The overextended mind

Shaun Gallagher
Philosophy and Cognitive Sciences
Institute of Simulation and Training
University of Central Florida (USA)
University of Hertfordshire (UK)
gallaghr@mail.ucf.edu

Clark and Chalmers (1998) introduced the concept of the extended mind, in part to move beyond the standard Cartesian idea that cognition is something that happens in a private mental space, “in the head.” In this paper I want to pursue a liberal interpretation of this idea, extending the mind to include processes that occur within social and cultural institutions. At the same time I want to address some concerns that have been raised about whether such processes actually constitute cognitive processes, so that we think of the mind as extended, or whether environmental and institutional factors are simply causal supports for cognition, which, essentially, continues to occur only in the head.\(^1\)

I start with what Clark and Chalmers have called the parity principle.

If, as we confront some task, a part of the world functions as a process which, *were it to go on 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)

On a strict interpretation this principle appears to measure cognition in terms of the Cartesian gold standard of what goes on in the head. It suggests that a process outside of the head counts as cognitive only if in principle it could be accomplished in the head (or at least imagined to be so) – conforming to the (minimal) Cartesian concept of mental process as

---

\(^1\) Different versions of this paper were presented in 2009 at the Concluding Meeting of the European Science Foundation supported Consciousness in Natural and Cultural Contexts (CNCC), University of Edinburgh, and the conference on Embodied Virtues and Expertise at the University of Wollongong, Australia. I thank participants at those conferences for their comments, and the ESF-CNCC-BASIC project for funding. In this paper I draw on some material found in Gallagher and Crisafi (2009) and Crisafi and Gallagher (2009). In those papers we made reference to Hegel’s idea of objective spirit and the externalization of the mind in social institutions. The idea can be explicated without reference to Hegel, and here I make no use of his philosophy.
something that would normally happen in the head. Thus, we might think of some mental processes as happening “out there” in the world, yet still have a principled reason to limit mental processes to the kinds of things that fit a standard although slightly extended model.

Clark (2008, p. 114) rejects this interpretation, insisting that the parity principle should not be interpreted as requiring any similarity between inner and outer processes. Wheeler (2006, 3) explains that the parity principle does not “fix the benchmarks for what it is to count as a proper part of a cognitive system by identifying all the details of the causal contribution made by (say) the brain [and then by looking] to see if any external elements meet those benchmarks.” This reading is consistent with the functionalist account that both Clark and Wheeler embrace. The worry that comes along with this liberal interpretation is that the concept of mind gets overextended to include any process in the world (the “cognitive bloat” worry). Thus, even as he allows for the liberal interpretation of the parity principle, Clark starts to tighten it up again with a set of additional criteria that need to be met by external physical processes if they are to be included as part of an individual’s cognitive process. He lists three criteria.

1. That the external resource be reliably available and typically invoked.

2. That any information thus retrieved be more-or-less automatically endorsed. It should not usually be subject to critical scrutiny (unlike the opinions of other people, for example). It should be deemed about as trustworthy as something retrieved clearly from biological memory.

3. That information contained in the resource should be easily accessible as and when required. (Clark 2008, 79)

The parity principle plus these criteria rule over the primary and much discussed example of extended cognition provided by Clark and Chalmers: the example of Otto and Inga.

First, consider a normal case of belief embedded in memory. [Inga wants to go to MOMA, and remembers where it is]…. It seems clear that Inga believes that the museum is on 53rd Street, and that she believed this even before she consulted her memory. It was not previously an occurrence belief, but then neither are most of our beliefs. The belief was somewhere in memory, waiting to be accessed. … Now consider Otto. Otto suffers from Alzheimer's disease, and like many Alzheimer’s patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him
everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults the notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum. (Clark and Chalmers 1998, 12-13)

The notebook, for Otto, clearly plays the same role that memory plays for Inga. The belief, in Otto’s case, supervenes on processes that lie “beyond the skin” when in fact Otto engages with those non-neural processes. Alzheimer’s disease aside, we all may use similar memory enhancing tricks as Otto. If we have poor memories for directions or addresses, then with paper and pencil we use technology to do something that we could do, with a little more effort, in our head. Or perhaps we enhance our cognitive performance with other technology. Imagine looking up directions on Google, downloading them into your phone, and then using your phone to find your way. We seemingly are able to store our memories, and activate beliefs about where things are located, using such instruments. I can’t remember where the restaurant is, but I plus my technology can.

One of the problems with this, and similar examples, as I see it, is that it starts with a concept of the mind that the extended mind hypothesis is really trying to challenge. Specifically, the three criteria are worked out in terms of how we might deal with beliefs or “beliefs embedded in memory.” Clark (2008) then generalizes the criteria to apply to all cognitive processes. The controlling conception of the mind that guides the application of the three criteria, then, is that the mind is constituted by beliefs, desires, and other propositional attitudes, as well as representations, informational states – where the term “resources” is taken to signify just such phenomena. But these criteria do not necessarily apply to all cognition, especially if one thinks of cognition in terms of enactive cognitive processes and activities, e.g., problem solving, interpreting, judging, rather than in terms of mental states, or contents, e.g., beliefs, information. In what follows I want to argue for a liberal extension of the extended mind hypothesis, and I do so by first introducing an example that views cognitive processes as more like solving problems than like holding a belief.

**Mental institutions**

Certain social institutions (including social practices) are what we might call ‘mental institutions’ (Gallagher and Crisafi 2009), in the sense that they are institutions that help us to accomplish certain cognitive processes. Indeed, without them, specific classes of cognitive processes would simply not exist. They are at least enabling conditions, and on the most liberal reading,
constitutive of those processes.\(^2\) Examples include things like legal systems, educational systems, cultural institutions, like museums, and even the institution of science itself. In each case a mental institution

1. includes cognitive practices that are produced in specific times and places, and
2. is activated in ways that extend our cognitive processes when we engage with them (that is, when we interact with, or are coupled to these systems in the right way).

Here I'll consider the legal system as a good example. One way to think of the institution of law is to begin by thinking of people making claims on property, the appropriation and use of which immediately puts us in certain kinds of relations to others, relations which grow in complexity, and necessitate the use of contracts. A contract (as a legal agreement, not just the piece of paper) is in some real sense an expression of several minds externalized and extended into the world, instantiating in external memory an agreed-upon decision, adding to a system of rights and laws that transcend the particularities of any individual’s mind. Contracts are institutions that embody conceptual schemas that, in turn, contribute to and shape our cognitive processes. As such they are used as tools to accomplish certain aims, to reinforce certain behaviors, and to solve certain problems. Institutions of property, contract, rights, and law not only guide our thinking about social arrangements, for example, or about what we can and cannot do, but allow us to think in ways that were not possible without such institutions. Insofar as we cognitively engage with such tools and institutions we extend and transform our cognitive processes.

The legal system is constructed in part in these cognitive processes. The use of the legal system in the administration of justice, or the application of law to particular cases, are cognitive processes – a cognitive event through and through. These are not processes, however, that happen simply in the individual brains of judge, jury, defense attorney, prosecutor, etc. Of course we usually think of judgments as happening in the privacy of one’s own head. But some judgments supervene on external practices and processes that allow manipulation of a large amount of empirical information. In a court of law, for example, evidence and testimony are produced, and judgments are made following a set of rules that are established by the system. The process in which the judgments get made will depend on a body of law, the relevant parts of which come to the fore because of the precise particulars of the case, and as we remain cognitively engaged as the proceedings develop.

Consider an example that involves three different scenarios.

1. Alexis is given a set of facts and is presented a collection of evidence and is asked to judge the legitimacy of a certain claim that is being made on the basis of her own subjective sense of fairness. To make her judgment Alexis must weigh the facts and consider the evidence entirely in her own head, without help or interference from others. In this process she draws up and considers three questions about the facts, tries to answer them the best she can, and in that way forms her judgment.

2. Alexis is given a set of facts and is presented a collection of evidence and is asked to judge the legitimacy of a certain claim that is being made. This time, however, she is given the three questions by a group of experts who provide a set of possible answers from which she may choose. She still has to figure out what principles to use in answering the questions and forming her judgment. She considers the three questions, tries to answer them the best she can, and in that way forms her judgment.

3. Alexis is given a set of facts and is presented a collection of evidence and is asked to judge the legitimacy of a certain claim that is being made. As in (2), she is asked to consider the same three questions from a group of experts who again provide a set of possible answers from which she may choose. This time they also provide a set of pre-established rules she must follow in answering the questions.

How much cognitive processing, or let’s say cognitive effort, is present in these cases? All three cases are similar in respect to cognitive effort. In the first case, however, Alexis does all of the work in her own head. In the second case, there may be less cognitive effort on her part since she did not have to draw up the questions, and the possible answers were already provided so she did not have to think them up. But overall, there seems to be an equal amount of cognitive effort going on, distributed across a number of participants – including the experts. In the third case there may be even less cognitive effort going on in Alexis’ head -- she not only doesn’t have to draw up the questions and possible answers, she doesn’t have to produce the principles or rules required to make the judgment. Quite possibly there is less cognitive effort going on in the heads of the experts too, since they may be simply informing her of possible answers and rules that have been pre-established, e.g., in the legal system. The possible answers and rules have been instituted by previous practice and are made available for the juror. Indeed, we could say that such questions, possible answers, and rules create the tracks along which the cognitive process must run to keep it, literally, legitimate. The possible answers and rules are part of a system – stored in a system – that these people become cognitively engaged with. Moreover, they were previously established in some kind of process that we would certainly
call cognitive. Over all, in a distributed sense, there is an equal amount (if not more) cognitive effort to be found in the third case.

Consider now a fourth scenario. Alex attended law school and has gained a certain expertise in legal matters. When faced with precisely the same situation as Alexis in (1) – i.e., left on his own resources to formulate a judgment about a certain state of affairs – his resources allow him to organize his judging process using precisely those questions, possible answers, and rules provided by the legal system to Alexis in (3). Even on the conservative reading of the parity principle, we would have to say that if we count as cognitive the kinds of processes going on in Alex’s head (although quite possibly Alex uses a variety of media and technology that are part of a specific legal system in order to do what he does), then we would also have to count as cognitive the kind of process that Alexis is engaged with, even if some of the elements are instituted in the system and only available to her externally.

Judgments, then, are not necessarily confined to individual brains, or even to the many brains that constitute a particular court. They emerge in the workings of a large and complex institution. Yet these judgments and legal proceedings are cognitive processes that then contribute to the continued working of the system in the form of precedents. The practice of law, which is constituted by just such cognitive and communicative processes, is carried out via the cooperation of many people relying on external (and conventional) cognitive schemas and rules of evidence provided by the legal institution itself. A judgment made in such contexts is a form of cognition that supervenes on a large and complex system without which it could not happen. Indeed, it’s a cognitive practice that in principle could not happen just in the head. Even in the case of Alex, who seemingly does do it in his head, what he does depends not only on the fact that he has engaged in the workings of the legal (or educational) system (by attending law school), but on the ongoing workings of the legal system since what he does is what it is – a legal judgment – only in that system. An individual required to make judgments about the legitimacy of certain arrangements interacts with the legal institution and forms a coupled system in a way that allows new cognitive processes to emerge. Take away the external part of this cognitive process – take away the legal institution – and “the system’s behavioural competence will drop, just as it would if we removed part of its brain” (Clark and Chalmers 1998, p. 9).

We create these institutions via our own (shared) mental processes, or we inherit them as products constituted in mental processes already accomplished by others. We then engage with these institutions to do further cognitive work. If we think of the mind not as a repository of propositional attitudes and information, but as a dynamic process involved in solving problems and controlling behavior and action – in dialectical, transformative relations with the environment – then we extend our
cognitive reach by engaging with tools, technologies, but also with institutions.

Cognition as it is extended in social institutions, transforms our cognitive practices. Such institutions allow us to engage in cognitive activities that we are unable to do purely in the head, or even in many heads. If we are justified in saying that working with a notebook or a calculator is mind-extending, it seems equally right to say that working with the law, the use of the legal system in the practice of legal argumentation, deliberation and judgment, as well as the enforcement of law for purposes of controlling behavior is mind-extending too. This view pushes us beyond the strictly defined parity principle and extends the mind to a degree that even the liberal interpretation might have reservations about. Is the concept of mental institutions too large – an overextended mind – or is Clark and Chalmers’ concept of the extended mind not large enough?

**Parity principle revisited**

Consider again Clark’s three criteria: “resources” should be

1. Reliably available and typically invoked.
3. Easily accessible

Each criterion involves matters of degree. What counts as reliably available (1), or easily accessible (3), for example? Call my attorney; I’m sure she’ll tell you. It certainly seems that a legal system, or some part of it, may be reliably available even if I don’t carry it in my pocket. It may be only a phone call away. It also likely introduces more long-term stability and reliability unavailable in a single biological system. My attorney can consult her law books and codes, and together, in this process, and relying on easily accessible information and the mechanisms of the law, we can think through the problem and answer the question in a reliable way, or take action that leads to that result. Answering the question, solving the cognitive problem, may in fact be impossible without that access to the legal system. Indeed, one could imagine a specific kind of question that would never even come up if there were no legal system. The legal system in effect helps to generate certain cognitive events, and helps to resolve them.

With regard to the second criterion, why should some process that would otherwise count as a cognitive process not count as a cognitive process because it requires critical scrutiny, which is itself a cognitive process? There are plenty of instances of taking a critical metacognitive perspective (which is, of course, a cognitive process) on other problem solving acts of cognition. Taking such a perspective is itself a cognitive process, and again, in certain instances, that process may necessitate an institution like the law. That is, some critical perspectives are clearly legal
perspectives that supervene on a legal institution, and do so in a way that is even more “trustworthy” than biological memory.

More prolonged and complex external processes involve many elements, including processes that depend on social institutions. Despite the complexity, however, they may be more (or less) reliable; they may be (more or) less easy to access as a whole; and they may require some critical metacognitive scrutiny. But such things should not disqualify them from being cognitive processes.

Lack of parity along these lines should not disqualify such processes from being considered cognitive if they are processes to which the human organism is linked in the right way, that is, “in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right” (Clark and Chalmers 1998, p. 8). Accordingly, we can start to see that human cognition relies not simply on localized brain processes in any particular individual, or on short-term uses of notebooks, tools and technologies, but often on social processes that extend over long periods of time.

**Answering some objections**

In this final section I want to give some quick indications of how I would respond to some of the standard objections that have been raised against the extended mind hypothesis, and would likely be raised against this over-extended conception of mind. For reasons of economy and space I will not be able to develop full-fledged responses, but only indications of the directions in which I would take my responses.

Larry Shapiro (in press), in his review of Adams and Aizawa (2009), writes:

> When I dig a hole, the shovel aids me in this task, and it may even be true that I could not dig the hole without the shovel. But this does not extend my musculature into [things that are] independent of my musculature. According to Adams and Aizawa the same story holds true, *mutatis mutandis*, for cognition and its external accessories.

But the point is not that I extend my musculature – the point is that my *digging* is something extended from my bodily musculature across the shovel and into the ground. No one claims that I extend my brain by using a notepad, but rather that I extend the cognitive process. Nonetheless, as we know from neuroscience, such extensions may transform our brain; e.g., there are plastic remappings of somatosensory cortex for ego-centric coding of arm position when we use tools (Bassolino et al. 2009); and there are the famous studies of the enlarged hippocampi of London taxi drivers (Mcguire et al. 2000). If prolonged cognitive practices in an urban transportation system can change one’s brain, then it is not unlikely that prolonged
cognitive practices in a legal system could do the same for attorneys, judges, etc.

The point here is that as we engage with externalities our cognitive processes – including our neural processes, but not just our neural processes – are transformed. Externalities are not neutral in this respect. Mental practices are not simply externalized in the objective accomplishments or products of cognition, they extend into some of their externalizations as we engage with them to do further cognitive activity.

One of the most serious objections to the extended mind is the idea that even if externalities play a role in cognition, they play only a contextual or an enabling or causal role rather than a constitutive role. All of the real constituting action of cognition happens in the brain, even if it is in some way supported or facilitated by external elements. I think this objection is too closely tied to the conception of the mind that the extended mind hypothesis is rejecting. I want to suggest that we get beyond this issue by taking up a new conception of the mind, and a new understanding of how the brain works. Here I want to take seriously the idea that the term ‘extended’ goes along with (and not in opposition to) other terms: ‘embodied’, ‘embedded’, and ‘enactive’. We need to conceive of the brain, not as the place where all the mental processing and representing happens, but as part of a larger, embodied and environmentally embedded system; and we need to conceive of the mind, not on a functionalist interpretation (pace Clark, Wheeler, and many others), but as enactively generated in the specific interactions of organism-environment (where environment is social as well as physical). The mind, not as a collection of propositional attitudes, representations, mental states, etc., all of which have intrinsic or non-derived content, but the mind as an enactive and emotionally embedded engagement with the world through which we solve problems, control behavior, understand, judge, explain, and generally do certain kinds of things. On this conception, the mind is constituted primarily by just such activities, and anything that we might consider propositional attitudes are derivative and are inexplicable except in reference to such activities.

To include mental institutions in the notion of extended mind seems to be a good example of cognitive bloat, where cognition extends to all kinds of processes that seem at odds with the very notion (or the very traditional notion) of cognition. If my mind is extended by my use of Google to solve a problem, does that mean that cognitive processing is ongoing everywhere in cyberspace? Rowlands (2009), in response to the cognitive bloat argument, has suggested that part of what qualifies a process as cognitive is that it is owned by the agent. This notion of ownership, however, doesn’t seem to apply to mental institutions – no one owns the legal system, for example. Here, however, we might adopt a Lockean notion of ownership – ownership is constituted by the work invested – i.e., it is the fact that I am engaged in the right way with mental institutions, and this engagement makes them a constituent part of my cognitive process. Only so far as I am engaged with these institutions (or with notebooks or pieces of technology), do they
contribute to the constitution of my cognitive processes; and if I am not engaged with them (just as some neuronal processes in my brain may remain unactivated in specific circumstances) then they are not cognitively activated. Rupert (2009, 131) discusses “densely interactive processes” – “those in which the organism and environment affect each other in an ongoing way.” It seems clear that such densely interactive engagement or enactive coupling with those non-neural aspects would push out the defining limit for the concept of cognition. What constitutes the cognitive is tied to the specific kind of engagement that’s involved.

One attempt to answer the question of where cognition stops and something non-cognitive begins is to ask about “the mark of the mental” (see Adams and Aziawa 2009). Adams and Aizawa (2009) suggest that it is non-derived content. But the concept of non-derived content is not on settled ground. As Shapiro (in press) notes, “there is today no received theory of how original content comes to be in the first place,” and this means that it might be possible to find a theory of non-derived content that is consistent with extended cognition. Others have appealed to intentionality (following Brentano’s idea) as that mark (e.g., Menary 2009). Again, however, there is a good deal of disagreement about intentionality. Surely, for example, the experience of pain is a mental experience, but is it intentional? (see Crane 1998 for discussion). And some, including Husserl and Searle, have argued that not all mental experience is intentional. Furthermore, the argument this solution is meant to solve is simply repeated again if, as some claim, some forms of intentionality are purely internalistic (see, e.g., Horgan and Tienson 2002).

Fortunately, however, we can set such considerations aside because the question about the mark of the mental is not the right question, whatever the right answer might be to it. The question about the mark of the mental is not a question that can be answered without first determining whether some cognitive processes are extended, otherwise it amounts to question begging. In any case, it is not clear that anyone is claiming that all mental processes are extended or that those processes that are extended will have the same common features of non-extended mental processes. Clark (2010) and others, for example, deny that consciousness per se is extended; and no one argues that all aspects of the mental can be viewed as extended. Regardless of what is to be said about mental experience overall, I have been focusing on cognitive aspects found in such things as problem solving, judging, evaluation, and so forth. These “epistemic actions” are what I am calling cognitive. What kind of engagements or processes count as cognitive – what makes the action an epistemic action?

In this regard I appeal to a hermeneutical measure. What constitutes cognition as cognition (and as the kind of mental process that can be extended) is its interpretive character, where interpretation involves an interpreter (an experiencing subject, an agent) engaged with something (which is often other than itself, but sometimes just itself) in a process that produces meaning for the interpreter. We might think of the “with something” as a way to bring
intentionality into the picture. But the idea that cognition is meaning-producing or sense-making brings something else into the picture. As enactive approaches to cognition have suggested, sense-making processes involve interaction and have a certain autonomy, that is, they are not necessarily under an individual agent’s complete control, but often transcend the agent’s subjective processes. De Jaegher, Di Paolo, and Gallagher (in press) make this case in the realm of social cognition where they define social interaction as a two-way phenomenon. Social interaction specifically involves others, and patterns of engagement that can acquire their own form of self-organisation through mutual regulation. In the current context of extended cognition, where we can speak of interaction with institutions as well as with tools, instruments, technologies, etc., the point is that cognition just is any interaction or engagement that produces meaning for the agent, even if it is not the case that all mental experience necessarily produces meaning (e.g., pain), i.e., even if it is not the case that all mental experience is cognition.

Conclusions
It seems clear that the use of a legal system to solve a legal problem constitutes a case of complex “epistemic action,” and is an instance of extended cognition. “In all these cases the individual brain performs some operations, while others are delegated to manipulations of external media” (Clark and Chalmers 1998, p. 8). A legal system is a good example of this kind of sense-making process which supervenes on institutions and practices that exist outside of the head, but so are other types of institutions, including political, military, economic, religious, and cultural institutions, as well as science itself.

It is important to note, finally, that this over-extended version of the extended mind hypothesis motivates a critical normative perspective. Institutional structures, especially, can shape the way that we use certain technologies, and can allow us to see certain possibilities even as they blind us to others. We should take a closer and critical look at how social and cultural practices either productively extend or, in some cases, curtail mental processes. Pieces of technology, as well as specific institutions, offer possibilities, which at the same time carry our cognitive processes in particular directions. Such processes can have profound effects on us, and on our thinking. We should ask what the mechanisms and institutions do to us as subjects of cognition.

References
Clark, A. 2009. Spreading the joy? Why the machinery of consciousness is (probably) still in the head. Mind 118: 963-993; doi:10.1093/mind/fzp110