Functionalism and Extended Cognition

Prof. Dr. Sven Walter

Institute of Cognitive Science
University of Osnabrueck
Germany
s.walter@philosophy-online.de

11th Symposium »Contemporary Philosophical Issues«, University of Rijeka, May 30, 2009

Outline

i. the *Hypothesis of Extended Cognition* (HEC)

ii. three possible arguments for HEC
   - a general account of cognition
   - the Parity Principle
   - functionalism

iii. troubles with functionalism
   - multiple realizability (Adams and Aizawa)
   - fine- and coarse grained roles (Rupert, Rowlands)
   - the Martian Intuition (Sprevak)

the Hypothesis of Extended Cognition

"I sometimes find, and I am sure you know the feeling, that I simply have too many thoughts and memories crammed into my mind. … At these times … I use the Pensieve. One simply siphons the excess thoughts from one's mind, pours them into the basin, and examines them at one's leisure."

the Hypothesis of Extended Cognition

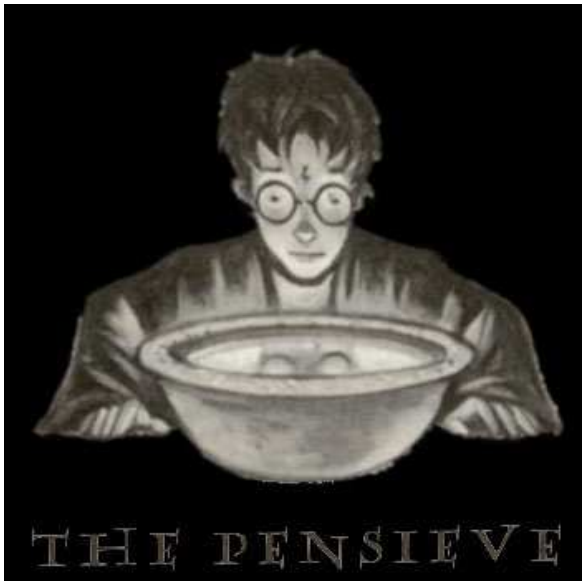
"I sometimes find, and I am sure you know the feeling, that I simply have too many thoughts and memories crammed into my mind. … At these times … I use the Pensieve. One simply siphons the excess thoughts from one's mind, pours them into the basin, and examines them at one's leisure."

Albus Dumbeldore, *Harry Potter and the Goblet of Fire*

the Hypothesis of Extended Cognition

"I sometimes find, and I am sure you know the feeling, that I simply have too many thoughts and memories crammed into my mind. … At these times … I use the Pensieve. One simply siphons the excess thoughts from one's mind, pours them into the basin, and examines them at one's leisure."

Albus Dumbeldore, *Harry Potter and the Goblet of Fire*



the Hypothesis of Extended Cognition

the central claim behind HEC comes under a variety of different labels:

extended mind (Clark and Chalmers 1998)

active externalism (Clark and Chalmers 1998)

vehicle externalism (Hurley 1998; Rowlands 2003)

environmentalism (Rowlands 1999)

locational externalism (Wilson 2004)

the Hypothesis of Extended Cognition

the common idea is that

> "there are conditions under which thinking and thoughts (or more precisely, the material vehicles that realize thinking and thoughts) are spatially distributed over brain, body and world, in such a way that the external (beyond-the-skin) factors concerned are rightly accorded fully-paid-up cognitive status" (Wheeler *forthcoming*, p. 1)

the Hypothesis of Extended Cognition

Consider the case of **Inga and Otto** (Clark and Chalmers 1998):

According to C&C, the state of Otto's notebook interacts with Otto's desires and other beliefs in a way similar to the way in which Inga's biomemory interacts with her desires and other beliefs.

E.g., exposure to new information causes Otto to modify the state of his notebook and Inga to modify her biomemory.

the Hypothesis of Extended Cognition

Moreover, the current state of Otto's notebook causes Otto to stop at 53rd St., and the current state of Inga's biomemory causes Inga to stop at 53rd St.

The **functional role** of the stored information—its "functional poise" (Clark 2007, 2008)—appears to be the same in both cases.

Hence, C&C conclude, just as Inga has a belief that MOMA is on 53rd St., so Otto has a belief, with the same content, that extends partially into the environment, viz., the notebook.

possible arguments for HEC (I)

The most convincing argument for HEC, it seems to me, would proceed by first giving a general account of **cognition** (an account of what it is for a process to be a **cognitive process**, or of what it is for a system to be a **cognitive system**), and then go and see which processes in the world fulfill these conditions – are they restricted to biological organisms, or subsystems thereof, or do they include extrabodily items?

possible arguments for HEC (I)

Problem

We have not the slightest idea what such an account of cognition could plausibly look like. There is just no received view on what it is for a process to be a cognitive process.

- Walter, S. and Kyselo, M., "Review of *The Bounds of Cognition*," *Erkenntnis*, forthcoming.
- Walter, S. and Kästner, L., "What Does 'Cognition' Mean in 'Extended Cognition'?", in preparation.

possible arguments for HEC (II)

Parity Principle:

> "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 we claim) part of the cognitive process." (Clark and Chalmers 1998, p. 8)

possible arguments for HEC (II)

The PP enables C&C to argue that if two processes are just like one another, except for one being internal and the other extended, then both have an equal right to be cognitive.

The PP thus guarantees **equal treatment** between internal and external cases.

Don't discriminate against the external!

possible arguments for HEC (II)

Problem

An "equal treatment principle" cannot, all by itself, support HEC because it is silent about the exact conditions under which we would grant that an extrabodily process is a process which, were it to go on in the head, we would unhesitatingly call "cognitive" and about whether these conditions are ever fulfilled.

- Kyselo, M. and Walter, S., "Review of *Supersizing the Mind*," *Philosophical Psychology*, forthcoming.

possible arguments for HEC (III)

What does the crucial work for C&C and others is an (explicit or implicit) commitment to **functionalism**.

What the PP says is that if an external process contributes in the same way to the overall system, i.e., **plays the same causal role**, as does a purely internal process (cf. Otto and Inga), then the mere fact that the former is external while the latter is internal should not make us conclude that the former is not a cognitive process, if the latter is.

HEC may thus be regarded as a mere footnote to Putnam (Wheeler 2008)

troubles with functionalism (I)

Fred Adams and Ken Aizawa (2008) argue that we should expect processes as distinctive as cognitive processes to be realized by correspondingly distinctive lower-level processes.

"Roughly speaking, lower-level processes should be as distinctive as the higher-level processes they realize" (Adams and Aizawa 2008, p. 68).

troubles with functionalism (I)

E.g., prior to the retina we find optical processes essentially similar to those present in non-organic optical machinery, but once the light enters the retina, there is a shift to distinctive molecular processes that, among other things, result in the color-sensitive, orientation-sensitive and motion-sensitive selective release of neurotransmitters, **and that can only be found there.**

According to A&A, this transition in lower-level processes also marks a transition from the non-cognitive to the cognitive.

troubles with functionalism (I)

Problem (I)

According to Wheeler (forthcoming, pp. 5-6), A&A's argument amounts just to the denial that there is something like **multiple realizability**, and he cites evidence from **functional convergence in evolution** designed to show that there actually is MR.

E.g., the enzymes that in humans and many other animals break down alcohols display no sequence similarity with each other, have fundamentally different tertiary structures, and catalyze alcohol into acetaldehyde using different chemical reactions.

troubles with functionalism (I)

A Problem with Problem (I)

This is true, but misses the point.

i.) A&A do not deny that cognitive processes are MR by different neurophysiological processes, only that they are realized by non-neural processes, and this is compatible with Wheeler's counterexample.

ii.) Even if Wheeler's counterexample worked, it would still be open to A&A to claim that there are no similar cases of convergent evolution in the cognitive domain (e.g., memory consolidation).

troubles with functionalism (I)

Problem (II)

A better response would be, I think, to point to actual cases of MR, where some realizers are clearly neurophysiological whereas others are equally clearly not neurophysiological, e.g., to **successful case studies in cognitive science** (artificial vision, learning neural nets etc.) or to **successful cases of neuroimplantation**.

troubles with functionalism (II)

Rupert (2004), Weiskopf (2008) and A&A (2001, 2008) have argued that there exist significant **functional differences** between human beliefs and memories on the one and Otto's allegedly extended "beliefs" and "memories" on the other hand that make any functional equivalence demanded by the PP impossible.

troubles with functionalism (II)

"Beliefs are, as I will say, normally *informationally integrated* with, and updated in concert with, other beliefs ... But most externally located mental states do not share this feature. So ... they cannot be beliefs" (Weiskopf 2008, p. 268).

When a believer acquires a new belief properly so called, her belief system is "automatically and unconsciously updated to reflect this new information" (Weiskopf 2008, p. 269), but not when she acquires an externally stored "belief," say a new entry in her notebook.

troubles with functionalism (II)

Likewise, Rupert and A&A argue with regard to memory that human memory is subject to effects (e.g., negative transfer, generation effects, etc.) not exhibited by Otto's alleged "memory". Hence, they argue, there is, again, no functional equivalence.

troubles with functionalism (II)

Problem (I)

One strategy would be to deny that human cognitive capacities actually have the characteristics claimed by the critics of HEC.

With regard to Weiskopf's claim that belief integration is typically rapid, automatic, and unconscious, this seems to be a plausible response.

- Walter, S. and Kyselo, M., "Belief Integration in Action: A Defense of Extended Beliefs," *Philosophical Psychology*, forthcoming.

troubles with functionalism (II)

Problem (II)

Another strategy would be to argue that the differences highlighted by the critics of HEC are found only at a **fine-grained level of functional analysis** and that it begs the question against HEC to let the actual neurophysiological/psychological details of human cognitive capacities set the standard for the "functional poise".

Instead, one might say, one has to adopt a rather more liberal characterization of the relevant functional role that is not "too chauvinistic".

troubles with functionalism (II)

A Problem with Problem (II)

"If Rupert's arguments against the extended mind are question-begging because they presuppose a chauvinistic form of functionalism, it is difficult to see why arguments for the extended mind are not question-begging given their predication on a liberal form of functionalism. Adjudicating between the extended mind and its critics, therefore, seems to require adjudicating between liberal and chauvinistic forms of functionalism. But this is a dispute that has been ongoing almost since functionalism's inception. In the absence of any satisfactory resolution of this dispute, the clear danger for the extended mind is one of stalemate." (Rowlands manuscript, pp. 6-7)

troubles with functionalism (II)

Response

I think the stalemate can easily be broken (see also Wheeler 2008, forthcoming): What would happen if we were to meet a human being (or, for that matter, a Martian) whose beliefs or memories did *not* exhibit the characteristics in question?

troubles with functionalism (III)

Part of the reason why my response to the previous problem is plausible is that we should respect what Sprevak (forthcoming) calls the "**Martian Intuition**":

Fine-grained differences should not prevent us from attributing cognitive states to Martians whose internal make-up, both materially and psychologically, is different from ours.

troubles with functionalism (III)

"The Martian intuition applies to fine-grained psychology as well as physiology: there is no reason why a Martian should have exactly the same fine-grained psychology as ours. A Martian's pain response may not decay in exactly the same way as ours; its learning profiles and reaction times may not exactly match ours; the typical causes and effects of its mental states may not be exactly the same as ours; even the largescale functional relationships between the Martian's cognitive systems (e.g. between its memory and perception) may not exactly match ours." (Sprevak forthcoming, pp. 5-6)

troubles with functionalism (III)

Appealing to the Martian Intuition, however, Sprevak argues, may solve one problem, but causes even more serious trouble for HEC.

According to Sprevak, "if the grain parameter is set at least coarse enough to allow for intelligent Martians, then it also allows many cases of extended cognition" (p. 8), and for cases which are so obviously implausible that they amount to a *reductio* of HEC.

troubles with functionalism (III)

Sprevak's strategy is the following: if we take some putative case of extended cognition, we can always imagine a functionally equivalent system that is located entirely inside the head of a Martian. On the strength of the Martian intuition, we would count the Martian-internal system as cognitive. So when, as functionalists, we fix the level of grain for our analysis, it must be set coarsely enough to generate that result. But if it is that coarse, then the (by hypothesis) functionally identical extended system too will count as cognitive.

troubles with functionalism (III)

For illustration: if I have a desktop computer which contains a program for calculating the dates of the Mayan calendar 5,000 years into the future, then, according to HEC, I possess an extended cognitive process that is capable of calculating the dates of the Mayan calendar.

The reason is that one could imagine a Martian with an internal process that is capable of calculating the dates of the Mayan calendar using the same algorithm as my desktop computer.

troubles with functionalism (III)

Problem (I)

Wheeler (forthcoming, pp. 19-20) argues that Sprevak's objection fails because it starts with a process which is clearly not cognitive (the program on my desktop computer), and then argues that a functionally equivalent process (the hypothetical one in the Martian) is cognitive, where the only difference is one in location – which would violate the PP.

troubles with functionalism (III)

A Problem with Problem (I)

Wheeler seems to get things backwards. Sprevak does not start with the non-cognitive desktop computer process and then concludes that the Martian process must be cognitive because it is in the head.

Rather, Sprevak begins with the claim that the Martian process is cognitive (because of the Martian Intuition) and then argues that the desktop computer must, given PP, be cognitive, too.

troubles with functionalism (III)

Problem (II)

I think Wheeler is right to claim that the advocate of HEC should deny that the Martian process is cognitive.

Yet, one what grounds should we do this? Sprevak argues that any reason must necessarily be arbitrary and an *ad hoc* maneuver designed only to save HEC.

troubles with functionalism (III)

Problem (II)

However, it need not be. Functionalism does not mean that anything goes.

Block's Chinese Nation Head, e.g., is rightly considered not to be a cognitive system, and not only because we want to save HEC.

This may mean that we need an independently plausible account of the cognitive, after all ...

to be continued ...

thanks to my collaborators

Miriam Kyselo

(cognitive science, Osnabrueck)

Holger Lyre

(philosophy, Bielefeld)

Lena Kästner

(cognitive science, Osnabrueck)