The Extended Mind Thesis and Mechanistic Explanations

Peter Fazekas

Institute of Philosophy, Research Centre for the Humanities
Hungarian Academy of Sciences
email: fazekas.peter@gmail.com

Abstract

The Extended Mind Thesis (EMT) is traditionally formulated against the bedrock of functionalism, and ongoing debates are typically bogged down with questions concerning the exact relationship between EMT and different versions of functionalism. In this paper, I offer a novel ally for EMT: the new mechanistic approach to explanation. I argue that the mechanistic framework provides useful resources not just to disambiguate EMT, and to show which objections fail to pose a serious challenge, but also to answer some of the deeper problems that stem from the functionalist roots of EMT.

1. Introduction

The Extended Mind Thesis (EMT), first proposed by Clark and Chalmers (1998), claims that the vehicles of our cognitive processes and mental states can and sometimes do extend beyond the boundaries of our brain and even body (cf. Clark, 2008, p. 76). If a physical object plays a part in the machinery realising a cognitive process, then that physical object *is* part of the cognitive process, no matter whether it is inside or outside the head. This is the so-called parity principle (Clark & Chalmers, 1998, p. 8), the heart and soul of EMT.

The most well-known illustration is the case of Otto and Inga. When Inga hears about a new exhibition at the Museum of Modern Art (MoMA) she retrieves from her biological memory that it is on 53rd Street and goes to the museum. Her counterpart, Otto, who suffers from a mild form of Alzheimer's disease, uses a notebook to jot down and retrieve information. So when he hears about the exhibition he turns to his notebook, and retrieves the address from there. Clark and Chalmers argue that, just like Inga, Otto walks to the museum because "he *wanted* to go to the museum and *believed (even before consulting his notebook)* that it was on 53rd Street" (Clark, 2008, p. 78). Both Otto and Inga have the same long-term beliefs, it is just Otto's beliefs are extended beyond his skull.

2. Objections to EMT

EMT has been widely criticised in the literature. In what follows, I am going to focus on some of the most characteristic objections: a set of related issues raised by Adams and Aizawa (2001, 2009, 2010a, 2010b), and an argument put forth by Mark Sprevak (2009).

First, Adams and Aizawa happily acknowledge that Otto's notebook is causally coupled to the cognitive processes of Otto, but resist the further claim that the notebook is part of the relevant cognitive process. Clark and Chalmers, they argue, commit a *coupling-constitution fallacy*: they confuse causal interactions between Otto's cognitive processes and the notebook with an extension of a cognitive process (Adams & Aizawa, 2010b, p. 68).

Clark and Chalmers support the constitution claim by emphasising that Otto's notebook plays "roughly" (Clark, 2008, p. 76) the same functional role as Inga's biological memory: their content are poised for the control of action, and interact with desires and other beliefs in roughly the same way. Adams and Aizawa agree. However, even if there is rough functional similarity, they argue, i.e. even if the coarse-grained functional profile of Otto's notebook matches that of Inga's biological memory, their fine-grained functional profiles are so vastly different (e.g. biological memories are dynamic systems prone to reorganisation, interpolation, creative mergers, etc., whereas notebooks are static storage devices) that Otto's notebook does not deserve to be called cognitive at all.

For Adams and Aizawa, fine-grained functional descriptions are important, because they believe that "the cognitive must be discriminated on the basis of underlying causal processes" (Adams & Aizawa, 2001, p. 52). Only uncovering these underlying fine-grained causal regularities yields interesting causal laws. From this perspective, EMT, by relying on coarse-grained functional characterisations, and thus permitting nearly anything into the scope of cognitive science (e.g. a range of tools that humans can use as mnemonic aids), looses every chance to be able to uncover interesting regularities (Adams & Aizawa, 2001, p. 61). That is, cognitive science allowing for extended cognitive processes just *ceases to be a science*.

Relatedly, Sprevak (2009) draws attention to that EMT faces a dilemma: in order to allow for cases of extended cognition EMT needs to set the grain parameter (at which two functionally characterised processes are identical) coarse enough, but if the grain parameter is too coarse, it yields a radical version of EMT by permitting cases of extended cognition which even proponents of EMT would not allow for.

Clark and Chalmers realise that extra conditions are needed in order to make EMT more modest and plausible. They offer the following "rough-and-ready set of additional criteria to be met by nonbiological candidates for inclusion into an individual's cognitive system": (1) "That the resource be reliably available and typically invoked.", (2) "That any information thus retrieved be more or less automatically endorsed [...] and deemed trustworthy", (3) "That information contained in the resource should be easily accessible as and when required.", and (4) "That the information in the notebook has been consciously endorsed at some point in the past and indeed is there as a consequence of this endorsement." (Clark, 2008, p. 79; see also Clark & Chalmers, 1998, p. 17).

For Sprevak, this set of criteria is problematic because the conditions are unjustified: Clark and Chalmers remain silent about why this set is to be accepted. Sprevak argues that without

providing well-established conditions for designating an intermediate level grain size, which proponents of EMT fail to do, EMT will not be able to balance between the two extremes: *radical-EMT*, where (at too coarse grain size) nearly any artefact could be seen as an extension of certain cognitive processes, and *no-EMT*, where (at too fine grain size) no artefacts would qualify as constitutive parts of cognitive processes.

3. The New Mechanistic Framework

Now that the stage is set, I turn my attention to demonstrating how the mechanistic framework (Bechtel, 2008; Craver, 2007; Machamer, et al., 2000) is able to help articulating the core idea behind EMT, and formulating systematic answers to the main objections.

The mechanistic approach identifies a phenomenon via identifying the causal roles played by a system producing the phenomenon, and accounts for this phenomenon in terms of the organised activity of the parts of the system. It is a multilevel approach—once the activity of an entity at a given level is decomposed into the lower level organised structure of its parts and their activities it is possible to apply the same methodology again in order to account for the lower level activities in terms of still lower levels, and so on.

If we apply this framework to the case of Inga and Otto, what we get is the following. Inga hears about an exhibition in MoMA and consequently she goes to the museum. This is the personal level phenomenon one can observe. Now in line with the mechanistic approach, it is possible to account for this phenomenon in terms of the organised activity of lower—arguably sub-personal—level entities and activities: e.g. a speech recognition process analysing the auditory input, an act of mental recall searching Inga's biological memory, an act of deliberate planning, and the activity of a motor control system moving Inga's body-parts. One can, of course, dig deeper, to find out what still lower level entities and activities produce the activities of these now middle-level entities. For example, at a lower level one can describe neural mechanisms responsible for how Inga's biological memory responds to recall processes and affects planning, and also for internal processes like consolidation, re-organisation, interpolation, and so on.

Similarly, when Otto hears about the exhibition in MoMA, he also goes to the right place. So the phenomenon as described at this personal level is the same. The underlying mechanism, however is different. Unlike in Inga's case, where at the sub-personal level one can find an act of mental recall searching biological memory, in Otto's case the set of sub-personal level entities and activities constituting the higher level phenomenon partly consists of an act of reaching for the notebook and reading from it. If one digs deeper, it turns out that Otto's notebook is quite unlike Inga's biological memory in the sense that it has no parts the activities of which would be responsible for such dynamic internal processes that are characteristic of biological memories. The notebook is a static storage device.

4. EMT Disambiguated

Distinguishing between these different levels helps systematising the problem space. Mental states typically get ascribed on the basis of personal (here: highest) level descriptions. Literally, neither

the sub-personal process of recalling from biological memory, nor reaching and reading from a physical notebook 'has a belief'—or 'is cognitive' for that matter (cf. Adams & Aizawa, 2010b for claiming otherwise)—mainly because the terms 'having a belief' and 'being cognitive' belong to a vocabulary characterising the personal level, but not the lower levels. Sub-personal level entities and their activities contribute to 'having believes' or 'being cognitive' by *constituting* a mechanism that, as a whole, is responsible for these phenomena.

EMT focuses on these sub-personal (here: middle) level mechanisms. EMT does *not* commit the coupling-constitution fallacy because what EMT is really interested in are exactly those cases where a physical object does play a part in a sub-personal mechanism. According to the mechanistic framework, mere coupling happens when an entity is causally connected to parts of a given mechanism, but it does not contribute to the overall performance of the mechanism. Contrary to this, constitution happens when the entity in question, and its activities, do contribute to the overall performance of the mechanism. EMT here can even rely on the specific method the mechanistic framework provides for deciding whether a given entity is part of a mechanism: the method of mutual manipulability—the idea that lower level intervention and higher level detection combined with higher level intervention and lower level detection is able to locate mechanism boundaries (Craver, 2007, pp. 152-160).

Interestingly, the top-down manipulability constraint readily disqualifies many of the famous counterexamples (e.g. me having standing beliefs about whatever I could find in Encyclopaedia Britannica, or calculate with a supercomputer, etc.), in which cases wiggling the whole mechanism would not necessarily wiggle the content or state of the artefact (cf. Craver, 2007, p. 153) due to lack of true read-write integration (as in the case of the Encyclopaedia Britannica), or constant and reliable access (as in the case of a supercomputer).

Note that some of the extra conditions Clark and Chalmers impose on EMT aim at disqualifying exactly these cases—e.g. (4) argues for true read-write integration, whereas (1) for constant and reliable access.

Other conditions—e.g. (2) arguing for automatic endorsement and trustworthiness, or Clark's returning emphasis that what matters from the point of view of EMT is "the achieved functional poise of the stored information" (cf. e.g. Clark, 2008, p. 88)—characterise the 'forward looking causal profile' (cf. Shoemaker, 2007) of the entities within the mechanism, i.e. how they affect other parts of the mechanism (e.g. how the notebook affects planning and ultimately motor control). Without these latter conditions, bottom-up style manipulability would fail, since if these relations were not in place (e.g. if the information stored was not trusted) wiggling the part would not necessarily wiggle the whole.

That is, the mechanistic framework helps us recognise that—contra Sprevak's (2009) charge,—the extra set of criteria proposed in order to avoid radical EMT is in fact well-founded: it is an alternative formulation of mutual manipulability, i.e. Clark and Chalmers' way of getting a grip on the constitution relation.

Next, note that EMT does not need to be concerned with the levels below the highest sub-personal (here: middle) level. The details about the entities and their activities at lower levels are

irrelevant from the perspective of the mechanism EMT focuses on. The only thing that matters is whether entities at this (middle) level contribute to a mechanism that as a whole produces the cognitive phenomenon in question. That is, Otto's notebook extends Otto's cognitive processes as far as it partly constitutes a mechanism that produces the same cognitive phenomenon that is also produced by the mechanism partly constituted by Inga's biological memory. The fine-grained functional profile of Otto's notebook does not need to match that of Inga's biological memory in order to have that.

Abstracting away from features of lower level implementation, contrary to what opponents of EMT claim, does not render EMT unscientific. First, EMT does not permit nearly anything into the scope of cognitive science—only those entities and activities that constitute the mechanisms EMT is interested in. Second, EMT does yield interesting causal regularities at this level (even if not at the levels below).

This is the very message of EMT: there are interesting causal regularities one can uncover when one investigates how the organised activity of (high level) sub-personal entities produces personal level phenomena. These regularities are insensitive to differences in lower level implementation, and most importantly, they sometimes span through the internal-external divide.

5. Conclusion

This paper argues that the mechanistic framework provides useful resources for proponents of EMT. Besides helping to articulate the core idea behind EMT, and revealing why objections based on the coupling-constitution fallacy style arguments trivially fail to pose a serious challenge for EMT, the mechanistic framework also offers a natural solution to the problem of how to set the grain size of the functional descriptions EMT relies on, and provides reasons for accepting the extra criteria that are typically imposed on EMT by its proponents as necessary amendments to general functionalism.

Literature

Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology, 14*(1), 43-64.

Adams, F., & Aizawa, K. (2009). Why the Mind is Still in the Head. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 78-95): Cambridge University Press.

Adams, F., & Aizawa, K. (2010a). *The Bounds of Cognition*. Oxford: Wiley-Blackwell.

Adams, F., & Aizawa, K. (2010b). Defending the Bounds of Cognition. In R. Menary (Ed.), *Extended Mind* (pp. 67-80): MIT Press.

Bechtel, W. (2008). *Mental mechanisms: philosophical perspectives on cognitive neuroscience*. New York: Lawrence Erlbaum Associates.

Clark, A. (2008). *Supersizing the Mind*. Oxford: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis, 58*(1), 7-19.

Craver, C. (2007). *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.

Sprevak, M. (2009). Extended Cognition and Functionalism. *The Journal of Philosophy, 106*(9), 503-527.

[word count: 2478]