



ELSEVIER

Journal of Cognitive Systems Research 1 (1999) 5–17

Cognitive Systems
RESEARCH

www.elsevier.com/locate/cogsys

Where brain, body, and world collide[☆]

Action editor: Ron Sun

Andy Clark

Philosophy/Neuroscience/Psychology Program, Department of Philosophy, Washington University, St. Louis, MO 63130, USA

Received 1 June 1999; accepted 20 June 1999

Abstract

The brain fascinates because it is the biological organ of mindfulness itself. It is the inner engine that drives intelligent behavior. Such a depiction provides a worthy antidote to the once-popular vision of the mind as somehow lying outside the natural order. However, it is a vision with a price. For it has concentrated much theoretical attention on an uncomfortably restricted space; the space of the inner neural machine, divorced from the wider world which then enters the story only via the hygienic gateways of perception and action. Recent work in neuroscience, robotics and psychology casts doubt on the effectiveness of such a shrunken perspective. Instead, it stresses the unexpected intimacy of brain, body and world and invites us to attend to the structure and dynamics of extended adaptive systems — ones involving a much wider variety of factors and forces. Whilst it needs to be handled with some caution, I believe there is much to be learnt from this broader vision. The mind itself, if such a vision is correct, is best understood as the activity of an essentially *situated* brain: a brain at home in its proper bodily, cultural and environmental niche. © 1999 Elsevier Science B.V. All rights reserved.

1. Software

Humans, dogs, ferrets; these are, we would like to say, mindful things. Rocks, rivers and volcanoes are not. And no doubt there are plenty of cases in between (insects, bacteria, etc.). In the natural order, clear cases of mindfulness always involve creatures with brains. Hence, in part, the fascination of the brain: understanding the brain looks crucial to the project of understanding the mind, but how should such an understanding proceed?

An early sentiment — circa 1970 and no longer much in vogue — was that understanding the mind depended rather little on understanding the brain. The brain, it was agreed, was in some sense the physical medium of cognition, but everything that mattered about the brain *qua* mind-producing engine turned not on the physical details but on the computational and information-processing strategies that the neural stuff (merely) implemented. There was something importantly right about this view, but something desperately wrong as well.

What was right was the observation that understanding the physical workings was not going to be enough; that we would need also to understand how the system was organized at a higher level in order to grasp the roots of mindfulness in the firing of neurons. This point is forcefully made by the cogni-

[☆]‘Where Brain, Body, and World Collide’ reprinted by permission of *Daedalus*, Journal of the American Academy of Arts and Sciences, from the issue entitled, ‘The Brain,’ Spring 1998, Vol. 127, No. 2.

E-mail address: andy@twinearth.wustl.edu (A. Clark)

tive scientist Brian Cantwell Smith who draws a parallel with the project of understanding ordinary computer systems. With respect to e.g., a standard PC running a tax-calculation program, we could quite easily answer all the ‘physiological’ questions (using source code and wiring diagrams) yet still lack any real understanding of what the program does or even how it works.¹ To really understand how mental activity yields mental states, many theorists believe, we must likewise understand something of the computational/information-processing organization of the brain. Physiological studies may contribute to this understanding, but even a full physiological story would not, in and of itself, reveal how brains work *qua* mind-producing engines.²

The danger, of course, was that this observation could be used as an excuse to downplay or marginalize the importance of looking at the biological brain *at all*. And so it was that, in the early days of Cognitive Science, it was common to hear real neuroscience dismissed as having precious little to offer to the general project of understanding human intelligence. Such a dismissal, however, could not long be sustained (for a powerful critique, see Reeke & Edelman, 1988). For although it is (probably) true to say that a computational understanding is in principle independent³ of the details of any specific implementation in hardware (or wetware), the project of *discovering* the relevant computational description (especially for biological systems) is surely not.

One key factor is evolution. Biological brains are the product of biological evolution and as such often fail to function in the ways we (as human designers) might expect (see e.g., Simon, 1962; Dawkins, 1986; Clark 1997, ch. 5). This is because evolution is both constrained and liberated in ways we are not. It is constrained to build its solutions incrementally via a series of simpler but successful ancestral forms. The human lung, to give one example, is built via a

process of tinkering (Jacob, 1977) with the swim bladder of the fish. The engineer might design a better lung from scratch. The tinkerer, by contrast, must take an existing device and subtly adapt it to a new role. From the engineer’s ahistorical perspective, the tinkerer’s solution may look bizarre. Likewise, the processing strategies used by biological brains may surprise the computer scientist. For such strategies have themselves been evolved via a process of incremental, piecemeal, tinkering with older solutions.

More positively, biological evolution is *liberated* by being able to discover efficient, but ‘messy’ or unobvious solutions: ones which may, for example, exploit environmental interactions and feedback loops so complex that they would quickly baffle a human engineer. Natural solutions (as we will later see) can exploit just about any mixture of neural, bodily and environmental resources along with their complex, looping, and often non-linear interactions. Biological evolution is thus able to explore a very different solution space (wider in some dimensions, narrower in others) than that which beckons to conscious human reason.

There are, of course, ways around this apparent mismatch. Computationalists lately exploit so-called ‘genetic algorithms’ (for a review, see Clark, 1997, ch. 5) that roughly mimic the natural process of evolutionary search and allow the discovery of efficient but loopy and interactive adaptive strategies. Moreover, hard neuroscientific data and evidence is increasingly available as a means of helping expand our imaginative horizons so as to better appreciate the way real biological systems solve complex problems. The one-time obsession with the ‘software level’ is thus relaxing, in favor of an approach which grounds computational modeling in a more serious appreciation of the biological facts. In Section 2 I highlight some challenging aspects of such recent research — aspects that will lead us directly to a confrontation with the situated brain.

2. Wetware, and some robots

Recent work in cognitive neuroscience underlines the distance separating biological and ‘engineered’ problem solutions, and displays an increasing aware-

¹See B.C. Smith (1996), p. 148. Note that Smith’s worry, at root, concerns the gap between physiological and semantic or intentional questions.

²Thus we read, for example, that computational approaches make possible “a science of structure and function divorced from material substance [that] ... can answer questions traditionally posed by psychologists” (Pylyshyn, 1986, p. 68).

³See e.g., David Marr’s (Marr, 1982) distinction between the levels of computation, algorithm and implementation.

ness of the important interpenetration — in biological systems — of perception, thought and action. Some brief examples should help fix the flavor.

As a gentle entry point, consider some recent work on the neural control of monkey finger motions. Traditional wisdom depicted the monkey's fingers as individually controlled by neighboring groups of spatially clustered neurons. According to this story, the neurons (in Motor Area 1, or M1) were organized as a 'somatotopic map' in which a dedicated neural sub-region governed each individual digit, with the sub-regions arranged in lateromedial sequence just like the fingers on each hand. This is a tidy, easily conceptualized solution to the problem of finger control, but it is the engineer's solution, not (it now seems) that of Nature.

Marc Schieber and Lyndon Hibbard (Schieber & Hibbard, 1993) have shown that individual digit movements are accompanied by activity spread pretty well throughout the M1 hand area, and that precise, single-digit movements actually require *more* activity than some multi-digit whole hand actions (such as grasping an object). Such results are inconsistent with the hypothesis of digit-specific local neuronal groups. From a more evolutionary perspective, however, the rationale and design is less opaque. Schieber (1990, p. 444) conjectures that the basic case, from an evolutionary perspective, is the case of whole hand grasping motions (used to grab branches, to swing, to acquire fruits, etc.) and that the fundamental neural adaptations are thus geared to the use of simple commands which exploit inbuilt synergies⁴ of muscle and tendon so as to yield such coordinated motions. The 'complex' coordinated case is thus evolutionarily basic and neurally atomic. The 'simple' task of controlling e.g., an individual digit represents the harder problem and requires more neural activity viz, the use of some motor cortex neurons to *inhibit* the synergetic activity of the other digits. Precise single-digit movements

require the agent to tinker with whole-hand commands, modifying the basic synergetic dynamics (of mechanically linked tendons, etc.) adapted to the more common task.

Consider next a case of perceptual adaptation. The human perceptual system can, we know (given time and training), adapt in quite powerful ways to distorted or position-shifted inputs. For example, subjects can learn how to coordinate vision and action while wearing lenses that invert the entire visual scene so that the world initially appears upside-down. After wearing such lenses for a few days, the world is seen to flip over — various aspects of the world now appear to the subject to be in the normal upright position. Remove the lenses and the scene is again inverted until readaptation occurs.⁵ Thach, Goodkin & Keating (1992) used a variant of such experiments to demonstrate the motor-specificity of some perceptual adaptations. Wearing lenses which shifted the scene *sideways* a little, subjects were asked to throw darts at a board. In this case, repeated practice led to successful adaptation,⁶ but of a motor-loop specific kind. The compensation did not 'carry over' to tasks involving the use of the non-dominant hand to throw, or to an underarm variant of the visual overarm throw. Instead, adaptation looked to be restricted to a quite specific combination of gaze angle and throwing angle: the one used in overarm, dominant-hand throwing.

Something of the neural mechanisms of such adaptation is now understood. It is known, for example, that the adaptation never occurs in patients with generalized cerebellar cortical atrophy, and that inferior olive hypertrophy leads to impaired adaptation. On the basis of this and other evidence, Thach et al. speculated that a learning system implicating the inferior olive and the cerebellum (linked via climbing fibers) is active both in prism adaptation and in the general learning of patterned responses to frequently encountered stimuli. The more general lesson, however, concerns the nature of the perception-action system itself. For it increasingly appears that the simple image of a general purpose perceptual system delivering input to a distinct and fully

⁴The notion of synergy aims to capture the idea of links that constrain the collective unfolding of a system comprising many parts. For example, the front wheels of a car exhibit a built-in synergy which allows a single driver 'command' (at the steering wheel) to affect them both at once. Synergetic links may also be learnt, as when we acquire an automated skill, and may be neurally as well as brute-physiologically grounded. See Kelso (1995), pp. 38, 52.

⁵For a survey of such experiments, see Welch (1978).

⁶In this case, *without* any perceived shift in the visual scene.

independent action system is biologically distortive. Instead, perceptual and action systems work together, in the context of specific tasks, so as to promote adaptive success. Perception and action, in this view, form a deeply interanimated unity.

Further evidence for such a view comes from a variety of sources. Consider, for example, the fact that the primate visual system relies on processing strategies that are not strictly hierarchic but instead depend on a variety of top-to-bottom and side-to-side channels of influence. These complex inner pathways allow a combination of multiple types of information (high-level intentions, low-level perception and motor activity) to influence all stages of visual processing. The Macaque monkey (to take one well-studied example) possesses about 32 visual brain areas and over 300 connecting pathways. The connecting pathways go both upwards and downwards (e.g., from V1 to V2 and back again) and side-to-side (between subareas in V1); see e.g., Felleman and Van Essen (1991). Individual cells at ‘higher’ levels of processing, such as V4 (visual area 4) do, it is true, seem to specialize in the recognition of specific geometric forms. However, they will each also respond, in some small degree, to many other stimuli. These small responses, spread unevenly across a whole population of cells, can carry significant information. The individual cells thus function not as narrowly tuned single feature detectors but as widely tuned filters reacting to a whole range of stimulus dimensions (see Van Essen & Gallant, 1994). Moreover, the responses of such cells now look to be modifiable both by attention and by details of local task-specific context (Knierim & Van Essen, 1992).⁷

More generally, back projecting (corticocortical) connections tend, in the monkey, to outnumber forward ones i.e., there are more pathways leading from deep inside the brain outwards towards the sensory peripheries than vice versa.⁸ Visual processing may thus involve a variety of criss-crossing

influences which could only roughly, if at all, be described as a neat progression through a lower-to-higher hierarchy.

Such complex connectivity opens up a wealth of organizational possibilities in which multiple sources of information combine to support visually guided action. Examples of such combinations are provided by Churchland, Ramachandran and Sejnowski (1994) who offer a neurophysiologically grounded account of what they term ‘interactive vision.’ The interactive vision paradigm is there contrasted with approaches that assume a simple division of labor in which perceptual processing yields a rich, detailed inner representation of the visual scene which is later given as input to the reasoning and planning centers which in turn calculate a course of action and send commands to the motor effectors. This simple image (of what roboticists call a ‘sense–think–act’ cycle) is, it now seems, not true to the natural facts. In particular:

1. Daily agent–environment interactions often do not require the construction and use of detailed inner models of the full visual scene.
2. Low-level perception may ‘call’ motor routines that yield *better perceptual input* and hence improve information pick-up.
3. Real-world actions may sometimes play an important role in the computational process itself.
4. The internal representation of worldly events and structures may be less like a passive data-structure or description and more like a direct recipe for action.

Evidence for proposition 1 comes from a series of experiments in which subjects watch images on a computer screen. Subjects are allowed to examine an on-screen pictorial display. Then, as they continue to saccade around the scene (focusing first on one area, then another) small changes are made to the currently unattended parts of the display. The changes are made during the visual saccades. It is an amazing fact that, for the most part,⁹ quite large changes go unnoticed: changes such as the replacement of a tree

⁷In a somewhat related vein, Caminiti, Johnson, Burnod, Galli & Ferrina (1990) show that the directional preference of individual cells encoding reaching movement commands varies according to initial arm position. See also discussion in Jeannerod (1997), ch. 2.

⁸Though much of the connectivity is reciprocal. See Van Essen & Anderson (1990), by Churchland, Ramachandran and Sejnowski (1994), p. 40.

⁹The exception is if subjects are told in advance to watch out for changes to a certain feature. See McConkie (1990), Churchland, Ramachandran & Sejnowski (1994).

by a shrub, or the addition of a car, deletion of a hat and so on. Why do such gross alterations remain undetected? A compelling hypothesis is that the visual system is not even attempting to build a rich, detailed model of the current scene but is instead geared to using frequent saccades to retrieve information *as and when it is needed* for some specific problem-solving purpose. This fits nicely with Yarbus' classic (Yarbus, 1967) finding that the pattern of such saccades varies (even with identical scenes) according to the type of task the subject has been set (e.g., to give the ages of the people in a picture, to guess the activity they have been engaged in, etc.). According to both by Churchland, Ramachandran and Sejnowski (1994) and Ballard (1991), we are prone to the illusion that we constantly command a rich inner representation of the current visual scene because we are able to perform fast saccades, retrieving information as and when required. (An analogy:¹⁰ a modern store may present the illusion of having a massive amount of goods stocked on the premises, because it always has what you want when you want it. However, modern computer ordering systems can automatically count off sales and requisition new items so that the necessary goods are available just when needed and barely a moment before. This fine-tuned ordering system offers a massive saving of on-site storage whilst tailoring supply directly to customer demand.)

Contemporary research in robotics avails itself of these same economies. One of the pioneers of 'new robotics,' Rodney Brooks (see e.g., Brooks, 1991) coined the slogan, "the world is its own best model" to capture just this flavor. A robot known as Herbert (Connell, 1989), to take just one example, was designed to collect soft drink cans left around a crowded laboratory, but instead of requiring powerful sensing capacities and detailed advance planning, Herbert got by (very successfully) using a collection of coarse sensors and simple, relatively independent, behavioral routines. Basic obstacle avoidance was controlled by a ring of ultrasonic sound sensors that brought the robot to a halt if an object was in front of it. General locomotion (randomly directed) was interrupted if Herbert's simple visual system detected a roughly table-like outline. At this point a new

routine kicks in and the table surface is swept using a laser. If the outline of a can is detected, the whole robot rotates until the can is centered in its field of vision. This simple physical action simplifies the pick-up procedure by creating a standard action-frame in which a robot arm, equipped with simple touch sensors, gently skims the table surface dead ahead. Once a can is encountered, it is grasped, collected and the robot moves on. Notice, then, that Herbert succeeds without using any conventional planning techniques and without creating and updating any detailed inner model of the environment. Herbert's 'world' is composed of undifferentiated obstacles and rough table-like and can-like outlines. Within this world the robot also exploits its own bodily actions (rotating the 'torso' to center the can in its field of view) so as to greatly simplify the computational problem involved in eventually reaching for the can. Herbert is thus a simple example both of a system that succeeds using minimal representational resources and one in which gross motor activity helps streamline a perceptual routine (as suggested in proposition (2) above).

The interactive vision framework envisages a more elaborate natural version of this same broad strategy viz., the use of a kind of perceptuo-motor loop whose role is to make the most of incoming perceptual information by combining multiple sources of information. The idea here is that perception is not a passive phenomenon in which motor activity is only initiated at the endpoint of a complex process in which the animal creates a detailed representation of the perceived scene. Instead, perception and action engage in a kind of incremental game of tag in which motor assembly begins long before sensory signals reach the top level. Thus, early perceptual processing may yield a kind of proto-analysis of the scene, enabling the creature to select actions (such as head and eye movements) whose role is to provide a slightly upgraded sensory signal. That signal may, in turn, yield a new proto-analysis indicating further visuomotor action and so on. Even whole-body motions may be deployed as part of this process of improving perceptual pick-up. Foveating an object can, for example, involve motion of the eyes, head, neck and torso. Churchland, Ramachandran and Sejnowski (1994, p. 44) put it well: "Watching Michael Jordan play basketball or a group of ravens steal a caribou corpse from a wolf

¹⁰Thanks to David Clark for pointing this out.

tends to underscore the integrated, whole-body character of visuomotor coordination.” This integrated character is consistent with neurophysiological and neuroanatomical data which show the influence of motor signals in visual processing. There are — to take just two small examples — neurons sensitive to eye position in V1, V3 and LGN (lateral geniculate nucleus), and cells in V1 and V2 that seem to know in advance about planned visual saccades (showing enhanced sensitivity to the target, see Churchland, Ramachandran & Sejnowski, 1994, p. 44; Wurz & Mohler, 1976).

Moving on to proposition (3) (that real-world actions may sometimes play an important role in the computational process itself), consider the task of distinguishing figure from ground (the rabbit from the field, or whatever). It turns out that this problem is greatly simplified using information obtained from head movement during eye fixation. Likewise, depth perception is greatly simplified using cues obtained by the observers own self-directed motion. As the observer moves, close objects will show more relative displacement than farther ones. That is probably why, as by Churchland, Ramachandran and Sejnowski (1994, p. 51) observe, head bobbing behavior is frequently seen in animals: “A visual system that integrates across several glimpses to estimate depth has computational savings over one that tries to calculate depth from a single snapshot.”

And so to proposition (4): that the neural representation of worldly events may be less like a passive data structure and more like a recipe for action. The driving force, once again, is computational economy. If the goal of perception and reason is to guide action (and it surely is, evolutionary speaking), it will often be simpler to represent the world in ways rather closely geared to the kinds of actions we want to perform. To take a simple example, an animal that uses its visual inputs to guide a specific kind of reaching behavior (so as to acquire and ingest food) need not form an object-centered representation of the surrounding space. Instead, a systematic metrical transformation (achieved by a point-to-point mapping between two topographic maps) may transform the visual inputs directly into a recipe for reaching out and grabbing the food. In such a set-up, the animal does not need to do any computational work on an action-neutral

inner model as to plan a reaching trajectory. The perceptual processing is instead tweaked, at an early stage, in a way dictated by the particular use to which the visual input is dedicated. This strategy is described in detail in P.M. Churchland’s (Churchland, 1989, ch. 5) account of the ‘connectionist crab,’ in which research in Artificial Neural networks¹¹ is applied to the problem of creating efficient point-to-point linkages between deformed topographic maps.

In a related vein, Maja Mataric of the MIT Artificial Intelligence Laboratory has developed a neurobiologically inspired model of how rats navigate their environments. This model exploits the kind of layered architecture¹² also used in the robot Herbert. Of most immediate interest, however, is the way the robot learns about its surroundings. As it moves around a simple maze, it detects landmarks which are registered as a combination of sensory input and current motion. A narrow corridor thus registers as a combination of forward motion and short lateral distance readings from sonar sensors. Later, if the robot is required to find its way back to a remembered location, it retrieves¹³ an interlinked body of such combined sensory and motor readings. The stored ‘map’ of the environment is thus immediately fit to act as a recipe for action, since the motor signals are part of the stored knowledge. The relation between two locations is directly encoded as the set of motor signals that moved the robot from one to the other. The inner map is thus *itself* the recipe for the necessary motor actions. By contrast, a more classical approach would first generate a more objective map which would then need to be *reasoned over* in order to plan the route.

The Mataric robot (which is based on actual rat neurobiology; see McNaughton & Nadel (1990)) and the connectionist crab exemplify the attractions of what I call ‘action-oriented representations’ (Clark, 1997, p. 49): representations that describe the world

¹¹For an accessible introduction, see P.M. Churchland (1995).

¹²This is known as a ‘subsumption’ architecture, because the layers each constitute a complete behavior-producing system and interact only in simple ways such as by one layers subsuming (turning off) the activity of another or by one layer’s co-opting and hence ‘building-in’ the activity of another. See Brooks (1991).

¹³By a process of spreading activation amongst landmark encoding nodes. See Mataric (1991).

by depicting it in terms of possible actions.¹⁴ This image fits in nicely with several of the results reported earlier, including the work on monkey finger control and the motor loop specificity of ‘perceptual’ adaptation. The products of perceptual activity, it seems, are not always action-neutral descriptions of external reality. They may instead constitute direct recipes for acting and intervening. We thus glimpse something of the shape of what by Churchland, Ramachandran and Sejnowski (1994, p. 60) describe as a framework that is ‘motocentric’ rather than ‘visuocentric.’

As a last nod in that same direction, consider the fascinating case of so-called ‘mirror neurons’ (Di-Pelligrino, Klatzky & McCloskey, 1992). These are neurons, in monkey ventral premotor cortex, that are action-oriented, context-dependent and implicated in both self-initiated activity and passive perception. They are active both when the monkey observes a specific action (such as someone grasping a food item) and when the monkey performs the same action, where sameness implies not mere grasping but the grasping of a food item (see also Rizzolatti, Fadiga & Fogassi, 1996). The implication, according to the psychologist and neuroscientist Marc Jeannerod is that “the action... to be initiated is stored in terms of an action code, not a perceptual one” (Jeannerod, 1997, p. 191).

Putting all this together suggests a much more integrated model of perception, cognition and action. Perception is itself tangled up with possibilities for action and is continuously influenced by cognitive contextual and motor factors. It need not yield a rich, detailed and action-neutral inner model awaiting the services of ‘central cognition’ so as to deduce appropriate actions. In fact, these old distinctions (between perception, cognition and action) may sometimes obscure, rather than illuminate, the true flow of effect. In a certain sense, the brain is

revealed not as (primarily) the engine of reason or quiet deliberation, but as the organ of *environmentally situated control*. Action, not truth and deductive inference, are the key organizing concepts. This perspective, however, begs to be taken a step further. For if brains are best understood as controllers of environmentally situated activity, then might it not be fruitful to locate the neural contribution as just one (important) element in a complex causal web spanning brains, bodies and world? This potential gestalt shift is the topic of Section 3.

3. Wideware

Let us coin a term, ‘wideware’, to refer to states, structures or processes that satisfy two conditions. First, the item in question must be in some intuitive sense environmental: it must not, at any rate, be realized within the biological brain or the central nervous system. Bodily aspects and motions, as well as truly external items such as notebooks and calculators, thus fit the bill. Second, the item (state, structure, process) must play a functional role as part of an *extended cognitive process*: a process geared to the promotion of adaptation success via the gathering and use of knowledge and information, and one that loops out in some non-trivial way, so as to include and exploit aspects of the local bodily and environmental setting.

Of course, even what is intuitively a fully internalized cognitive process will usually involve contact with the external world. That much is demanded by the very ideas of knowledge acquisition and information-gathering. The notion of wideware aims to pick out instead a somewhat more restricted range of cases — ones in which the *kind* of work that cognitive science has typically assigned to the inner workings of the brain is at least partly carried out by processes of storage, search and transformation realized using bodily actions and/or a variety of external media. Understanding the human mind, I shall argue, will require us to attend much more closely to the role of such bodily actions and external media than was once anticipated. To better fix the notion of wideware itself, consider first a very simple case involving ‘bodily backdrop’.

Computer controlled machines are sometimes used

¹⁴Such representations bear some resemblance to what the ecological psychologist, J.J. Gibson called ‘affordances,’ although Gibson himself would reject our emphasis on inner states and encodings. For an affordance is the potential of use and activity that the local environment offers to a specific kind of being: chairs afford sitting (to humans), and so on. See Gibson (1979). The philosopher Ruth Millikan has developed a nice account of action-oriented representation under the label ‘pushmipullyu representation’. See Millikan (1995).

to fit small parts into one another. The error tolerance here is very low and sometimes the robot arm will fail to make a match. A computationally expensive solution exists in which the control system includes multiple feedback loops which signal such failures and prompt the machine to try again in a minutely different orientation. It turns out, however, that a much cheaper, more robust and more efficient procedure is simply to mount the assembler arms on rubber joints that ‘give’ along two spatial axes. This bodily backdrop allows the control device to dispense with the complex feedback loops. Thanks to the rubber mountings the parts “jiggle and slide into place just as if millions of tiny feedback adjustments to a rigid system were being continuously computed” (Mitchie & Johnson, 1984, p. 95).

Mere bodily backdrop, however, does not really count as an instance of an extended cognitive process. The computational and information based operations are reduced, courtesy of the brute physical properties of the body, but they are not themselves extended *into* the world. Genuine cognitive and computational extension requires, by contrast, that the external or bodily operations are themselves usefully seen as performing cognitive or information-processing operations. A simple phototropic (light following) robot — to take another negative example — does not constitute an extended cognitive or computational system. For although the presence of some external structure (light sources) is here vital to the robot’s behavioral routines, those external structures are not doing the kind of work that cognitive science and psychology has typically assigned to inner neural activity. They are not acting so as to manipulate, store or modify the knowledge and information that the organism uses to reach its goals. Sometimes, however, external structures and bodily operations do seem to be proper parts of the cognitive and computational processes themselves. Thus consider the use, in recent interactive vision research, of so-called *deictic pointers*. A classical (non-deictic) pointer is an inner state that can figure in computational routines but that can also ‘point’ to further data-structures. Such pointing allows for both the retrieval of additional information and the binding of one memory location content to another. One use of classical pointers is thus to bind information about spatial location (e.g., ‘top left corner of visual field’)

to information about current features (e.g., ‘bright yellow mug’), yielding complex contents (in this case, ‘bright yellow mug in top left corner of visual field’; see Pylyshyn (1986) for discussion). Since neural processing involves (more or less) distinct channels for e.g., object properties, object location and object motion, binding is an important element in neural computation, but binding, it now seems, can sometimes be achieved by the use of actual bodily orientations instead of linking inner data structures. The story is complex, but the basic idea is straightforward. It is to set up a system so that bodily orientations (such as saccadic eye motions leading to object fixation) directly yield the kinds of benefits associated with classical binding. This Ballard, Hayhoe, Pook & Rao (in press) show how to use visual scene fixation to directly associate the features represented (the properties of the perceived object) with an external spatial location. Another example is the use of so-called ‘do-it-where-I’m-looking’ processing routines in which a bodily motion (e.g., grasping) is automatically directed to whatever location happens to be currently fixated in the visual field. In all these cases, the authors comment:

The external world is analogous to computer memory. When fixating a location the neurons that are linked to the fovea refer to information computed from that location. Changing gaze is analogous to changing the memory reference in a silicon computer (Ballard, Hayhoe, Pook and Rao, in press, p. 6).

Deictic binding provides a clear example of a case in which bodily motion does the *kind* of work that cognitive science has typically assigned only to inner neural activity. The deictic strategies use a combination of inner computation and gross bodily action to support a type of functionality once studied as a property of the inner neural system alone. Taking this perspective a step further, we next consider the use of external media as both additional memory and as potent symbol-manipulating arenas.

Portions of the external world, it is fairly obvious, often function as a kind of extra-neural memory store. We may deliberately leave a film on our desk to remind us to take it for developing. Or we may write a note ‘develop film’ on paper and leave that

on our desk instead. As users of words and texts, we command an especially potent means of off loading data and ideas from the biological brain onto a variety of external media. This trick, I think, is not to be underestimated. For it affects not just the quantity of data at our command, but also the kinds of operation we can bring to bear on it. Words, texts, symbols and diagrams often figure *intimately* in the problem-solving routines developed by biological brains nurtured in language-rich environmental settings. Human brains, trained in a sea of words and text, will surely develop computational strategies that directly ‘factor-in’ the reliable presence of a variety of external props and aids. The inner operations will then complement, but not replicate, the special manipulative potentials provided by the external media.

Consider, for example (borrowed from Clark, 1995), the process of writing an academic paper. You work long and hard and at day’s end you are happy. Being a good physicalist, you assume that all the credit for the final intellectual product belongs to your brain: the seat of human reason, but your are too generous by far. For what really happened was (perhaps) more like this. The brain supported some re-reading of old texts, materials and notes. Whilst re-reading these, it responded by generating a few fragmentary ideas and criticisms. These ideas and criticisms were then stored as more marks on paper, in margins, on computer discs, etc. The brain then played a role in re-organizing these data, on clean sheets, adding new on-line reactions and ideas. The cycle of reading, responding and external re-organization is repeated, again and again. Finally, there is a product; a story, argument or theory. However, this intellectual product owes a lot to those repeated loops out into the environment. Credit belongs to the agent-in-the-world. The biological brain is just a part (albeit a crucial and special part) of a spatially and temporally extended process, involving lots of extra-neural operations, whose joint action creates the intellectual product. There is thus a real sense (or so I would argue) in which the notion of the ‘problem-solving engine’ is really the notion of the *whole caboodle*: the brain and body operating within an environmental setting.

Consider, by way of analogy, the idea of a swimming machine. In particular, consider the

bluefish tuna.¹⁵ The tuna is paradoxically talented. Physical examination suggests it should not be able to achieve the aquatic feats of which it is demonstrably capable. It is physically too weak (by about a factor of 7) to swim as fast as it does, to turn as compactly as it does, to move off with the acceleration it does, etc. The explanation (according to the fluid dynamicists M. & G. Triantafyllou) is that these fish actively create and exploit additional sources of propulsion and control in their watery environments. For example, the tuna use naturally occurring eddies and vortices to gain speed, and they flap their tails so as to actively create additional vortices and pressure gradients which they then exploit for quick take-off’s, etc. The real swimming machine, I suggest, is thus the fish *in its proper context*: the fish plus the surrounding structures and vortices that it actively creates and then maximally exploits. The *cognitive machine*, in the human case, looks similarly extended (see especially Dennett, 1995, chs. 12 and 13). We actively create and exploit multiple linguistic media, yielding a variety of contentful structures and manipulative opportunities whose reliable presence is then factored deep into our problem-solving strategies.

4. Implications

Software, wetware and wideware, if our story is to be believed, form a deeply interanimated triad. The computational activities of the brain will be heavily sculpted by its biological ‘implementation’. And there will be dense complementarity and cooperation between neural, bodily and environmental forces and factors. What the brain does will thus be precisely fitted to the range of complementary operations and opportunities provided by bodily structure, motion and the local environment. In the special case of human agency, this includes the humanly generated ‘whirlpools and vortices’ of external, symbol-laden media: the explosion of wideware made available by

¹⁵The story is detailed in Triantafyllou & Triantafyllou (1995) and further discussed in Clark (1997). A 49 inch eight-segment anodized aluminum and lycra robot tuna is being used at MIT (Massachusetts Institute of Technology) to test the details of the theory.

the ubiquitous devices of language, speech, and text. The picture that emerges is undeniably complex, but what does it really mean both for our understanding of ourselves and for the practice of scientific inquiry about the mind?

Certain implications for our vision of ourselves are clear. We must abandon the image of ourselves as essentially disembodied reasoning engines; and we must do so not simply by insisting that the mental is fully determined by the physical, but by accepting that we are beings whose neural profiles are profoundly geared so as to press maximal benefit from the opportunities afforded by bodily structure, action, and environmental surroundings. Biological brains, at root, controllers of embodied action. Our cognitive profile is *essentially* the profile of an embodied and situated organism. Just how far we should then press this notion of cognitive extension, however, remains unclear. Should we just think of ourselves as cognitive agents who co-opt and exploit surrounding structures (e.g., pen and paper) so as to expand our problem-solving capacities, or is there a real sense in which the cognitive agent (as opposed to the bare biological organism) is thus revealed as an extended entity incorporating brain, body and some aspects of the local environment? Normal usage would seem to favor the former, but the more radical interpretation is not as implausible as it may initially appear.

Consider an example. Certain Alzheimer's sufferers maintain an unexpectedly high level of functioning within the normal community. These individuals should not — given their performance on a variety of standard tests — be capable of living as independently as they do. Their unusual success is explained only when they are observed in their normal home environments (see Edwards, Baum & Morrow-Howell, 1994; Baum, 1991), in which an array of external props and aids turn out to serve important cognitive functions. Such props and aids¹⁶ may include the extensive use of labels (on rooms, objects, etc.), the use of a 'memory book' containing annotated photos of friends and relatives, the use of a

diary for tasks and events, and simple tactics such as leaving all important objects (check book, etc.) in open view so as to aid retrieval when needed. The upshot, clearly, is an increased reliance on various forms of wideware (or 'cognitive scaffolding'; see Clark, 1997) as a means of counterbalancing¹⁷ a neurally based deficit, but the pathological nature of the case is, in a sense, incidental. Imagine a whole community whose linguistic and cultural practices evolved so as to counterbalance a normal cognitive profile (within that community) identical to that of these individuals. The external props could there play the same functional role (of complementing a certain neural profile) but without any overtone of pathology-driven compensation. Finally, reflect that our own community is just like that. Our typical neural profile is different, to be sure, but *relative to that profile*, the battery of external props and aids (laptops, filofaxes, texts, compasses, maps, slide-rules, etc.) play just the same role. They offset cognitive limitations built into the basic biological system. Now ask yourself what it would mean — in the case of the Alzheimer sufferer — to maliciously damage that web of external cognitive support? Such a crime has, as Daniel Dennett once remarked, much of the flavor of a harm to the *person*, rather than simple harm to property. The same may well be true in the normal case: deliberate theft of the poet's laptop is a very special kind of crime. Certain aspects of the external world, in short, may be so integral to our cognitive routines as to count as *part of the cognitive machinery* itself (just as the whorls and vortices are, in a sense, part of the swimming machinery of the tuna). It is thus something of a question whether we should see the cognizer as the bare biological organism (that exploits all those external props and structures) or as the organism-plus-wideware. To adopt the latter perspective is to opt for a kind of 'extended phenotype' view of the mind, in which the relation between the biological

¹⁶Some of this additional structure is maintained and provided by family and friends. (but similarly, much of our own wideware is provided by language, culture and institutions which we do not ourselves create).

¹⁷Such counterbalancing is, as Marcel Kinsbourne has usefully reminded me, a somewhat delicate and complex matter. The mere provision of the various props and aids is useless unless the patient remains located in a stable, familiar environment. And the ability of different patients to make use of such added environmental structure itself varies according to the nature and extent of the neurally-based deficit.

organism and the wideware is as important and intimate as that of the spider and the web (see Dawkins, 1982).

The implications for the specific study of the mind are, fortunately, rather less ambiguous. The vision of the brain as a controller of embodied and situated action suggests the need to develop new tools and techniques capable of investigating the brain (literally) in action: playing its part in problem-solving routines in (as far as possible) a normal ecological context. Very significant progress has already been made, of course. Non-invasive scanning techniques such as PET and fast MRI represent a giant leap beyond the use of single-cell recordings from anaesthetized animals, but we should not underestimate the distance that remains to be covered. For example, experimenters recently carried out some neural recordings from a locust in free flight (see Kutsch, Schaarz, Fischer & Kautz, 1993).¹⁸ This kind of ecologically realistic study is clearly mandated by the kinds of consideration we have put forward. Yet such investigations remain problematic due to sheer technological limitations. In the locust case, the researchers relied upon tiny radio transmitters implanted in the insect, but the information pick-up remained restricted to a scanty two channels at a time.

Moreover, it is not just the information-gathering techniques that need work, but also the analytic tools we bring to bear on the information gathered. If we take seriously the notion that brain, body and world are often united in an extended problem-solving web, we may need to develop analytic and explanatory strategies that better reflect and accommodate this dense interanimation. To this end, there are (so far) two main proposals on the table. One proposal (see e.g., Thelen & Smith, 1994; Port & Van Gelder, 1995) is to use the tools of dynamical systems theory. This is a well established mathematical framework for studying the temporal evolution of states within complex systems. Such a framework appeals, in part, because it allows us to use a single mathematical formalism to describe both internal and external organizations and (hence) allows us to treat complex looping interactions (ones that criss-cross

brain, body, and world) in a deeply unified manner. The other proposal (see e.g., Hutchins, 1995) is to take the kind of analysis we traditionally applied only to the inner states and extend it outwards. This means sticking with talk of representations and computations, but applying these ideas to extended organizations encompassing e.g., multiple individuals, maps, texts, and social institutions. (My own view, which I will not defend here, is that we need to combine the two approaches by defining new, dynamical (process-based) ways of understanding key terms like ‘representation’ and ‘computation’ (see Clark, 1997; Crutchfield & Mitchell, 1995; Mitchell, Crutchfield & Hrabner, 1994; Van Gelder, 1995).)

Finally, the pervasive notion of a neural code or codes is now in need of major overhaul. If the notion is to apply to real biological systems, it must be relieved of a good deal of excess baggage. Natural neural codes, as we saw in Section 2, will often be closely geared to the particular details of body and world. The coding for single finger motion is unexpectedly baroque, courtesy of the need to actually combat basic synergies created by the system of mechanically linked tendons. Other codes (for example, for whole hand grasping motions) prove unexpectedly simple and direct. Moreover, there is evidence of the widespread use of task-specific, motor-oriented and context-dependent encodings. Mirror neurons, recall, code for context-specific actions (e.g., grasping food) and function both in passive perceptions and active grasping; and interactive vision routines press large benefits from minimal and often task-specific forms of internal encoding. In all these cases, we discern a much more austere and action-oriented vision of neural encoding — a vision stripped of the excess baggage of a single, rich, language-like inner code (e.g., Fodor, 1975; Fodor & Pylyshyn, 1988).

The simple, almost common-sensical, notion of the brain as a system evolved so as to guide the actions of embodied agents in rich real-world settings thus yields substantial and sometimes challenging fruit. Gone is the vision of the environment as simply a source of problem-specifying inputs and an arena for action. Instead, both basic and imposed aspects of environmental order and complexity now emerge as fundamental components of natural problem-solving behavior. Gone too is the vision of the

¹⁸Thanks to Joe Faith for drawing this example to my attention.

human brain as an organ of pure reason. In its place, we encounter the brain as a locus of action-oriented and activity-exploiting problem-solving techniques, and as a potent generator and exploiter of cognition-enhancing wideware. Taking this vision seriously, and turning it into a concrete and multi-disciplinary research program, presents an exciting new challenge for the sciences of the mind.

Acknowledgements

I am very grateful to Stephen Graubard and the participants at the Daedalus Authors Meeting (Paris, October 1997) for a wealth of useful advice, good criticism and wise counsel. Special thanks to Jean Pierre Changeux, Marcel Kinsbourne, Vernon Mountcastle, Guilio Tonini, Steven Quartz, and Semir Zeki. As usual, any remaining errors are all my own.

References

- Ballard, D. (1991). Animate vision. *Artificial Intelligence* 48, 57–86.
- Ballard, D., Hayhoe, M., Pook, P. & Rao, R. (in press). Deictic codes for the embodiment of cognition. *Behavioral & Brain Sciences*.
- Baum, C. (1991). Addressing the needs of the cognitively impaired elderly from a family policy perspective. *American Journal of Occupational Therapy* 45(7), 594–606.
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence* 47, 139–159.
- Caminiti, R., Johnson, P., Burnod, Y., Galli, C., & Ferrina, S. (Johnson, Burnod, Galli & Ferrina, 1990). Shift of preferred directions of premotor cortical cells with arm movements performed across the workspace. *Experimental Brain Research* 82, 228–232.
- Churchland, P., Ramachandran, V., & Sejnowski, T. (Ramachandran & Sejnowski, 1994). A critique of pure vision. In: Koch, C., & Davis, J. (Eds.), *Large-Scale Neuronal Theories of the Brain*, MIT Press, Cambridge, MA.
- Churchland, P. M. (1989). *The Neurocomputational Perspective*, MIT/Bradford Books, Cambridge.
- Churchland, P. M. (1995). Learning and conceptual change. In: Clark, A., & Millican, P. (Eds.), *Essays in Honour of Alan Turing*, Oxford University Press, Oxford, UK.
- Clark, A. (1995). I am John's brain. *Journal of Consciousness Studies* 2(2), 144–148.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*, MIT Press, Cambridge, MA.
- Connell, J. (1989). A Colony Architecture for an Artificial Creature. *MIT AI Lab Tech Report II 5*, 1.
- Crutchfield, J., & Mitchell, M. (1995). The evolution of emergent computation. *Proceedings of the National Academy of Sciences* 92, 10742–10746.
- Dawkins, R. (1982). *The Extended Phenotype*, Oxford University Press, Oxford.
- Dawkins, R. (1986). *The Blind Watchmaker*, Longman, London.
- Dennett, D. (1995). *Darwin's Dangerous Idea*, Simon & Schuster, New York.
- DiPelligrino, J., Klatzky, R., & McCloskey, B. (Klatzky & McCloskey, 1992). Time course of preshaping for functional responses to objects. *Journal of Motor Behavior* 21, 307–316.
- Edwards, D., Baum, C., & Morrow-Howell, N. (Baum & Morrow-Howell, 1994). Home environments of inner city elderly: Do they facilitate or inhibit function? *The Gerontologist: Program Abstracts of 47th Annual Scientific Meeting* 34(1), 64.
- Felleman, D., & Van Essen, D. (1991). Distributed hierarchical processing in primate visual cortex. *Cerebral Cortex* 1, 1–47.
- Fodor, J. (1975). *The Language of Thought*, Crowell, New York.
- Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3–71.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*, Houghton-Mifflin, Boston.
- Hutchins, E. (1995). *Cognition in the Wild*, MIT Press, Cambridge, MA.
- Jacob, F. (1977). Evolution and Tinkering. *Science* 196, 1161–1166.
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*, Blackwell, Oxford.
- Kelso, S. (1995). *Dynamic Patterns*, MIT Press, Cambridge.
- Knierim, J., & Van Essen, D. (1992). Visual cortex: Cartography, connectivity and concurrent processing. *Current Opinion in Neurobiology* 2, 150–155.
- Kutsch, W., Schaarz, G., Fischer, H., & Kautz, H. (Schaarz, Fischer & Kautz, 1993). Wireless transmission of muscle potentials during free flight of a locust. *Journal of Experimental Biology* 185, 367–373.
- Marr, D. (1982). *Vision*, W.H. Freeman, San Francisco, CA.
- Mataric, M. (1991). Navigating with a Rat Brain: A neurobiologically inspired model for robot spatial representation. In: Meyer, J. A., & Wilson, S. (Eds.), *From Animals to Animats I*, MIT Press, Cambridge.
- McConkie, G.W. (1990). *Where vision and cognition meet*. Paper presented at the H.F.S.P. Workshop on Object and Scene Perception, Leuven, Belgium.
- McNaughton, B., & Nadel, L. (1990). Hebb-Marr networks and the neurobiological representation of action in space. In: Gluck, M., & Rumelhart, D. (Eds.), *Neuroscience and Connectionist Theory*, Erlbaum Associates, Hillsdale, NJ, pp. 1–64.
- Millikan, R. (1995). Pushmi-Pullyu representations. In: Tomberline, J. (Ed.), *Philosophical Perspectives 9: Connectionism and Philosophical Psychology*, Ridgeview, CA.
- Mitchell, M., Crutchfield, J., & Hraber, P. (Crutchfield & Hraber, 1994). Evolving cellular automata to perform computations. *Physica D* 75, 361–391.
- Mitchie, D., & Johnson, R. (1984). *The Creative Computer*, Penguin, UK.

- Port, R., & Van Gelder, T. (1995). It's about time: An overview of the dynamical approach to cognition. In: Port, R., & van Gelder, T. (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, MIT Press, Cambridge, MA, pp. 1–44.
- Pylyshyn, Z. (Ed.) (1986). *The Robots Dilemma The Frame Problem in Artificial Intelligence*, Ablex, Norwood.
- Reeke, G., & Edelman, G. (1988). Real brains and artificial intelligence. *Daedalus, Winter*, pp. 143–173.
- Rizzolatti, G., Fadiga, L., & Fogassi, L. (Fadiga & Fogassi, 1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3, 131–141.
- Schieber, M. (1990). How might the motor cortex individuate movements? *Trends in Neuroscience* 13(11), 440–444.
- Schieber, M., & Hibbard, L. (1993). How somatotopic is the motor cortex hand area? *Science* 261, 489–492.
- Simon, H. (1962). *Models of Bounded Rationality*, Vols. I and II, MIT Press, Cambridge, MA, Reprinted 1982.
- Smith, B. C. (1996). *On the Origin of Objects*, MIT Press, Cambridge, MA.
- Thach, W., Goodkin, H., & Keating, J. (Goodkin & Keating, 1992). The cerebellum and the adaptive coordination of movement. *Animal Review of Neuroscience* 15, 403–442.
- Thelen, E., & Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press, Cambridge, MA.
- Triantafyllou, M., & Triantafyllou, G. (1995). An efficient swimming machine. *Scientific American* 272(3), 64–71.
- Van Essen, D., & Anderson, C. (1990). Information processing strategies and pathways in the primate retina and visual cortex. In: Zornetzer, S., Davis, J., & Lau, C. (Eds.), *An Introduction to neural and Electronic Networks*, Academic Press, New York, pp. 43–72.
- Van Essen, D., & Gallant, J. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron* 13, 1–10.
- Van Gelder, T. (1995). What might cognition be if not computation. *Journal of Philosophy* 92(7), 345–381.
- Welch, R. (1978). *Perceptual Modification: Adapting to Altered Sensory Environments*, Academic Press, New York.
- Wurz, R., & Mohler, C. (1976). Enhancement of visual response in monkey striate cortex and frontal eye fields. *Journal of Neurophysiology* 39, 766–772.
- Yarbus, A. (1967). *Eye movements and vision*, Plenum Press, New York.