intensified. Watson's brainc behaviourism -- had conquered the American academies within a few Yet it was sneeringly described (by Tolman) as "mere Muscle Twi and (by McDougall) as "a most misshapen and beggarly dwar description that, in the world of haute couture, would be d indeed. More recently, Chomsky has ridiculed Skinner by sayin according to his views on reinforcement, the best way of encourag artist would be to stand in front of one of his paintings "BEAUTIFUL!!!" at the top of one's voice.

Nor are behaviourists the only ones to be attacked. Pia dismissed by many as a pretentious fraud, while even some admirers (such as Bruner) complain that his central theoretical is mere "confusing imagery" and "surplus baggage". Freud h accused (by Cioffi) of systematic intellectual dishonesty, and F theory is seen as a classic case of the Emperor's new clothes by -- who has mocked it with a suitably italicized account of a journey, during which "the train enters a dark tunnel", the sign "rise as we approach and fall again as we pass", and we "sha pencil" to write a postcard, but -- horror of horrors -- "th drops off".

Perhaps the most scathing dismissal of all -- "When men writ volumes of such stuff, are they not mad, or seek to make others has been applied to the most recent psychological fashion. This "computational" approach, wherein the mind is compared to a c program and described by concepts drawn from artificial intel [Boden, 1977].

Artificial intelligence uses computer programming to stu structure and functioning of knowledge. The programs concerned rigid and inflexible, like those used to calculate tax rebates match courting partners. For work in artificial intellige concentrated on intelligent information-processing abilities, enable a system to cope flexibly with changing and largely situations. The relevant programs specify computations e computers to do such things as: conversing (by teletype) in language; understanding spoken speech; recognizing objects s widely varying positions or lighting conditions; planning comple involving unpredictable conditions; making sensible guesses specific knowledge is not available ... and the like.

It is crucial to realize that "computation" here does no "counting", but any symbolic process of inference, compari association. The symbolism may be numerical (for counting is one of computation), or it may be of some other form (such as visual, or logical). The examples mentioned in the previous pa indicate the wide range of computations required for the manipulation tasks that are the focus of artificial intelligence.

Seen from this viewpoint, the mind is a symbol-manipulating system. It contains many internal representations of aspects of the world (and possible worlds), and a variety of rules for building, changing, comparing, and inferring from them. Psychological questions, accordingly, concern the structure and content of mental representations, and the ways in which they can be generated, augmented, and transformed. Thinking, experience, and motivation – and the myriad differences between individual people that lead to the fascinating human pastime of gossip – are grounded in computational processes.

Cast your mind back, for a moment, to the Middle Ages. The thoughts and actions of the mediaeval people who went in search of unicorns – expecting to find them in the forest, their heads resting in the laps of virgins – were guided by a specific mental representation: the goal of finding a unicorn in those circumstances.

We in the twentieth century can form similar representations, which is to say that we can think about the mediaeval beliefs and ideas about unicorns. But we do not guide our footsteps into the forests accordingly, because in our minds the notion that unicorns exist is represented as false (not to mention our suspicion that the Likelihood of finding a virgin hereabouts is today even smaller than that of finding a forest). An essential precondition for intentional action and voluntary choice – that the goal be believed, rightly or wrongly, to be at least possibly attainable – is thus not satisfied, in your mind or in mine.

We could of course decide to suspend our disbelief in unicorns, or to discount it, so as to venture into the woods for a fanciful picnic – dressed appropriately in tunic and hose and carrying a silken halter. This would require the (temporary) transformation of our representation of unicorns so that their non-existence was either not recognized, or else not allowed to veto the afternoon's plans. That is, the check on whether the plans were realistic would not be carried out, so that we would not draw the inference that "There is no point in forming the goal of finding a unicorn[11] from the judgment "There are no unicorns".

Searching for unicorns, then – and also refusing or pretending to do so – are human activities that depend on the functioning of specific rules and representations in the mind. If these are transformed (by learning, reasoning, or fancy) then the person's thought and action relating to unicorns will be different.

These differences in behaviour and experience may be subtle or coarse-grained, for the representations concerned are varied and complex. Planning a unicorn-hunt, whether for fun or for real, requires that our minds contain more than the concept of "unicorn". We must also understand the concepts of "virgin" and "forest", and must be able to represent their probable location and recognize them when we get there. We must be able to plan how to reach the forest, and how to creep up on virgin and unicorn without frightening either. And if we are to have any hope of catching the unicorn we would be wise not to forget the halter. If a silken halter is not available, would a hemp one do instead? According to our current sensibilities, it probably would. But according to an older, magical, viewpoint, it might not.

The psychological interest here is not in which conception of unicorn-hunting is true and which false, but in how it is possible for people to have and be guided by such mental representations, whether realistic or not. Understanding how something is possible is more important here than predicting what will actually happen. To be sure, if we know that someone believes a unicorn to be a sea-creature, half fish half woman, we can predict that they will not search for unicorns in the forest. But the theoretical interest is in how the familiar concept of unicorn can be integrated with a person's powers of perception, planning, and persuasion, so as to generate a unicorn-hunt. This integration may be extremely complex, involving comparisons of priorities (what else might one do this afternoon?) and individual life-styles (which of one's friends would appreciate the enterprise?).

Even in the Middle Ages, life was not focussed solely on unicorns. People can solve other problems, of varying types -- from cooking a meal through designing motorbikes to writing sonnets. Similarly, people hold beliefs of different sorts about different kinds of things, beliefs that are largely idiosyncratic and not always consistent with each other ("Do I contradict myself? Very well then, I contradict myself. I contain multitudes."). An adequate theoretical psychology should help us to understand how it is possible for all these problems and beliefs to coexist in individual minds.

How do we recognize different types of problem, and classify some as tractable and others as hopelessly beyond our grasp? What mental processes enable us to build or acquire our various beliefs? How do we inter-relate them, inferring one belief from another with more or less confidence, or recognizing an actual or potential inconsistency? If we decide not to tolerate an inconsistency, how do we transform the content and/or organization of our minds accordingly? For example, how do we relate evolutionary biology and the various theological forms of Christianity: are they possible mental bedfellows, and if so what sort of conceptual bolster might need to be put down the middle of the bed?

No-one, at present, is able to answer all these questions. But the important point is that they are the sort of question to which computational psychology is especially well-suited. For they concern the ways in which we store, access, compare, and transform various sorts of symbolically represented information. Indeed, psychological questions in general -- whether they concern belief, problem-solving, purpose, choice, language, perception, memory, or even emotion -- are to be understood as computational questions about mental rules and representations.

A connoisseur of the history of psychological fashions might observe that these newly arrived computationalists are not the first to style the mind as a domain of symbolic representation and transformation. Freud, for example, offered some suggestive ideas about various sorts of symbolic transformation, conceptual association and mental compartmentalization. He thought of the "defence mechanisms" as involving different sorts of psychological transformation. "Introjection" and "displacement", for instance, transform the object of one's neurotic hatred in distinct ways: the former changes it from another person to oneself; the latter shifts the hatred to some third person, conceived of as somehow analogous to (symbolic of) the original. Similarly, both the dream-work and slips-of-the-tongue involve strings

and comparisons as well as on generally interpretable symbols (such as tunnels).

But Freud's ideas – like those of other non-computational psychologists – are suggestive rather than specific, vaguely expressed rather than rigorously defined. The computational approach, by contrast, offers precisely definable concepts, because a program has to be expressed clearly (as a set of instructions defining specific symbol-manipulations) if the computer is to accept it. If the program is written in a high-level programming language, the programmer can ignore the more basic information-processing involved (much as we may think of a task in terms of high-level goals, being unable to specify the details of how we tackle it). But, since clarity is essential, artificial intelligence is a rich source of clear distinctions between many types of symbolic representation and interpretative process. (Some programmed models of Freud's theory of repression exist; while they are admittedly very crude, they enable one to see – for instance – that there are various different sorts of "denial", which are not clearly distinguished in Freud's writings CBoden, 1977, chs. ii-iiij.)

Moreover, this style of theorizing highlights process as well as structure, since a program has to tell the computer not only what result to produce but also how to produce it. Non-computational psychologists often take psychological change for granted, assuming that it can be sufficiently specified by stating the initial and final mental states involved. However, the process of mental transformation is itself problematic. In a programming context, a failure to suggest any way in which the change might be effected will show up as a glaring gap in the program, a gap over which the uninstructed computer is unable to leap. Some computational account of how to make the leap must be supplied if the program is to function. In short, the pictures of the mind that are designed in the computational style are more like movies than pin-ups.

For example, what goes on in our minds when we understand what someone means by the word "it" on any particular occasion? You may remember the following snatch of conversation from Alice's Adventures in Wonderland;

> "... Even Stigand, the patriotic Archbishop of Canterbury, found it advisable—"
>
> "Found what?" asked the duck.
>
> "Found it," replied the mouse, rather crossly. "Of course you know what •Tt{ means?"
>
> "I know what Mt[1] means when 1^ find a thing," said the duck. "It's generally a frog, or a worm. The question is, what did the Archbishop find?"

Obviously, in the duck's last remark, the word "it" refers to the thing found by the duck, whereas in the first sentence of the exchange the word "it" does not refer to a thing at all (which is why the question here is fiat "what did the Archbishop find?").

I said these facts are "obvious," and at an intuitive level so they are. But stating the grammatical principles involved here (so as to say precisely what the duck's mistake was, in asking what the Archbishop had found) is very difficult. And it is wellnigh impossible to suggest a series of psychological processes for interpreting the word "it" in its

might explain what goes on in our minds when we understand e
language.

Or, rather, it is wellnigh impossible without the discipl
programming. If a computer program is to be able to interp
translate) language correctly, it has to be given both the r
grammatical rules and specific procedures for applyin
intelligently in distinct contexts. For instance, a wellknown
language-program [Winograd, 1972] engaged in a "conversation" w
programmer that involved seventeen distinct uses of the word "it
you specify the differences between these three out of the sevent
"Find a block which is taller than the one you are holding and
into the box"; "How many objects did you touch while you wer
it?"; and "Does the shortest thing the tallest pyramid's
supports support anything green?" -- "Yes, the green pyramid." -
colour is it?" (For that matter, can you say what processes went
your mind to interpret the words "... support supports support
This man-machine interchange was generated by a complex p
precisely specifying ways of integrating various sorts of knowled
context-sensitive way. In the examples quoted, the program us
knowledge of grammatical rules, of the current state of the (bl
the) world, and of people's probable interests (one does not n
enquire as to the colour of something that has just been identi
"green").

However, this program is not as clever as it looks, for the
many everyday uses of language which it simply cannot handl
example, it cannot decipher incomplete or grammatically
sentences. Only in such extreme cases as the tortured phrases
Watergate tapes are people similarly flummoxed. No-one knows just
are able to cope with incomplete sentences in ordinary langua
since this program (unlike some others) cannot achieve this res
all, it cannot be an adequate model of human language use, and ca
interpreting language in quite the same way that we do.

This example brings to mind a common criticism of the comput
approach to psychology. Critics often point out that if a c
program can achieve a certain result (such as recognizing a u
playing chess, or interpreting the word "it") it does not foll
people reach that result in the same way. We should remember that
are many different levels at which, or respects in which, on
specify "the way" in which a program does something; it might do
the same way" as people do under one description, but in "a di
way" under another. Nevertheless, this does not affect the main
of the criticism, that one cannot pass directly from computati
program to thinking in a person.

This is of course correct. It is a special case of the
truth that a theory that fits the known facts may not be t
possible theory of those facts (remember Ptolemy and Copernicus
may indeed be incompatible with facts discovered in the fut
principle, science can offer us no cast-iron epistemological guar
So scientists have to do the best they can with the best they hav

If an alternative theory is available, which accounts for im
aspects of the subject-matter that the first theory ignores or
then some workers will concentrate on that alternative and t
approaches can be compared. Sometimes two conflicting theories

used by the same scientists: for many years in the history of optics, apparently incompatible theories of light were employed (the wave theory and the corpuscular theory), each of which was able to explain specific experimental results inexplicable by the other. Only with the advent of field-theory in the nineteenth century were these two approaches theoretically reconciled. To reject a theory when one has nothing (or nothing of comparable plausibility and rigour) to put in its place is irrational, unless there are compelling reasons for believing it to be useless. Even authors who see the computational approach as in some sense radically misconceived often admit that it may be scientifically useful for generating psychological hypotheses. This is why many psychologists who doubt that this methodology will answer all their questions are nonetheless prepared to use it until its limits can be established.

This newly fashionable talk of "computers" and "programs", however, is not acceptable in all salons. For many people see this mode of theorizing as the "punk-gear" of psychology, as an aggressive rejection of traditional styles that manages also to threaten our own self-confidence.

From this viewpoint, the computational approach appears to offer a chilling picture of humanity that is not only false but also dangerously dehumanizing. It has been criticized as an "obscene" and "deeply humiliating" view, one that will deaden our personal responses and our valuation of purpose, desire, and emotional life [Weizenbaum, 1976]. It is bad enough, such critics complain, to say (with Freud) that we are driven by irrational drives and uncontrollable anxieties, or (with Skinner) that like rats or pigeons we are slaves responsive only to environmental conditioning. But to put us on a par with computers is even worse than bringing us down to the level of unreasoning beasts. Little wonder, then, that such humanists accuse proponents of the computational approach of being "mad", or of seeking to make others so.

Like beauty, however, madness and threat may be in the eye of the beholder. These common fears of the computational approach are mistaken. They rest on a failure to realize that describing a system (whether person or computer) as a symbol-manipulating system is conceptually quite distinct from describing the physical hardware that embodies the computational powers concerned. The former type of description requires computational concepts, whereas the latter employs the terms of physics, chemistry, and physiology.

As the poet Blake foresaw, the natural sciences have encouraged a "single vision" that has to some degree undermined people's sense of personal autonomy and responsibility. This is unsurprising, for no science that lacks the concept of "representation" can even acknowledge humanity, still less explain it. This is equally true of those styles of theoretical psychology which, modelling themselves on the natural sciences, likewise eschew representational notions. The modes of psychology based on phenomenology, existentialism, and other "subjectivist" traditions escape this trap, and have some intuitive plausibility. But they are unacceptably vague, and are more successful in reminding us of what psychological phenomena are possible than in explaining how they are possible.

Computational psychology does not support the mechanization c world-picture that has been brought about by the natural science by such "scientific[11] styles of psychology as behaviourism. Fai being dehumanizing, it is – potentially – positively rehumar This is not to deny that some (though not all) of the technological applications of artificial intelligence may tf healthy social relations CEvans, 19803. But the conceptual nati this new approach enables it to distinguish "subjective[11] truths ideas, aspirations, and beliefs) from "objective" truths (about and other physical things). And, what is even more signifies concentrates firmly on the former.

Computational psychology emphasizes the richness and subth our mental powers, a richness that hitherto has often been intu* glimpsed (at least by poets and novelists) but never theorei recognized by psychologists. It admits the influence on our I" shared cultural beliefs, of individual ideas, interests, purpose* choice, and of self-reference and self-knowledge CBoden, Hofstadter, 1980D. And it provides rigorous hypotheses about the processes that underlie such influences and make them possible.

But, the power of fashion in psychology being so great, perh< computational style is a mere passing fancy? Is it a trendy fad *I* a technological society, doomed to obsolescence because of its irrelevance to human realities? Or is it a lasting contributic seed of the long-awaited general paradigm of psychology?

Its being currently "fashionable" need not debar it from the role. For although fashion is largely ephemeral, some modes last, rag-trade for instance, hats and halter-necks may be in or ou1 colours change from season to season – but shoes endure, c changes in their detailed form. How could they not, being so use* soft-footed walking creatures?

The computational style in psychology will survive likewise, is so well-suited to the representational anatomy of our mir offers us a lasting insight into important mental features – f< that other psychological approaches have recognized less cleai perhaps even wholly ignored. It illuminates not only our cognit intellectual powers, but also our capacities for purposive act moral choice. It will change, to be sure, and some of its change: doubtless be as shocking as next month's cover of Vogue. Many currently favoured types of computation will be superceded by \ even today, different types are preferred by different theorist! the computational approach will endure, for it has provided a si of rigour and clarity that must make us permanently dissatisfi(less.

ooooo0OOooooo

REFERENCES

BODEN, M. A. (1972) Purposive Explanation jn Psychology (Cam! Mass.: Harvard Univ. Press).

BODEN, M. A. (1977). Artificial Intelligence and Natural Man. York: Basic Books; Brighton: Harvester Press).

EVANS, C. (1980) The Mighty Micro (London: Gollancz).

HOFSTADTER, D. (1980) Godel, Escher, Bach: An Eternal Golden B (London: Penguin).

WEIZENBAUM, J. (1976) Computer Power and Human Reason: From Judgment Calculation (San Francisco: Freeman).

WINOGRAD, T. (1972) Understanding Natural Language (Edinburgh U Press).