Stalking the Wild Epistemic Engine

PATRICIA SMITH CHURCHLAND

PAUL M. CHURCHLAND

THE INSTITUTE FOR ADVANCED STUDY, PRINCETON
and
THE UNIVERSITY OF MANITOBA

INTRODUCTION

In some manner, devolving from Evolution's blind trials and blunders, densely crowded packets of excitable cells inevitably come to represent the world. The conglomeration which is the human brain standardly evolves an awesomely complex world-representation in short order and on the basis of scanty input. Less distinguished beasts such as slugs and sloths are presumed to have world-representations which are less rich, or anyhow, different. It is perhaps salutory here to bear in mind that some animals have sensory detectors where we are stony blind. Pigeons have tiny ferro-magnets for detecting the earth's magnetic field; rattlesnakes have infra-red detectors; electric fish have organs which discern small variations in electric fields, and so on [1]. It is remarkable also that in the human case the world-representation evolves, and it evolves not only during the lifetime of one human brain, but across the life-spans of collections of brains. But *how* can a brain be a world-representer? How can brains change so that some of their changes consist in learning about the world? How are representations used by a brain such that the output yields purposive and intelligent behavior?

Broadly speaking, research on the question of how the mind-brain works follows one of two methodological colors. The first is in substantial degree part of the rationalist tradition, emphasizing the linguistic and rule-following aspect of cognition, and is now prominently represented by cognitive/computational psychology, or by a substantial movement within that field.[1] The second is naturalistic in character, and is part of the tradition containing such thinkers as de la Mettrie, Darwin, Helmholtz and Hebb, and is the guiding framework for most neuroscientists and physiological psychologists. On closer inspection, the distinction fuzzes and smears at border spots, but the general contrasts are distinct enough. Which approach, if any, will succeed in treeing answers is an empirical question, and in recent work[2] we have argued that the odds favor the naturalistic approach. In this paper we turn to semantic questions, and confess at the outset to some trepidation. For one thing, semantics is a tar-baby. It is difficult to handle without becoming horribly stuck, and worse, once stuck, it is difficult to avoid the conviction that one is embraced by a verity. Additionally, the bounds of the paper exact brevity of presentation, and there are places where we have had to be ruthlessly synoptic.

Unadorned, the gist of the paper is twofold. The first and more familiar point is that computational psychology should seek a wider conception of cognitive processes than is embodied in a sentential/rationalistic model. The second point, however is our main concern. We argue that, because computational psychology is quite properly methodologically solipsistic (we will explain what this means shortly), it cannot provide, and should not be expected to provide, a theory of how a representational system *hooks up to the world*. Insofar, it cannot explain how the representing creature survives and flourishes in the environment the creature is struggling to represent. To make good this deficit, we probe the possibilities for a naturalistic strategy. But first, a few remarks are needed on the contrast between naturalistic and non-naturalistic approaches.

Simplifying to the very bone, the dominant hypotheses of the rationalist version of computational psychology are as follows.

(a) The paradigm of the information-bearing or representational state is the propositional attitude, where the obj ct of the attitude (its content) is a sentence.

(b) In cognitive activity, the transitions between representational states are a function of the logical relations holding between the *contents* of those states.

(c) Such representations, and the transitions between them, can thus be modelled or realized in a computer. [10], [12], [16], [20].

Church's Thesis says that whatever is computable is Turing computable. Assuming, with some safety, that what the mind-brain does is computable, then it can in principle be simulated by a computer. So what needs to be done is to figure out the program that mimics what cognitive organisms do. Fortunately, goes the rationale (and here is where we start to disagree), the essentially correct basis for devising that program can be found in the propositional attitudes of folk psychology. Extensions and innovations are to be expected, but folk psychological characterizations of the nature of representational structures are fundamentally correct. This is providential, since it means that part of the theory sought is already in hand, and moreover the work can be done without so much as opening a skull and implanting an electrode, and no one has to feed the animals and clean the cages. For those who are squeamish about looking nature in her occasionally noissome face, this assurance of remoteness comes as a relief.

A different strategy inspires the naturalist. The naturalist is moved by two large-scale intellectual visions: the evolution of complex nervous systems from simpler nervous systems, and secondly, the displacement of primitive theories, and their ontologies, by more encompassing and more powerful theories. Enthusiasm for the computational strategy is gutted by the observation that folk theories about the way one or other part of the universe works have typically lost out in the competition for explanatory space. They have, in the light of new theories, been revealed as misdirected, narrow, animistic, and misconceived in varying degrees—despite their having passed as uncontestable truths of common sense for eons. The history of science is littered with the dry bones of folk theory. Even as folk theories of the nature of fire, of the sky, of matter, of heat, of light, of space, of life, of numbers, of weather and climate, of birth and death and disease—even as these folk theories have succumbed to the sharper tooth and fleeter foot of modern biology, chemistry, physics, etc., so it would not be surprising to find folk psychology primitive and inadequate in competition with newer theories of how the mind-brain works. We do not say folk psychology *must* be as inadequate as, say, alchemy, but only that it would be astonishing if it alone amongst folk theories happened to be good enough to survive. A far, far more complex and devious object of wonder than heat or light, the brain is unlikely to have been adequately groped by folk theory in the misty dawn of emerging verbalization [5].

In loosening the grip of the bonds of common sense psychology, the naturalist suggests we view ourselves as epistemic engines [3]. Call an epistemic engine any device that exploits a flow of environmental energy, and the information it already contains, to produce more information, and to guide movement. So far as natural (wild) epistemic engines are concerned, survival depends on a fit between the information contained and the world it inhabits. For example, a simple bottom creature who has sensory neurons which happen to be responsive to changes in magnetic field will not benefit from such responses unless the changes are related to its feeding, fleeing, fighting or reproducing. The choice of the word 'engine' is more apt than might be supposed, since it original meaning is 'native intelligence' or 'mother wit,' and is the source of 'ingenuity.' The planet abounds with a wondrous profusion of epistemic engines; building nests and bowers, peeling bark, dipping for termites, hunting wildebeests, and boosting themselves off the planet altogether. The human brain is but one result of Evolution's blind maunderings, and like other creatures with a nervous system, we too are epistemic engines. Accordingly, cast within the naturalist's framework, the problem consists in figuring out how epistemic engines work.

In considering the problem, the naturalist suggest we dethrone language as the model for the structure and dynamics of representational activity generally. Representations—information-bearing structures—did not emerge of a sudden with the evolution of verbally competent animals. As Sellars [21] remarks, "...the generic concept of a representation admits of many gradations between primitive systems and the sophisticated systems on which philosophers tend to concentrate" (p. 15). Whatever information-bearing structures humans enjoy, such structures have evolved from simpler structures, and such structures are part of a *system* of information-bearing structures and structure-manipulating processes. If we want to understand how epistemic engines work, we might have to understand simpler systems first, and that means we cannot avoid penetrating the skull, implanting electrodes, and looking nature full in the face.

THE FORMALITY CONDITION: EPISTEMIC ENGINES ARE SYNTACTIC ENGINES

The central insight of computational psychology is that intentional and purposive behavior is the outcome of mental states and operations, where a mental state is characterized as standing in a relation to a representation, and where mental operations are defined over representations. Computers are formal machines, in the sense that they operate on symbols in virtue of the form of the symbol, not in virtue of how the symbol may be interpreted. As Fodor [14] puts it:

> Formal operations are the ones that are specified without reference to the semantic properties of representations, as, e.g., truth, reference, or meaning (p. 277).

The basic point can be put in the following way: the machine goes from one state to another because it is caused to do so. If the machine treats two tokens differently, it will be because they have a formal difference in virtue of which the machine can discriminate them, and if they are formally indistinguishable, then the machine cannot distinguish them either.

The *formality condition* says that cognitive states are type-distinct only if the representations which constitute their objects are formally distinct. Fodor [14] has argued that computational psychology should honor the formality condition. This yields the position known as *methodological solipsism:* the causal explanation of cognitive processes must proceed without reference to whatever semantic properties our cognitive states may or may not have. So far as cognitive activity is concerned, semantics enters the picture inessentially, and only insofar as it has a purely syntactic image.

The question spawned by this methodological point is this: what *is* the semantics which has a syntactic image (i.e. syntactic stand-ins for semantic features)? The question is more approachable and familiar if put this way: what criteria of ascribing content to mental states specifies content which has a syntactic image, and what criteria fail to so specify content?

Churchland [3] and Stich [24] have made the observation that ascription of beliefs and desires (and propositional attitudes generally) to others is fundamentally akin to translation. Stich has developed the point, showing that when I ascribe the belief that p to Trudeau, this is to be analyzed roughly as saying that Trudeau has a mental state which is like the one I would be in were I to sincerely say this: p. Wtihout tarrying over the niceties, notice that the ascription is a *similarity* judgement, and that it makes ineliminable reference to oneself and *to one's own representational system*. In that respect, such ascriptions are observer relative. Deepening his analysis, Stich has then convincingly argued that like similarity judgements in other domains, these similarity judgements (e.g. belief ascriptions) vary as a function of *which* criteria are used, and that a hodge-podge of criteria jockey for position. Depending on purpose, context, and sundry other considerations, one criterion may be preferred to another, and application of different criteria may well give conflicting descriptions of what Trudeau believes. Sometimes sameness of natural-kind-reference is made to count (cf. Putnam, [19]), sometimes it is not. Sometimes conceptual role counts more, sometimes ideological similarity counts, sometimes linguistic practice figures in (see Burge [2]), sometimes social practice takes precedence over conceptual role and so on (see Stich [24], [25]).

The question raised earlier can now be asked again: if computational psychology is to abide by the formality condition, which way of ascribing content will specify content which has a syntactic stand-in? Evidently the criterion which counts beliefs as different if the references are different will not do. Stich's own view, and we concur, is that the best choice will be the criterion which specifies similarity in terms of functional role (which he calls 'narrow causal role,' and is close enough to 'conceptual role'). The point here is this: if what is wanted in computational psychology are generalizations describing routes from input to output, and the only semantic features relevant are features the machine can detect, then the semantic content of representations must be fixed by their conceptual role, by their narrow causal role, because this is what co-varies with differences in a representation's intrinsic formal structure. This kind of content we call *translational content*, since similarity of conceptual role is what faithful translation attempts to capture. Thus, when I ascribe to Trudeau the belief that dopamine is a neurotransmitter, I am saying that Trudeau stands in the belief-relation to a representation that plays the same inferential/causal/functional role in his representational system that "Dopamine is a neurotransmitter" (or its internal analogue) plays in *my* representational system (See Sellars [22].) A translational mapping has been postulated between Trudeau and me. Notice that the relation is not between Trudeau's representations and some part of *the world*, but between Trudeau's representional system and *my representational system*.

Already it will be evident that a computational psychology which confines its semantics to conceptual role semantics will depart from folk psychology in certain minor ways. That is surely inevitable and part of what progress here requires. It is also important to mention that Stich argues that for a scientific psychology, even translational content is too much semantics, and that the semantics will have to be laundered out altogether. His worry here stems from the fact that the description of content is observer-relative. He finds this troublesome because for one thing, to the extent that another organism's representational system diverges from mine, I cannot ascribe content (translational content is what we are confined to now) to the other's representations. If the generalizations of cognitive psychology require specification of content, then the generalizations will be incapable of reaching cases where content specification is uncertain. Most obviously this happens in the case of pre-verbal children, humans from different cultures, humans with brain damage, and as well, the entire animal kingdom. Such a psychology threatens to be a psychology of Me-And-My-Friends, and chartered provincialism is a methodological *faux pas*. Hence Stich's attempt to see if computational psychology can define operations over uninterpreted syntactic objects [25].

We are convinced that computational psychology should honor the formality condition. Moreover, P.M. Churchland [4] has argued that methodological solipsism can be derived from entirely naturalistic assumptions. The brain is evidently a syntactic engine, for a neuron cannot know the distant causal ancestry or the distant causal destiny of its input and outputs. An activated neuron causes a creature to withdraw into its shell not because such activation represents the presence of a predator—though it may indeed represent this—but because that neuron is conne ted to the withdrawal muscles, and because its activation is of the kind that causes them to contract. The 'semantic' content of the state, if any, is causally irrelevant.

What does adherence to the formality condition signify for the research program of computational psychology? It means, for one thing, that questions about how mental states hook up to the world are questions it simply shelves as not within its proper province. On the one hand this is fine, but on the other it means that a completed computational psychology is nonetheless a radically *incomplete* theory of how humans work. For if it has nothing whatsoever to say about how representational systems represent features in the world, it has left out a crucial part of the theory. It is like a genetic theory which tells us how genes produce phenotypic traits, but which throws up its hands on the matter of the relation between the traits and the world the organism inhabits.

Permit us to milk the point briefly. Organisms are syntactic engines. Yet via the nervous system an organism exhibits behavior suited to its surroundings. For example, honey bees remove dead bees from the hive, a herring gull chick pecks its mother's bill for food, a person drinks polio vaccine. Now even if computational psychology did dope out the internal cognitive program, it would still seem miraculous that a person's being in a certain state, syntactically described as say, I-P38, is followed by his drinking polio vaccine, rather than by his moving his rook or by his firing his bazooka or what have you (see also Fodor [15]). The point is that *it seems* that brains do what they do in virtue of the referents of their assorted states, inasmuch as there is a stupendously good fit between representational systems and the world. Of course if I should specify the content of Trudeau's state by saying he intends to drink polio vaccine, I am specifying the *translational content*; I am saying his representational system is like mine, and his hooks up to the world the way mine does, *whatever way that is*. And that is no *theory* of the way representational sytems hook up to the world.

Perhaps this can all be avoided by saying that mind-brain states have intrinsic of-ness or original intentionality or ultimate aboutness, call it what you will. To some it may seem a plain, observable fact that epistemic engines—fancy ones anyhow—operate on states with intrinsic of-ness. It may even be conjectured that this is what makes mental states mental. The suggestion, to the extent that it makes sense, is unappealing. For one thing, it gives up just when things get particularly exciting. And it is a bit like explaining the nature of life by citing 'original vitality,' or the nature of neuronal responsiveness by citing 'original excitability,' and, insofar, it is a way of trumping up a virtue out of being stumped. Moreover, the exasperating thing about 'plain facts' is that they often turn out to be neither plain nor factual. Intrinsic of-ness is an illusion, like intrinsic up-ness or intrinsic down-ness.

Another possibility here will be to base a theory of how representations hook up to the world on the idea that the content of a subject's mental state is linked to truth conditions for the content sentence. The analysis of 'Jones knows the meaning of p' in terms of 'Jones knows what conditions could make p true' is the basis. While tbere is more to be said here, our simple response is that this strategy will not work because it connects internal aspects of Jones' representational system—namely *his taking* (believing, etc.) the meaning to be thus and such with *his taking* the truth conditions to be thus and such—and hence does not clear the fence and tell us anything whatever about how representational systems hook up to the world.

A more promising suggestion addresses the *causal* relations that hold between representations and states of the world. For example, it may be that a specific representation R_1 occurs in a creature's 'perceptual belief-register' only when something in its environment is F. We could thus ascribe "$(\exists x)$ (Fx)" as R_1's propositional content. Other representations may be similarly keyed to other aspects of the environment, and we can thus ascribe content to each of them. This would constitute the first stage of a theory which would then go on to develop a wider account of how representations less tightly keyed to the environment acquire content (see, e.g., Stampe [23], Dretske [26].

For perceptually-sensitive representations, one can indeed ascribe propositional content in this way, and on the basis of real causal connections with the world. We call content thus ascribed, *calibrational content*, since this procedure is just another instance of calibrating an instrument of measurement or detection. The states of living creatures do indeed carry systematic information about the environment, in virtue of their law-governed connections with it.

Our enthusiasm for this approach to how representations hook up to the world must be dimmed, however, by three serious problems. First, it is very difficult to see how to make the jump from ascribing content to representations at the sensory periphery to assigning content to the dominant mass of representations not so conveniently tied to aspects of the environment.

Second, and more important, even for perceptual representations, the contents assigned in this way are not identical with their more familiar translation contents. These two kinds of content can and often do diverge radically. A Neanderthal's representation might have the calibrational content, "The wind is producing atmospheric resonances," but have the translational content, "The Storm God is howling at us." An Aristotelean's representation might have the calibrational content, "The planet beneath me has a non-zero angular momentum," but have the translational content, "The crystal sphere above me is turning." A Puritan's representation might have the calibrational content, "She is epileptic," but have the translational content, "She is possessed by Satan." In general, calibrational contents do nothing to reflect how the representing creature happens to conceive of things. This approach fails to explicate how our familiar *translational* contents hook up with the world.

Third, and equally important, the dynamic or functional properties of a representation, within one's overall cognitive economy, are not determined by its calibrational content, but by its 'formal' or 'structural' properties. An account of how a representational system hooks up to the world should make some contact with the system's behavior over time, but calibrational content is dynamically irrelevant. This is just the thesis of methodological solipsism showing itself again.

In sum, a causal approach must disappoint some of our original expectations regarding a general account of how epistemic, representing creatures 'hook up' to the world. But we should not despair immediately. Perhaps it is those original expectations that need schooling. Perhaps we should not expect that all epistemic creatures must have representations that are somehow like sentences, and that a satisfactory account of the important hook-ups must address the relation between singular terms and things, the relation between predicates and properties, and the relation between sentences and states of affairs. Perhaps we should not expect this even for ourselves. Let us explore the problem without making these assumptions.

NEUROSCIENCE: CALIBRATIONAL AND COMPUTATIONAL

In its bones, neuroscience is also solipsistic; it must honor the formality condition. How then does neuroscience expect to deal with the question of how representational systems hook up to the world? For if it sees the brain as syntactic, then it does seem miraculous that a sequence of events in a herring gull's brain results in its asking for food, or a sequence in a bee's brain results in its taking a particular flight path to nectar-heavy blossoms.

In the case of such animals as bees and slugs, we confidently expect to be able to defrock the mystery by giving evolutionary/neurobiological/neuroethological explanations. The idea is to treat the organism's nervous system as something which evolution calibrates (i.e. as something which, by random mutation and on random selection, is tuned to measure, via the excitable cells, certain features in the environment). When responses involved in measuring such features happen to be linked to motor responses relevant to survival, then the probability is enhanced that the organism's genes will be passed on. Bees are 'calibrated,' by natural selection, to detect oleic acid, and are tuned to produce a motor sequence which results in their lugging its source (dead bees) out of the hive. Herring gull chicks are calibrated to detect small red spots on moving objects (this picks out their mother's bill), and are tuned to peck at them, which results in their being fed. The neuronal story of how this works is not beyond us, and for simple creatures ascription of calibrational content to states of their humble nervous sytem is well underway.

What then of complex organisms such as humans? The story which defrocks the miracle in the case of the bees and the herring gull is relatively simple, in that their behavior is essentially fixed-action-pattern stuff. The story which defrocks the miracle in the case of persons will be much harder to ferret out. For here the enigma ramifies because these organisms are spectacular learners. Seemingly, they learn about the world, though in the syntactic spirit one would say, roughly, that their syntactic organization is fancied up with the end result that they do new things, where some of those things enhance their survival chances. The miracle now is how the syntactic engine ends up as advantageously tuned as it is. How is it that the fancy organisms "make hypotheses die in their stead"? And learn to become increasingly proficient at doing so? How is it that a syntactic engine evolved so that certain of its states seem to have intentionality? How can a person come to have an I-P38 state such that this state typically causes it to drink polio vaccine? Not, evidently, in virtue of evolution's *directly* selecting for a match between P-38 states and drinking polio vaccine, but rather, one guesses, in virtue of evolution's selecting 'learner-planner rigmaroles.' If the person acquires new concepts, that is, of course, a syntactical affair, but what is the causal story in virtue of which it can invent concepts which surpass the old when what is on the receiving end of the predictions are events in the world? The miracle is that the organism has become so "well-tuned" that it seems to have *an evolving world-picture*, rather as though the organism has tricked up an analogue of the evolutionary process itself.

Here then the job for neurobiology and neuroethology is Herculean, but the bets are that the story for complex organisms will build on the more basic story of calibrational semantics for simpler organisms, following the steps of evolution itself. The backbone of what we are calling *calibrational content* is the observation that there are reliable, regular, standardized relations obtaining between specific neural responses on the one hand, and *types* of states in the world. The notion exploits the fact that specific neural responses are regularly caused by types of state in the organism's normal environment. Inching closer to a working definition, we suggest the following:

> A state S of a system O contains the calibrational content P if and only if O would not be in S unless P, with some high degree of probability n/m.

For example, normally the receptor cells of the rattlesnake's pit organ respond only if there is a warm object within half a metre or so of the pit. That is, with very high probability, the receptors are not excited unless there is a warm object in the vicinity. The probability is less than 1 because receptors might be caused to respond by oddball things, such as an ethologist's injecting a drug into the pit, or by malfunction, as when the tired old receptors of a senescent rattlesnake fire spontaneously. In any case, these occurences are rare. Crudely formalized, we can say, where x ranges over the relevant receptor cells and y ranges over objects in the environment:

1. $(x)\ (\text{Excited}\ [x] \rightarrow (\text{Prob}\ (\exists y)\ (\text{Warm}[y]) = .98)$.

Moreover, given the snake's environment, there is a decent probability, let us say, .7, that the warm object is warm-blooded prey, such as a rabbit, mouse, etc. The probability is less than 1 because the pit receptors can be excited by something warm which is not the customary comestible—like a sun-heated rock, or a smoldering ember. The probability will vary with night and day, being higher at night, and of course it will plummet if the snake is put in an ethologist's laboratory filled with light bulbs and kettles. But assume his standard environment. Then we can crudely formalize the relation between warmth and food:

2. $(x)\ (\text{Warm}\ [x] \rightarrow (\text{Prob}(\text{Food}[x])) = .7)$.

Accordingly, the excitation of receptors is a moderately reliable indicator of the presense of food; .7 x .98 = .686. In a primitive infrared engine, an excited cell may be a sign for warm blooded prey often enough that it can rely on that simple connection to guide its motor response. In fact of course evolution cranks up the probability as it stumbles on better correlations between neural responses and food, but some of the fine tuning of responses will not be at the sensory periphery, but will be deeper in the neural network. The rattlesnake is eminently better tuned than the simple infrared engine. Information from the pit organ is sent to the optic tectum in a two-stage relay, and in the tectum there is integration of visual and infrared information. Some tectal cells, for example, respond with a brief high-frequency burst of impulses only when there is a *small, moving, warm* object nearby, where the visual system provides the movement data, and the infrared system provides the temperature data. In particular, these cells do not respond to hot rocks [17]. These cells represent small, moving, warm objects—their excited state contains calibrational content to the effect that there are small, moving warm objects nearby. Of these deeper cells, we might now say, again crudely:

3. (x) Excited$[x] \supset$ Prob $(\exists y)$ (Small$[y]$ & Moving$[y]$ & Warm$[y]$)$=.98$))

where x ranges ovr the relevant tectal cells, and y ranges over objects. Now in the rattlesnake's environment, the probability is, say, .97, that small, moving warm objects are mice, so we can say with high probability ($.98 \times .97 = .95$) that the relevant cells would not be excited unless there were mice in the snake's vicinity. Excitation of these cells represents the presence of mice nearby.

The 'computations' executed by the preceding system are of course trivial, and yet the story provides us with a useful conception of the snake's representational attunement to certain aspects of its environment, a conception which helps explain how the snake survives and flourishes. And none of the story ascribes representations with a sentence-like syntax, or talks about the reference of terms or the meaning of predicates.

But what of more talented creatures, creatures whose computational activities are more broadly directed and more intricately constituted? In particular, what of creatures in whom *learning* is a major element in their progressive attunement to the environment? In such cases, are we not forced to postulate an entire *system* of representations, manipulable by the creature? It seems that we are. But here we must resist our parochial impulses concerning the structure of such a system. In the first place, *such a system need not and almost certainly will not be monolithic or uniform at all*. More likely, we possess an integrated hierarchy of quite different computational/representational systems, facing very different problems and pursuing quite different strategies of solution. Why should we expect the representational systems used by the visual system, the auditory system, the proprioceptive system, and the motor system all to be the same? Even the cytoarchitecture of the relevant brain areas is different for each of these cognitive sub-systems.

Will *some* sub-system of this functional mosaic display the familiar structures of human language? In humans, presumably yes, though other species need not possess it. And even in humans it may play a relatively minor role in our overall cognitive activities, serving a mainly social function. The bulk of cognition may take place in other subsystems, and follow principles inapplicable in the linguistic domain. What those other representational systems are, and how they are knit together to form human cognition, these are empirical questions, begging empirical answers. Lesion studies from neurology are one source of answers: the accidental destruction of isolated brain areas leaves people with isolated and often very curious cognitive deficits. The direct examination of active neural netw is another, though here animal studies must dominate. The computer simulation of proposed representational sytems will also be invaluable, if it is neurophysiologically guided, since computers will allow us to defeat the problems of sheer functional complexity in the systems we discover. Research is well-established in all three of these areas, and it wants only our attention.

Our conclusion is that computational psychology cannot afford to embrace a principle of categorical aloofness from or methodological disdain for neuroscience, for at least two reasons. First, if we want to know how cognitive creatures hook up to the world they inhabit, neuroscience holds out the best hope for an enlightening account. And second, even if we restrict our concern to the brain's abstract computational activities, empirical neuroscience will provide authoritative data on just what those activities are, and on their many varieties. In particular, neuroscience holds out the best hope for understanding the individual evolutionary process we call *learning*, since the elements of variation, and the mechanisms of selection, whatever they are, are there under the skull, awaiting our exploration. A truly informed story of how the human cognitive system hooks up to the world must await their discovery and examination.

REFERENCES

[1] T. H. Bullock, R. Orkand, and A. Grinnel, *Introduction to Nervous Systems* (San Francisco: W. H. Freeman, 1977).

[2] Tyler Burge, "Individualism and the Mental," in P. A. French, T. E. Uehling and H. K. Wettstein, eds., *Midwest Studies in Philosophy, Vol IV* (Minneapolis: University of Minnesota Press, 1979).

[3] Paul M. Churchland, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979).

[4] ———, "In Defense of Naturalism," *The Behavioral and Brain Sciences*, 3(1980): 74-75.

[5] ———, "Eliminative Materialism and the Propositional Attitudes," *The Journal of Philosophy*, 78(1981): 67-90.

[6] ———, "Is *Thinker* a Natural Kind?," *Dialogue*, (forthcoming).

[7] Patricia S. Churchland, "A Perspective on Mind-Brain Research," *The Journal of Philosophy*, 77(1980): 185-207.

[8] ———, "Neuroscience and Psychology: Should the Labor be Divided?," *The Behaviorial and Brain Sciences*, 3(1980): 133.

[9] ———, "Mind-Brain Reduction: New Light From Philosophy of Science," *Neuroscience*, 7(1982): 1041-1047.

[10] Daniel Dennett, "Artificial Intelligence as Philosophy and as Psychology," *Brainstorms* (Cambridge, MA: MIT Press/Bradford, 1978).

[11] ———, "Three Kinds of Intentional Psychology," R. Healey, editor, *Reduction, Time and Reality* (Cambridge: Cambridge University Press, 1981): 37-61.

[12] Jerry A. Fodor, *The Language of Thought* (New York: Thomas Y. Crowell, 1975).

[13] ———, "Computation and Reduction," C. Wade Savage, editor, *Minnesota Studies in the Philosophy of Science* Vol. 9 (Minneapolis: University of Minnesota Press, 1978); reprinted in his *Representations* (Cambridge, MA: MIT Press/Bradford, 1981): 146-174.

[14] ———, "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," *The Behavioral and Brain Sciences*, 3(1980); reprinted in his *Rrepresentations*(Cambridge, MA: MIT Press/Bradford, 1981): 225-253.

[15] ———, "Tom Swift and His Procedural Grandmother," *Cognition*, 6(1978); reprinted in his *Representations* (Cambridge, MA: MIT Press/Bradford, 1982): 204-224.
[16] John Haugeland, "Semantic Engines: An Introduction to Mind Design," *Mind Design* (Cambridge, MA: MIT Press/Bradford, 1981): 1-34.
[17] Eric A. Newman and Peter H. Hartline, "The Infrared 'Vision' of Snakes," *Scientific American*, 246(1982): 116-127.
[18] Hilary Putnam, "The Meaning of 'Meaning'," in Keith Gunderson, editor, *Minnesota Studies in the Philosophy of Science*, Vol. 7 (Minneapolis: University of Minnesota Press, 1975): 131-193.
[19] A. Pellionisz and R. Llinas, "Brain Modeling by Tensor Network Theory and Computer Simulation: Distributed Processor for Predictive Coordination," *Neuroscience*, 4(1979): 323-348.
[20] Zenon Pylyshyn, "Computation and Cognition: Issues in the Foundations of Cognitive Science," *The Behavioral and Brain Sciences* 3(1980): 111-132.
[21] Wilfrid Sellars, "Behaviorism, Language, and Meaning," *Pacific Philosophical Quarterly*, 61(1980): 3-25.
[22] ———, "Mental Events," *Philosophical Studies*, 39(1981): 325-345.
[23] Dennis W. Stampe, "Toward a Causal Theory of Linguistic Representation", in French, Uehling and Wettstein, editors, *Midwest Studies in Philosophy* Vol. 2 (Minneapolis: University of Minnesota Press, 1972): 82-102.
[24] Stephen P. Stich, "On the Ascription of Content," in Andrew Woodfield, editor, *Thought and Object* (Oxford: Clarendon Press, 1981): 153-206.
[25] ———, *Folk Psychology and Cognitive Science: Tbe Case Against Belief* (Cambridge, MA: MIT Press/Bradford, forthcoming).
[26] Fred I. Dretske, *Knowledge and the Flow of Information* (Bradford/M.I.T. Press, 1981).

NOTES

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada, no. 410-81-0182.

[1]See especially Jerry Fodor [12], [13], and [14], and Zenon Pylyshyn [20].

[2]See Paul M. Churchland [3], [4], [5], and [6], and Patricia S. Churchland [7], [8], and [9].