COMMENTS WELCOME

MIND & WORLD: BREACHING THE PLASTIC FRONTIER

Andy Clark

- definitely need a better title!

Where does mind stop and the rest of the world begin? The question invites two standard replies. Either we go with the intuitive demarcations of skin or skull, or we ~~assume that the question is really about the analysis of meaning and proceed to~~ debate the pro's and con's of the (broadly) Putnam-esque doctrine that "meaning just ain't in the head" (Putnam (1975)). I propose, however, to pursue a third position, viz that cognitive processes are no respecters of the boundaries of skin or skull. That is to say, I ~~claim (1) that the intuitive notion of the mind ought to be purged of its internalist learnings, and (2) that the reasons for so doing do not depend~~ argue for an externalist position that in no way depends on the debatable role of truth-conditions and real-world reference in fixing the meaning of mental or linguistic tokens.

This is a bit distracting, almost misleading. (~~Complex systems apply to~~ suggests networks of humans – which maybe isn't the image you want to bring up, just yet!)

Consider, first of all, a networked computer. Such a device will often farm out jobs to other machines on the network. This can occur if, for example, one of the other machines has already loaded the software needed to carry out a given sub-task, or if it is currently under-used and hence can devote more processing power to the task, thus speeding overall problem-solving (see e.g. Malone et al (1988)). In such cases, we have no hesitation in describing the overall problem-solving activity as distributed across several machines, each of which carry out distinct parts of the computation.

Now consider three possible cases of human problem solving:

(1) A person sits in front of a computer screen which displays images of various 3 D geometric shapes and is asked to answer questions concerning the potential fit of such shapes into depicted 'sockets'. To assess fit, the person

must mentally rotate the shapes so as to align them with the sockets.

(2) A person sits in front of a similar computer screen, but this time can choose either to physically rotate the image on the screen, by pressing a rotate button, or to mentally rotate the image as before. We can also suppose, not unrealistically (see below) that some speed advantage accrues to the physical rotation operation.

(3) Sometime in the cyberpunk future, a person sits in front of a similar computer screen. This agent, however, has the benefit of a implant which can perform the rotation operation as fast as the computer in the previous example. The agent must still choose which internal resource to use (the implant or the good old fashioned mental rotation), as each resource makes different demands on attention and other concurrent brain activity.

Cases could have been multiplied. We might have imagined genetically engineered versions of case (3), or human mutants who spontaneously exhibited such faster mental rotation skills. Such minor variations may help those who set store by some notion of the 'natural' array of cognitive capacities found in human brains. Either way, the question remains: what are we to say about the various cases? While case (2) allows for computations to be distributed across agent and computer, case (3) displays (by hypothesis) an identical computational/informational structure, but this timer internalized within the agent. If the

latter case is (as it seems to be) a case of extended cognitive processing, by what right do we count case (2), using the external rotation button, as fundamentally different? We cannot simply point to the skull/skin boundary as justification, since the legitimacy of that boundary is precisely what is at issue. But nothing else seems different.

~~Moreover,~~ The kind of case just described is by no means as exotic as it may at first appear. It is not just the presence of advanced external computing resources which raises the issue, but rather the general tendency of human reasoners to lean heavily on environmental supports. Thus consider the use of pen and paper to perform long multiplication (see McClelland, Rumelhart, Smolesky & Hinton (1986), Clark (1989) Ch. 6), the use of physical re-arrangements of tiles to prompt word recall in Scrabble (see Kirsh (to appear)), the use of instruments such as the nautical slide rule (see Hutchins (1995)), and the general paraphernalia of language, books, diagrams and culture. In all these cases, the individual brain performs some operations (roughly, those based on associative recall and pattern recognition) while others (the ones involving re-combination of gross symbol structures, complex mathematics, the tracking of dependencies in logic etc.) are delegated to manipulations of external media. Had our brains been different, this distribution of tasks would doubtless have varied. Likewise, the jobs farmed out by the networked computer reflect its particular processing profile, whether temporary (it doesn't have a specific item of software loaded) or permanent (it lacks the memory needed to run a particular package).

In fact, even the mental rotation cases described in scenarios (1) and (2) are real. The cases reflect options available to players of the computer game Tetris! In Tetris, geometric figures (zoids) descend from the top of the screen and must be built into compact

walls as the game progresses. A falling zoid must be directed, rather rapidly, to an appropriate slot in the emerging structure. The physical rotation operation is available to players. Kirsh & Maglio (1994 p. 530) calculate that the physical rotation of a zoid through 90° takes about 100 milliseconds, plus about 200 milliseconds to select the rotate button. To achieve the same result by mental rotation takes about 1,000 milliseconds. Kirsh & Maglio go on to present compelling evidence that this physical rotation operation is used not just to position a zoid ready to fit a slot, but often to help identify the shape of a candidate zoid. The latter use constitutes a case of what Kirsh & Maglio call an 'epistemic action'. Epistemic actions alter the world so as to aid and augment cognitive processes such as recognition and search. Merely pragmatic actions, by contrast, alter the world because some physical change is desirable for its own sake (e.g. putting cement into a hole in a dam).

Epistemic action, I suggest, demands spread of epistemic credit. If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so I claim) part of the cognitive process. Cognitive processes ain't (all) in the head! (See also Rutkowska (1993), Clark (1989) Ch. 4, (1993) Ch. 6).

This conclusion may at first seem unpalatable. The reason, I think, is that many philosophers identify the cognitive with the conscious (Searle (1992) at least makes a virtue of this conflation and attempts to argue the case). And I assuredly do not seek to claim that individual consciousness itself extends outside the head. It seems clear, however, that not everything that occurs in the brain, and which constitutes a cognitive process in current scientific usage, is tied up with conscious processing. The Vestibula Ocular Reflex (VOR),

↳ many would happily say the VOR is not cognitive.
maybe find a better example, something more high level
– e.g. linguistic processing?

to take just one example from dozens, stabilizes the image of the world on the retina so as to offset head movement (see e.g. Churchland & Sejnowski (1994) p. 353-365). This operation is, of course, crucial for human vision. And human consciousness apprehends the world in a way which depends on the correct operation of the VOR. But the computational steps performed by the VOR circuitry do not figure among our conscious contents. If the computational transformations on which the VOR depends were in fact carried out by some external device (akin to the external high speed image rotator from case (2)), the interplay between conscious states and VOR computations could remain unaltered. So whatever role is played by the presence of consciousness (whatever exactly that means) somewhere in the loop, that role cannot itself afford grounds for rejecting the characterization of some external data transformations as part of our cognitive processing. Or rather, it could do so only if we bite the bullet and reject all processes which are not themselves consciously known -- i.e. if we reject both the external high-speed rotation and the VOR as cognitive processes.

One final worry deserves mention. It may be said that what keeps real cognitive processes in the head is some consideration of portability. That is to say, we are moved by a vision of what might be called the Naked Mind: a vision of the resources and operations we can always bring to bear on a cognitive task, regardless of whatever further opportunities the local environment may or may not afford us.

right. it's what stays constant across explanatory contexts.

I am quite sympathetic to this objection. It seems clear that the brain (or perhaps, on this view, the brain and body) is a proper and distinct object of study and interest. And what makes it such is, precisely, the fact that it comprises some such set of basic, portable

cognitive resources. These resources, as noted parenthetically above, may incorporate bodily actions as integral parts of some cognitive processes (as when we use our fingers to off-load working memory in the context of a tricky calculation). But they will not encompass the more contingent aspects of our external environment; the ones which may come and go, like the presence of a pocket calculator.

Nonetheless, I do not think that the portability consideration can bear sufficient conceptual weight, and for several reasons.

First, they run a risk of begging the question. If we ask why portability should matter to the constitution of cognitive processes, the only answer seems to be that we want our cognitive processes to come in a distinct, individually mobile package. But this, of course, is just to invoke the skin/skull boundary all over again -- and it is the legitimacy of this very boundary for defining cognitive processes which is in question.

weak!

no it's not - it's to motivate the sk/skull boundary.

Second, it would be easy (albeit a little tedious for the reader) to construct a variety of troublesome cases. What if some people always carried a pocket calculator; what if we one day have such devices implanted in our brains? What if we have body docks for a variety of such devices and 'dress' each day with add-on-devices appropriate for that day's prescribed problem-solving activity? Nor can the vulnerability of such devices to discrete damage or malfunction serve to distinguish them, for the biological brain, likewise, is at risk of losing specific problem-solving capacities through lesion or trauma.

sure, and to the extent that that were the case, a limited extension would be true.

Third, what if the biological brain has itself evolved in ways which factor in the reliable presence of a manipulable external environment? It seems rather plausible to suppose (see Clark (1989) Ch. 4) that evolution will favor on-board problem-solving

vague + a bit hard to picture what you mean - example?

- also, reliable presence of an environment ≠ reliable presence of specific environmental feature/object, (if feature is sometimes absent - constancy goes way down)

capacities which are especially geared to parasitizing the local environment so as to reduce memory load, and even to transform the nature of the computational problems themselves (see e.g. the growing body of work on the tricks and strategies of so-called Animate vision systems: systems which act so as to maximize the role of external structure in simplifying internal computation (Ballard (1991), Churchland, Ramachandran and Sejnowski (1994), Clark (to appear)).

All of which leads us, at last, to the methodological heart of the matter. To a first approximation it surely matters very little whether we choose to include external structures and manipulations as proper parts of cognitive processes, or merely as additional props and aids. It is, perhaps, a matter of some philosophical interest if (as I have claimed) there exists no good reason to restrict cognitive processes to the bounds of the body, and some good reasons to allow them to range more freely. But beyond a raised eyebrow or two, who would care? Methodologically, however, there is a real battle to be fought. For suppose that inner structure and processes are indeed exquisitely adapted (by evolution or lifetime learning) so as to distribute cognitive tasks, and to solve problems by closely coupled physical interactions with further resources. In that case, it makes no sense to study the brain (or brain/body) with as little regard to real environmental interactions as has been generally customary (with some notable exceptions -- see e.g. Gibson (1979), Thelen and Smith (1994), and the Animate vision literature mentioned earlier). This does not mean that the brain/body is not itself a legitimate target of interest, as conceded earlier. Rather it means that *if* that is your target, you will only understand its true nature by paying close attention to its properties as embedded in some ecologically realistic setting. Treating

cognitive processes as the kind of things which regularly criss-cross the body/environment boundary is, I believe, a crucial step in promoting such a practice.

What, finally, of that most problematic entity, the self? Does the putative spread of cognitive processes out into the world imply some correlative (and surely unsettling) leakage of the self into the local surroundings? The answer looks to be (sorry!) "Yes and No". No, because we already conceded that conscious contents supervene on individual brains. But Yes, because such conscious episodes are at best snapshots of the self considered as an evolving psychological profile. Thoughts, considered only as snapshots of our mental activity, are fully explained, I am willing to say, by the current state of the brain. But the flow of thoughts, the temporal evolution of ideas and attitudes, these are determined (and explained) by the intimate continued interplay between the individual brain and the body and world. In constructing this paper -- which is, I promise, a reflection of my attitudes and beliefs! -- I relied on constant interactions with external media (paper, pen, computers, books). These media allowed ideas to gel and rearrange and mutate in ways which simply would not have occurred sitting in a darkened room. The evolving flow of my thoughts and beliefs is thus not determined or explained simply by some confrontation between my on-board reasoning devices and a world of data and inputs. Instead, the world (at times) is itself the locus of operations and transformations which, *were they to occur in the head*, we would have no hesitation in calling cognitive processes and in directly associating with some specific psychological profile.

In sum, let the idea of self fall where it will. My concern has been only to argue that we have good reasons, both philosophical and methodological, to embrace a more liberal

notion of cognitive processes -- one which explicitly allows the spread of such processes across brain, body, world, and artifact. The time is ripe to usurp the hegemony of skin and skull, and to recognize the complex conspiracy that is adaptive success.

Philosophy/Neuroscience/Psychology Program
Department of Philosophy
Washington University
St. Louis, MO 63130
e-mail: andy@twinearth.wustl.edu

The overall point seems to be that ewt feature may be cognitive insofar as they are closely coupled with cognitive processes (affected by them, + affect them back). Maybe worth making the role of coupling explicit. Note that this distinguishes from the Putnam/H_2O case – where the relevant ewt. features (that constitute contact) will be distal + historical, i.e. uncoupled, + so much less obviously relevant.

But then this coupling point can help motivate the 'portability' objection, i.e.: as a matter of contingent fact, by far the most constant + reliable coupling takes place within the ~~skull~~ skin. To the extent that there was constant external coupling, some externalism would be true – but one might argue that given the actual facts about coupling, internalism isn't far off.

Note that to the extent that coupling is unconstant – e.g. calculators that come and go – we are tempted to describe the process as the interaction of a system with its environment (as after all, we need an expl. framework that also handles cases when the calculator isn't there). But to the extent that coupling is constant, we might describe the process as one big system, with no need for the 'factoring' strategy. And we could ~~say~~ more easily say the mind is in the world, without worries about personal identity and the like.

Incidentally, I would also note that these coupled systems can plausibly support intentional explanations that are essentially external. e.g., X went to this house because X thought John lived there, when actually John's address was written in X's pocketbook, rather than in memory – seems to be equally good explanation. So we have a form of intentional externalism much more interesting than the Putnam/Burge kind (again, because of coupling).

REFERENCES

Ballard, D. (1991). Animate Vision. *Artificial Intelligence* , *48* , 57-86.

Churchland, P., Ramachandran, V., & Sejnowski, T. (1994). A Critique of Pure Vision. In C. Koch & J. Davis (Ed.), *Large-Scale Neuronal Theories of the Brain* . Cambridge, MA: MIT Press.

Churchland, P., & Sejnowski, T. (1992). *The Computational Brain* . Cambridge, MA: MIT Press.

Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing* . Cambridge: MIT Press.

Clark, A. (to appear). Moving Minds: Situating Content in the Service of Real-Time Success. In J. Tomberlin (Ed.), *Philosophical Perspectives* (Vol. 10,).

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception* . Boston: Houghton-Mifflin.

Hutchins, E. (1995). *Cognition in the Wild* . Cambridge, MA: MIT Press.

Kirsh, D. & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science, 18*, 513-549.

Malone, T., Fikes, R., Grant, K., & Howard, M. (1988). Enterprise: A market-like task scheduler for distributed computing environments. In B. Huberman (Ed.), *The Ecology of Computation* . North-Holland.

McClelland, J., Rumelhart, D., & Hinton, G. (1986). The appeal of Parallel Distributed Processing. In J. McClelland, D. Rumelhart, & P. R. Group (Ed.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. II, pp. 3-44).

Cambridge: MIT Press/Bradford Books.

Putnam, H. (1975). The meaning of "meaning". In H. Putnam (Ed.), *Mind, Language and Reality* (pp. 215-275). Cambridge: Cambridge University Press.

Rutkowska, J. (1993). *The Computational Infant* . New York: Harvester Wheatsheat.

Searle, J. (1992). *The Rediscovery of the Mind* . Cambridge, MA: MIT Press.

Thelen, E., & Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action* . Cambridge, MA: MIT Press.

You make grand claims in terms of "mind", but at the level of detailed argument it is always "cognitive processes" — rhetorical strategy? maybe would be harder to make case directly with "mind"? can you find an intuitive example of a mental state (perhaps unconscious) that's external? (e.g. a belief, c.f. the address in the pocketbook). that would help, though I imagine there would be a few strategies of resistance. In any case doing the dirt-work with "cognitive process" makes things too easy for you.

—

Maybe ext-supported intentional states are like tacit beliefs — you have to do some work to get them.
(One intuitive difference in these ext-coupling cases is that the coupling is usually low-bandwidth compared to the brain. Which raises the question — how do we think of low bandwidth coupling within the brain? Maybe memory, tacit belief, etc, is like this.)

—

— Also, you need a nifty name for your position
— "coupled externalism"? or "The Extended Mind: ..."
or something along those lines.

—

How about a nice Twin case somewhere
— e.g. the guy with "220 Baker St" in his pocketbook
internal + the guy with "22 Baker St".
they're twins, but do different things (go to different addresses) because of differences in what they're coupled with. → different cognitive explanations.
You might even argue "A believes Sherlock lives at 220"
and "B believes Sherlock lives at 22"
— at least this would seem to function in belief/desire explanation in precisely the usual way. (I could make analogy with memory, etc).

—

I wonder what you'd say about distributed cognition — e.g. between workers in a team. Is this a single "group mind"? If the previous arguments work, this same sort of argument ought to go through in this case.