The Where and What of Cognition: The Untenability of Cognitive Agnosticism and the Limits of the Motley Crew Argument*

1. Cognitive Extension and Cognitive Agnosticism

According to classical cognitivism, cognitive processes are a special kind of computational processes over (symbolic or subsymbolic) representations. In the case of humans, these processes and representations seem to be found in and only in the brain—maybe not necessarily so, but arguably as a contingent matter of fact (Adams & Aizawa, 2001, 2008). Such a view squares nicely with, although it does not entail, what Susan Hurley (1998, 2001) dubbed the “sandwich model of cognition,” according to which cognition is the intracranial “filling” between perceptual input from and behavioral output to the world. Famously, Clark & Chalmers (1998) suggested a radical alternative to intracranialism. Cognitive processes, they held, can (and at least sometimes do) extend beyond the brain into the extracranial parts of the body and the extrabodily parts of the environment. This hypothesis of cognitive extension is best understood as a claim about the location of the material realizers of cognitive processes: while intracranialists locate the realizers of cognitive processes within the bounds of the skull, advocates of cognitive extension hold that cognitive processes can be materially realized (partially) by processes beyond the skull or skin.

* The paper is original and ownership possessed by the authors; it is neither published nor under consideration for publication elsewhere; its publication is approved by all authors; if accepted, it will not be published elsewhere in the same form without written consent of the copyright-holder.
During the past decade, the idea of cognitive extension\(^1\) has sparked a heated debate.\(^2\) A key issue has been how to argue for such a seemingly absurd view. Considerations of parity initially seemed pivotal: an extended process should be deemed cognitive if a relevantly similar, purely intracranial, process would (Clark & Chalmers, 1998). More recently, functionalism has been invoked: why should extracranial factors not qualify as realizers of cognitive processes, if they play the right kind of functional role (Wheeler, 2010)? Strikingly, opinions on these matters diverge widely (AUTHOR-a, in press; see also Drayson, 2010). While Wheeler (manuscript, ch. 5) thinks parity considerations provide the only plausible basis for cognitive extension, Clark (2008a, p. 77) denigrates them to a mere “rule of thumb,” and others deny they can play any substantive role at all (Coleman, in press; Di Paolo, 2009; Rupert, 2009, 2010). Similarly, while Clark (2008a) argues that Clark and Chalmers’ original argument is best viewed as a simple argumentative extension of functionalism and Mark Sprevak (2010) and Wheeler (2010, in press) maintain that functionalism entails cognitive extension, Shapiro (2008) and Rupert (2004, 2009) argue that functionalism is at best inconclusive and at worse at odds with cognitive extension.

These quarrels about the vices and virtues of our best current arguments for cognitive extension would obviously be pointless if we had what Adams & Aizawa (2001, 2008, 2009) call

---

\(^1\) Clark & Chalmers (1998) talk about the “extended mind” or “active externalism.” Other labels include “vehicle externalism” (Hurley, 1998), “wide computationalism” (Wilson, 1994), “locational externalism” (Wilson, 2004), “environmentalism” (Rowlands, 1999), and “integrationism” (Menary, 2006).

\(^2\) Not the least in this journal; see the special issue (vol. ??, 2010) on “Extended Mind,” edited by Leslie Marsh.
a “mark of the cognitive.” If we knew what makes a process cognitive, we could (at least in principle; see note 12) simply go and see where in the world those processes that fulfill the mark occur, rendering any disputes about cognitive extension unnecessary. Unfortunately, we seem to be rather clueless regarding the What of cognition. We just do not seem to know what, exactly, cognition is. After all, if we had a clue, we would not have been arguing about its Where so vigorously for over a decade now without making any progress.

We find this situation bewildering. How can we squabble about the location of cognitive processes if we do not even know which processes are cognitive to begin with? How can we expect to make progress on the Where-question, barring an answer to the What-question? After all, as Robert Rupert puts it, “the author who asserts that cognition extends into the environment had better be prepared to tell the rest of us what it is that extends into the environment” (in press, p. 3). Likewise, one should add, the intracranialist had better be prepared to tell us what it is that can occur within the confines of our skulls only. In other words: Cognitive agnosticism is untenable, where cognitive agnosticism is the view that one can fruitfully discuss the pros and cons of cognitive extension in the absence of an account of cognition, so that the failure to provide a mark of the cognitive should not prevent one from worrying about whether cognitive processes are an intracranial affair only, encompass extracranial parts of the body, or even stretch out into the extrabodily environment. Such a view is unsustainable: barring an answer to the

3 Apparently agnosticism-friendly proponents of cognitive extension include, e.g., Clark and Anthony Chemero. Clark’s pluralism regarding the cognitive (his view that cognitive extension is fully continuous with the cognitivist heritage of classical computational and representational approaches (Clark, 2008a, p. 198; see also pp. 152–156)), for instance, suggests an agnosticist reading (see also his [2008a, p. 239] approval of Richard Samuels’ remark “that we do not
What-question the Where-question is irresolvable, and this is unacceptable because the Where-question is a substantive issue that has to be resolved some way or other.

Section 2 claims that the Where-question cannot sensibly be pursued in the absence of an answer to the What-question because any attempt to establish or eschew cognitive extension fails unless a mark of the cognitive is provided. Section 3 argues that even a mild version of cognitive agnosticism—according to which our failure to provide a mark of the cognitive is unproblematic because there is no need to address the Where-question—fails, too, because cognitive extension is not only a fancy philosophical hypothesis but a natural consequence of successful research in cognitive science. The Where-question is thus a legitimate and important encounter anything like Adams and Aizawa’s demand for a ‘mark of the cognitive’ in other scientific fields”). Chemero (2009, p. 212) is even more explicit: “I disagree that proponents of radical embodied cognitive science [Chemero’s version of cognitive extension; AUTHORS] actually require a definition of ‘cognition.’ … There is no such thing.” Shapiro (in press, ch. 6.10) also seems to deny that a mark of the cognitive is needed (although he is not explicitly defending cognitive extension). Laudable opponents of agnosticism include Adams & Aizawa (2001, 2008, 2009) among the critics and, recently, Rowlands (2009a) among the proponents of cognitive extension. The view Sprevak (2010) calls “quietism” closely resembles cognitive agnosticism, and Sprevak rejects it for similar reasons we are eschewing cognitive agnosticism: “The hope was that one could appeal to general theories of mentality to decide whether extended cases were mental or not. If quietism is correct, then there is no way of resolving these cases: they are simply cases where competent observers differ” (2010, p. 523).

Matthew Barker (2010, p. 9) also comes to the conclusion that we have to say what cognition is in order to make any progress on the question where it is, but his argument is different from ours.

---

4 Matthew Barker (2010, p. 9) also comes to the conclusion that we have to say **what** cognition is in order to make any progress on the question **where** it is, but his argument is different from ours.
one we cannot afford to leave unresolved, and this means providing a mark of the cognitive is indispensable. However, instead of adding our own account to the growing list of suggestions, section 4 suggests that it may be worthwhile to start with a more general question: what kind of concept could “cognition” be? Along the way, section 4 will teach us something about the limits of a recent objection against cognitive extension sometimes referred to as the “Motley Crew Argument” (Shapiro, in press; Adams & Aizawa, 2010).

2. Intracranialism, Cognitive Extension, and the Mark of the Cognitive

The claim to be defended in this section is that any dispute about the pros and cons of cognitive extension, and thus about the Where of cognition, is futile unless it is theoretically backed by an adequate mark of the cognitive.

As indicated above, advocates of cognitive extension have appealed to parity considerations in order to show that extended processes can be cognitive processes. For instance, Clark & Chalmers (1998, p. 8) argued that an extracranial object may (partly) realize an agent’s cognitive processes if it fulfills the following “Parity Principle” (PP): “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 … part of the cognitive process.” Consider their example of three different ways to determine whether a falling Tetris block fits a socket in the bottom lines: (1) rotate the block’s image mentally; (2) rotate the block’s image physically on the screen by pressing a button; (3) rotate the block’s image mentally by means of a neural implant. Since the mental rotation in (1) is cognitive, the implant rotation in (3) seems cognitive, too. But then PP apparently demands that the physical
rotation in (2) is also cognitive, given that (2) is not fundamentally different from (3), except for
the fact that it involves an extended rather than an intracranial process.  

However, PP cannot, all by itself, establish cognitive extension. Note that PP is silent
both about the conditions under which we would grant that a part of the world functions as a
process we would, were it done in the head, have no hesitation in recognizing as cognitive and
also about whether these conditions are ever fulfilled. PP says that if we would not hesitate to
recognize a process that is as a matter of fact extended as cognitive if it were intracranial, then
the fact that it is extended should be no reason do deem it non-cognitive. What PP does not
tell us, however, is what exactly the conditions are under which we would consider the imagined
internal process to be cognitive. What the appeal to PP presupposes, in other words, is a mark of
the cognitive. Absent such a mark, PP is inapplicable. The same holds for intracranialists who
deny PP. Why should an extended process that is in all important respects like an imagined
intracranial cognitive process not be regarded as cognitive? The only plausible answer seems be
that, despite their alikeness, the intracranial process fulfills the mark of the cognitive (whatever it
is) while the extended one does not. Disputes about PP and its role in establishing cognitive
extension will thus be irresolvable unless a mark of the cognitive is provided.

Shapiro (in press) disagrees: “Marks of the cognitive would surely be useful were one
unsure whether the process under investigation were cognitive, but when this is not an issue, as it
should not be in the case, for instance, of memory or spatial reasoning, demanding their
characterization might fairly be dismissed as a red herring” (manuscript, p. 255). First, however,
although memory and spatial reasoning may be obvious cases, whether, for instance, emotions,

5 Coleman (in press) argues convincingly that the story has to be more complicated than Clark
and Chalmers thought, but for the current (illustrative) purpose the simple version suffices.
affectivity, consciousness, comprehension, or instinct are cognitive phenomena is a question which will not receive an unanimous answer from cognitive scientists. Absent a mark of the cognitive that decides the issue, a parity argument for these phenomena will thus not get off the ground. Second, consider an objection against cognitive extension raised by Sprevak (2010). In a nutshell, Sprevak argues that PP would commit us to claiming that someone owning a PC running software that calculates the dates of the Mayan calendar is subject of an extended cognitive process stretching out into the PC. Why? Because one can imagine a Martian using the same algorithm the PC uses to calculate the dates of the Mayan calendar, but intracranially. Since the process in the Martian’s case is intracranial and functions as the extended process in the case of someone owning a PC with appropriate software, the latter process should be deemed cognitive, too. That, however, follows only if the intracranial process in the Martian is cognitive to begin with. But how are we to decide this issue? Who is able to say, in the absence of a principled way of distinguishing the cognitive from the non-cognitive, whether a process in a Martian’s head that implements the same algorithm by which ordinary PCs calculate the dates of the Mayan calendar is cognitive? This issue cannot be resolved simply by pointing out, as Shapiro does, that “cognitive scientists [are] familiar enough with cognitive processes to recognize one when they see it” (manuscript, p. 255). Without a principled way of determining, in controversial cases, whether the imagined intracranial process is cognitive, PP does not universally deliver a straightforward verdict.6

6 One may object that it does not matter that cognitive scientists have no clear-cut intuitions regarding Sprevak’s Martian because it is just a fancy philosophical thought experiment. This, however, misses the point. The point is that cases like Sprevak’s Martian are in fact appealed to
One might object that even if PP is inapplicable in some controversial cases, Shapiro is right that it is applicable in cases where it is obvious that the intracranial process with regard to which parity is assessed is cognitive (if, say, it is a case of memory or spatial reasoning). PP can thus be used to establish cognitive extension for at least these cases, and this is enough to decide the Where-question in favor of cognitive extension. This response, however, ignores a more subtle way in which PP presupposes a mark of the cognitive, which can be exposed by making explicit a connection between PP and a general commitment to functionalism that is often left implicit.

Functionalism has come to play an increasingly prominent role in the debate about cognitive extension (Clark, 2008a; Drayson, 2010; Shapiro, 2008; Sprevak, 2010; Wheeler, 2010). The general idea is that cognitive processing is extended because what makes a process cognitive is its having the right “functional poise” (Clark, 2008a, p. 78) and at least some extracranial processes do indeed contribute in the right way to the functional roles characteristic of our cognitive capacities. The appeal to functionalism seems to have a clear advantage over the appeal to PP. In contrast to the latter strategy, the former does not require us to compare an extended process with an intracranial one whose status as cognitive has first to be established independently. Rather, the argument is simply that an extended process plays the right functional role within an agent’s cognitive life.7

---

7 In fact, however, the two issues are often mangled insofar as the appeal to functionalism is combined with a comparison of an extended and an intracranial process. Consider again the Tetris example: Since “case (2) with the rotation button displays the same sort of computational
Yet, the dispute about the viability of functionalism as a means of establishing cognitive extension also presupposes a mark of the cognitive. On the one hand, advocates of cognitive extension who appeal to functionalism must assume that cognitive processes are individuated at such a coarse level that an extended process can fulfill the functional role characteristic of a given cognitive capacity despite the fact that it potentially differs from the more familiar intracranial realizers both at finer functional levels and at the level of material implementation. Clark, for instance, reverts to some kind of “‘commonsense functionalism’ concerning mental states … [according to which] normal human agents already command a rich (albeit largely implicit) theory of the coarse functional roles distinctive of various familiar mental states” (2008a, p. 88). On the other hand, intracranialists must hold that the fact that extended processes differ from the more familiar intracranial realizers of our cognitive capacities at a finer functional or implementational level renders them non-cognitive, even if they possess the right coarse functional poise. The point of disagreement is thus again what exactly it is for a process to be cognitive. Is our conception of the cognitive such that differences at a finer functional or implementational level matter qua cognitive or not? This dispute can be resolved by and only by a mark of the cognitive. Absent such an independently motivated mark, advocates of cognitive extension will accuse intracranialists of begging the question against a more liberal account of the cognitive according to which only the coarse functional profile matters, while intracranialists structure as case (3),” Clark and Chalmers (1998, p. 7) argued, the processes in (2) and (3) seem to be functionally equivalent, so that (2) should be cognitive if (3) is. More on the connection between PP and functionalism below.

8 On the importance of coarse vs. fine grained functional roles in the debate about cognitive extension see Rupert (2004) and Wheeler (2010).
will accuse advocates of cognitive extension of begging the question against a more restrictive account of the cognitive according to which finer functional or implementational differences indeed do make a difference.

This also shows why, without a mark of the cognitive, PP is not applicable even in cases where it is obvious that the intracranial process with regard to which parity is assessed is cognitive. Suppose we know that an intracranial process is one of memory and thus undoubtedly cognitive. Still, the answer to the question with which the appeal to PP stands and falls—viz., Does the extended process function just as the intracranial process does?—depends upon the level of grain at which the functional profile of memory is characterized. If memory is characterized in terms of its coarse functional role, then the application of PP may yield an extension-friendly result, but if not, intracranialism will remain undefeated. And which level of grain is the correct level is, as we saw above, a question that can only be settled by providing a mark of the cognitive.

The foregoing shows that any dispute about the Where of cognition based on either PP or functionalism is futile unless a mark of the cognitive is provided. Therefore, the attempt to argue for cognitive extension by appeal to either PP or functionalism faces a dilemma: absent a mark of the cognitive, the argument will not get off the ground, but given such a mark, it will be superfluous, because once we have a mark, we can simply go and see where in the world the processes fulfilling it fall (AUTHOR-a, in press; see also Rupert, 2010, p.3, who independently comes to the same conclusion regarding PP). Something similar holds for other attempts to argue for (or against) cognitive extension.

Consider Robert Wilson’s argument for what he calls “wide computationalism” (1994, 2004; Wilson & Clark 2009). According to Wilson, cognitive processes are computational
processes, and in some cases the computations do not involve intracranial representations, but extracranial representations or the extracranial objects themselves. As a consequence, Wilson claims, cognition is extended in the sense that “at least some of the computational systems that drive cognition reach beyond the limits of the organismic boundary” (2004, p. 165). Thus, Wilson and the traditional intracranialist computationalist agree that cognition is computation, but they disagree over what the objects of computation are. Importantly, however, not all computational processes are cognitive processes. What we need to know, in other words, is which computational processes are such that they are cognitive and which computational processes are such that they fail to be cognitive. The traditional computationalist will argue that Wilson’s wide computational processes belong to the latter sort, i.e., are non-cognitive computational processes, while Wilson will insist that they are of the former sort, i.e., computational and cognitive. Once again, the crucial question is how to settle this dispute, and once again the answer seems to be that what we need to know is what exactly it is that makes the wide computational process a cognitive process (if Wilson is correct) or a non-cognitive one (if the intracranialist classicists are correct). In order to decide the issue in their favor, therefore, both parties need to go beyond a mere appeal to computationalism and provide an independently motivated mark of the cognitive.

Finally, consider the attempt to decide the Where-question in favor of cognitive extension by appealing to (real or imagined) cases in which it is “just obvious” that the cognitive process in question is extended. Imaginary cases could include Clark & Chalmers’ (1998) example.

---

9 See, e.g., Adams & Aizawa (2010): the “computational theory of the mind maintains not that any computation is a kind of cognition, but that only some specific forms of computation (yet to be discovered and characterized) constitute cognition” (manuscript, pp. 23–24).
concerning Inga and Otto (see section 3), real cases might include the popular problem-solving game Rush Hour for children (Wilson 2004, p. 193) or case studies from cognitive science, like Randall Beer’s (2003) simulated robot that is controlled by an evolved artificial neural network and able to categorize circle-shaped and diamond-shaped objects in its artificial environment (discussed as a case of cognitive extension in, e.g., Chemero (2009)) or Dana Ballard et al.’s (1997) work on minimal memory strategies (discussed in, e.g., Clark (2008a)).

Arguments on a case by case basis are of course notoriously problematic. For the current purpose, however, the important point is a different one. Advocates of cognitive extension who think that in the examples above the extracranial factors are indeed constitutive of the cognitive processes in question will inevitably face resistance from intracranialists who think they are (at best) an important causal contributor to, but not constitutive of, the cognitive processes.10 The nub is that no example is, all by itself, decisive. Examples must be interpreted, and any interpretation presupposes a theoretical background against which it is made. Whether an example is interpreted in an extension-friendly or an intracranialist way depends upon a theoretical account of what it means for a process to be cognitive. Whether or not an extended process is taken to be cognitive, i.e., whether the extracranial factors are (partly) constitutive of an extended cognitive process or only a causal contributor to an intracranial cognitive process, is something that can be decided only against the background of a mark of the cognitive. Absent

---

10 This is of course related to what Adams & Aizawa (2001, 2008) dub the “coupling/constitution fallacy” (see also Aizawa, 2010). In Rupert’s (2004) terminology the debate is over whether HEC (the hypothesis of extended cognition) or HEMC (the hypothesis of embedded cognition) is the correct approach to cognition to adopt (more on embedded cognition in section 3).
any account of what it means for a process to be cognitive it is impossible to say whether an example (real or imagined) is a case of cognitive extension or of intracranialism.

Once again, either side of the debate needs a mark of the cognitive to make its case against the other. Barring a mark of the cognitive, the debate threatens to dwindle into an endless and unrewarding back and forth over who begs the question against whom by presupposing either too liberal a conception of the cognitive that mistakenly equates causal contribution with constitution, or too restrictive a conception of the cognitive that mistakenly assigns to a genuine constituent of a cognitive process the role of a causal contributor only. As Shapiro (in press) aptly puts it: “the controversy over Constitution [i.e., cognitive extension; AUTHORS], if it is to avoid dwindling into linguistic insignificance, must confront questions over the meaning of mental, or cognitive, processes. Whether the body and world [...] may be constituents of cognition or merely causally related to cognition depends, first, on what we mean by cognition” (manuscript, p. 221).11

What has just been said points towards the most fundamental, and entirely general, reason why quarrels about the pros and cons of cognitive extension are futile unless an independently motivated mark of the cognitive is provided. Any argument for cognitive extension must appeal to some feature of extended processes—be it their parity with imagined intracranial processes, the fact that they fulfill a certain coarse grained functional role, the fact that they are computational processes, their intimate coupling to an agent or some other feature we have failed to mention—that is allegedly sufficient (even if perhaps not necessary) for rendering them cognitive. In response, the intracranialist must deny that the feature in question is sufficient for

11 How to square this with other passages (like the one cited on p. XXX above) in which Shapiro seems to deny that a mark of the cognitive is needed is, however, something that eludes us.
rendering a process cognitive and/or appeal to some allegedly necessary condition for a process’ being cognitive that is exhibited (if only as a contingent matter of fact) by intracranial processes only. In response, the advocate of extended cognition must then deny that possessing the intracranialist’s favored feature is indeed necessary for a process’ being cognitive. And so on. Such disputes, unless they are underwritten by a theoretical background telling us what the sufficient and/or necessary conditions of cognition are, will denigrate into an endless theoretically unfounded intuition mongering. In other words: Before we can sensibly expect to make any progress on the Where-question, we need a mark of the cognitive, i.e., an answer to the What-question. Cognitive agnosticism is untenable.12

3. Mild Cognitive Agnosticism

Cognitive agnosticism, we said, is the view that one can discuss the pros and cons of cognitive extension while refraining from providing a mark of the cognitive. Thus understood, we argued, cognitive agnosticism is untenable. In response, one may grant the link between the Where- and the What-question established in section 2, insist that there is no answer to the What-question and still avoid the problem by denying that the Where-question is a reasonable question to ask in the first place. If there is no need to make progress on the Where-question, we may remain agnostic vis-à-vis the What-question. This section argues that even this mild form of cognitive agnosticism is untenable (see also AUTHOR-a & AUTHOR-b, in press).

12 Nothing that has been said entails that once we have an answer to the What-question the Where-question is thereby eo ipso also resolved. It may happen that it is difficult, or even impossible, to say exactly where in the world those processes that fulfill the mark of the cognitive fall.
In order to dismiss the Where-question as irrelevant, one could try to discredit it as idle philosophical theorizing relevant only for those interested in abstract and ultimately extraneous intellectual gymnastics. Consider, for instance, Clark & Chalmers’ (1998) (in)famous thought experiment concerning Inga, a normal human believer who hears about an art exhibit at the MoMA, recalls that the MoMA is on 53rd street and heads off, and Otto, who is suffering from Alzheimer’s disease, relying on a notebook in which he records information he thinks he will later need to remember and who, upon deciding he wants to see the exhibit, consults his notebook, finds out that the MoMA is on 53rd street and heads off, too. Under certain conditions, Clark and Chalmers argued, the information in Otto’s notebook functions just like the information stored in Inga’s biomemory and so should, given PP, be considered to be (part of) the material realizer of Otto’s dispositional belief about MoMA’s location. But why, one may wonder, should we care at all about whether the notebook entries (partially) realize Otto’s cognitive life? After all, Otto is an entirely fictitious cognitive agent, an invention of philosophers. If philosophers want to indulge in futile guesswork about the Where of Otto’s cognitive life, and if it turns out that they must provide a mark of the cognitive in order to do so, then, one may argue, so be it. So much the worse for philosophy if it turns out there is no mark. But the rest of us should simply dump such issues and resort to mild cognitive agnosticism.13

13 It has been argued (e.g., Rupert, 2004, 2009) that the ontologically more conservative alternatives to the hypothesis of cognitive extension are explanatorily superior, which could be seen as supporting such a case for mild cognitive agnosticism—be not ontologically profligate beyond explanatory necessity! For the claim that it is actually the hypothesis of cognitive extension that is explanatorily superior see Clark (2008a, ch. 6). Barker (2010) and Sprevak (in
This attempt to save mild cognitive agnosticism fails. Cognitive extension is not primarily a philosophical speculation for the sake of intellectual gymnastics, but a hypothesis that flows quite naturally from two successful and independent (and potentially incompatible) strands of empirical cognitive science research—work done under the heading of “dynamical systems theory” on the one hand and “embodied” and “embedded cognition” on the other.

A first route to cognitive extension is via a dynamicist approach to cognition that emphasizes the importance of the dynamical coupling of cognitive agents to their environments in immediate, real-time interactions under continuous reciprocal causal influence. The idea is that brains are seamlessly integrated into bodies and environments in such a way that intracranial, bodily, and extrabodily processes form a single, dynamically changing whole: “The cognitive system is … a structure of mutually and simultaneously influencing change. The cognitive system does not interact with other aspects of the world by passing messages or commands; rather, it continuously coevolves with them” (Van Gelder & Port, 1995, p. 2). Due to this continuous co-evolution, the agent and its environment can be modeled as a nonlinearly coupled dynamical system which, as such, is unified and not decomposable into separate parts: “It is only for convenience (and from habit) that we think of the organism and environment as separate; in fact, they are best thought of as forming just one nondecomposable system” (Chemero, 2009, p. 36). The result, of course, is cognitive extension. Since cognitive systems are dynamically changing, nondecomposable wholes, “[c]ognitive processes span the brain, the body, and the environment” (Van Gelder & Port, 1995, p. ix).

press) argue that considerations of explanatory value in cognitive science (can) neither support nor disconfirm cognitive extension.
A second route to cognitive extension is to view it as something that follows naturally from *embodied* and *embedded* approaches to cognition.\(^{14}\) According to the embodied approach, an agent’s cognitive life is grounded in her physical features and abilities. This idea is sometimes expressed by saying that “the presence of a humanlike mind depends quite directly upon the possession of a humanlike body” (Clark, 2008b, p. 43; emphasis added). Thus understood, the embodied approach holds that cognitive processes, although dependent upon extracranial bodily processes, are still an entirely intracranial affair. According to a stronger reading, however, cognitive processes are not only dependent upon but actually constituted by extracranial bodily processes, so that cognitive processing leaks out into the body.\(^{15}\) The *embedded* approach is a natural extension of the two embodied approaches by additionally stressing the role of the environment and its active structuring by the agent for the purpose of facilitating cognitive processing. An explanation of cognitive processes can thus not be limited to their (causal or constitutive) grounding in extracranial but still intrabodily processes, but must appeal to the way an agent uses her extrabodily environment’s structure or actively structures her extrabodily

\(^{14}\) There is, unfortunately, no uniform terminology (Robbins & Aydede, 2009). What we say below strikes us as plausible readings of “embodied” and “embedded” cognition (AUTHOR-a, 2010). Terminological disagreements do not, however, affect the general point that work bearing that name naturally leads to cognitive extension.

\(^{15}\) Support for this stronger claim comes, e.g., from studies on vision (Ballard *et al.*, 1997; Noë 2004; O’Regan & Noë, 2001) that can be interpreted as showing that bodily movements are “as much part of vision as the detection of disparity or the calculation of shape from shading” (Shapiro, 2004, p. 188), so that “[v]ision for human beings is a process that *includes* features of the human body” (p. 190; both emphases added).
Like the first embodied approach, the embedded approach characterizes the grounding relation as a dependence relation, but it extends the dependence base from extracranial bodily to extrabodily processes. Cognitive processing thus takes place in the brain and the extracranial parts of the body only, although it is dependent upon extrabodily processes. Once the idea of embedded cognition is on the table, the hypothesis of cognitive extension is just a reasonable next step. If extrabodily processes are so crucial for our cognitive life, then why should they not be said to realize it? Just as the embedded approach simply extends the dependence base of the first version of embodiment from extracranial bodily to extrabodily processes, the hypothesis of cognitive extension simply extends the constitution base of the second version of embodiment from extracranial bodily to extrabodily processes.

None of this shows that cognitive processes do extend. But it shows that the idea of cognitive extension is at least continuous with two widely held and (rightly) feted approaches to cognition that are well-entrenched in empirical cognitive science research. This is ample reason to not only raise the Where-question, but regard it as a central and utmost important question that need not only concern philosophers in their metaphysical moods but anyone in cognitive science. The fact that the idea of cognitive extension squares nicely with two so radically different, if not even incompatible, research programs, only adds to the argument’s force. The idea of cognitive extension is not a consequence of a specific commitment to, say, computational as opposed to

For instance, recent research on visual processing (Ballard et al., 1997) suggests that instead of creating detailed internal representations as the basis for later stage cognitive processing, human subjects extract the relevant information “on the fly” from the world itself. David Kirsh & Paul Maglio’s (1994) research on epistemic actions highlights a similar kind of environmental “offloading.”
non-computational, or representational as opposed to non-representational or abstract as opposed to applied approaches, but rather an idea accompanying all of what may be called “second-generation cognitive science.” Therefore, we have to take it seriously. Therefore, we have to address the Where-question. Therefore, even mild cognitive agnosticism is untenable.

Section 3 argued that the Where of cognition is an issue that ought to be resolved. Section 2 showed that without an independently motivated account of the What of cognition, quarreling about its Where is pointless. Time therefore, it seems, to lay our cards on the table. What do we think cognition is? Instead of tackling this question head-on by adding yet another account to the growing list of suggestions, we propose to ask a less demanding question first: what kind of concept should we expect “cognition” to be? We will sketch the consequences the various possible answers to this question have for the debate about the Where of cognition. Along the way, we will learn something about the limits of a recent objection against cognitive extension, sometimes referred to as the “Motley Crew Argument” (Shapiro, in press; Adams & Aizawa, 2010).

17 If it eventually turns out that the Where-question has to be answered in an intracranialist way, then the research discussed in this section, although it naturally leads to, or suggests, cognitive extension, would of course not loose its importance. An intracranialist answer to the Where-question would only lead to a re-conceptualization of the empirical research. This does not, however, mean that contrary to what has been claimed in this section the Where-question is just an extraneous add-on to the empirical research.
The arguably most natural approach is to take “cognition” to be a concept that picks out a natural kind. Natural kind concepts are concepts for which there is a scientifically discoverable essence in the form of a set of conditions the satisfaction of which is both necessary and sufficient for something’s falling under the concept. For instance, something falls under the concept “aluminum” if and only if it has atomic number 13, and something falls under the concept “tiger” if and only if it has the distinctive genetic code characteristic of all and only tigers. Something along these lines may hold for “cognition” as well. Among the critics of cognitive extension, for instance, Adams and Aizawa could be read as defending a natural kind view of cognition.\textsuperscript{18} Cognitive processes, they hold, “are natural kinds of processes” (2008, p. 80) that involve computational operations over states with non-derived content and are implemented by special kinds of mechanisms. The fact that this scientifically discoverable essence of cognition is currently found in the brain only supports what they call “contingent intracranialism,” by providing “defeasible reason to suppose that cognitive processes are typically brain bound and do not extend from the nervous system into the body and the environment” (2008, 70).\textsuperscript{19} Since, for reasons to be discussed below, the view that cognition is a

\textsuperscript{18} “Could” because Aizawa (2010) explicitly denies that they intended the characteristics discussed below to (jointly) sufficient for something’s being cognitive.

\textsuperscript{19} Incidentally, it seems doubtful whether Adams and Aizawa’s mark of the cognitive has indeed the consequence that cases of cognitive extension are at best possible, but not actual. In Adams & Aizawa (2001, pp. 50–51), for instance, they stress that while a cognitive process must involve non-derived content, not every aspect of it must, all by itself, have non-derived content, and in Adams & Aizawa (2010) they point out that all they had in mind was “that if you have a process that involves no intrinsic [i.e., non-derived] content, then the condition rules that the process is
natural kind is intimately tied to intracranialism, the search for a mark of the cognitive is a rather straightforward matter if “cognition” is a natural kind concept—go and study the brain until you can spell out adequate necessary and sufficient conditions for something’s being cognitive.

Not every account of cognition in terms of a set of necessary and sufficient conditions is a natural kind view, however. Mark Rowlands (2009a, p. 8), e.g., argues that a process is cognitive if and only if it involves information processing that has the proper function of making available, by means of representational states, either to the subject to which it belongs or to subsequent operations, information hitherto unavailable. Since Rowlands is an advocate of cognitive extension, he thinks that among the processes fulfilling this condition are at least some that “extend into the cognizing organism’s environment in that they are composed, partly ..., of manipulative, exploitative, and transformative operations performed by that subject on suitable environmental structures” (2009a, p. 1). Precisely this integration of external operations on worldly items, however, seems to render Rowlands’ account incompatible with a natural kind.

---

20 The idea of a process’ being owned by or belonging to a subject does not commit Rowlands to intracranialism, because ownership is “determined not by spatial containment but by a certain kind of integration” (2009a, p. 17).
view; for in contrast to the causally relatively homogenous brain processes Adams and Aizawa focus on, the various operations by which cognitive agents manipulate, exploit, and transform environmental structures seem to be causally heterogeneous. The extended processes that advocates of cognitive extension like Rowlands want to count as cognitive do not form a natural kind.

Moreover, if extended processes are not a natural kind, then—since science is interested in causal taxonomies of things only (i.e. taxonomies that groups things with the same causal powers together [Fodor, 1987])—they form, as Shapiro (in press) puts it, a “foul hybrid,” “a nasty mess that is not amenable to scientific scrutiny” (manuscript, p. 260). This is the punch line of what is sometimes called the “Motley Crew Argument” against cognitive extension.

According to the Motley Crew Argument, advocates of cognitive extension are committed to viewing “cognition” as something like an umbrella concept, i.e., a concept subsuming a diversity of processes that share only the fact that they fall under this concept. In that case, “cognition” would be nothing but a label for a causally heterogeneous lot, for a mere disjunction of causally diverse processes, and effectively a hypernym, comparable to concepts like “things in our living room” or “things that can be used as a paper weight.” And just as the things in our living room or the things that can be used as a paper weight do not form a kind suitable for scientific investigation, there could be no science of the cognitive if “cognition” were indeed an umbrella concept and cognitive processes a complete motley.

One thing to note about the Motley Crew Argument is that strictly speaking it is not so much an argument against cognitive extension as an argument against any account of the cognitive that treats “cognition” as an umbrella concept. And although the umbrella view seems to square better with an extended approach to cognition—which is, by the very nature of the
approach, more liberal with regard to what processes can count as cognitive than intracranialists—there is no principled reason why intracranialists should not adopt it, thereby rendering their account subject to the kind of criticism raised by the Motley Crew Argument.

Another thing to note about the Motley Crew Argument is that one should distinguish the claim that “cognition” is not a natural kind term from the claim that “cognition” is an umbrella term. Although advocates of the Motley Crew Argument sometimes seem to suggest that the problem is that the processes the advocate of cognitive extension wants to count as cognitive do not form a natural kind, one can hold that cognitive processes are not a natural kind and yet deny that they are a motley that falls under an umbrella concept only and therefore succumbs to the criticism raised by the Motley Crew Argument.

Adams and Aizawa (2010), e.g., argue that “[t]ools do not constitute a natural kind; tools are, after all, artifacts. It is for this reason that a would-be brain-tool science would have to cover … a genuine motley” (manuscript, p. 14). The fact that the processes of manipulation, exploitation, and transformation of or by means of tools involve artifacts admittedly shows that they are not a natural kind. From the point of view of science, however, what matters, one may say, is kindhood, not natural kindhood. Robotics studies robots, computer science and artificial intelligence study, among other things, algorithms and artificial neural networks, and physics studies, among other things, superconductors and oscillators, although robots, algorithms, artificial neural networks, superconductors, and oscillators are not natural kinds, but artifacts. The mere fact that extended processes involve artifacts and so fail to constitute a natural kind does therefore not entail that there could be no science of the cognitive if cognitive processes

21 Laura Taylor (2005), e.g., is not defending cognitive extension and nevertheless takes “cognition” to be “an umbrella term that refers to all of the mental activities we engage in” (p. 2).
were extended. What matters is not whether the processes in question are natural kinds, but whether they are kinds.

“Cognition” may pick out a class of processes that are causally heterogeneous, but that, despite their causal heterogeneity at a lower level, exhibit enough higher level causal homogeneity for them to be worthwhile objects of scientific studies.22 “Cognition,” in other words, may be more like “robot” or “algorithm,” and unlike both natural kind concepts like “tiger” or “aluminum” and umbrella concepts like “things in our living room” or “things that can be used as a paper weight.” If this is true, a mark of the cognitive would not be spelled out in terms of lower level communalities among all and only cognitive processes but by appeal to higher level causal characteristics.

The Motley Crew Argument would thus not only have to claim that extended processes are not natural kinds, but that they lack even the sort of “higher level unification despite mechanistic dissimilarities” (Clark 2008a, p. 96) we are familiar with from robots, algorithms, artificial neural networks, superconductors, and oscillators. The latter would be the case if “cognition” were what is sometimes called a nominal kind concept, i.e., a concept that does not

---

22 This is, of course, reminiscent of Fodor’s (1974) point about the indispensability of special science taxonomies: “Some monetary exchanges involve strings of wampum. Some involve dollar bills. And some involve signing one’s name to a check. What are the chances that a disjunction of physical predicates which covers all these events … expresses a physical natural kind? The point is that monetary exchanges have interesting things in common … But what is interesting about monetary exchanges is surely not their commonalities under physical description” (1974, 103–104).
provide a scientifically discoverable, but a nominal essence, i.e., a definition that specifies form, function, or origin of the concept’s referents.\textsuperscript{23}

A nominal kind view of cognition is neutral with regard to the question whether cognitive processes are intracranial or extended. Rowlands’ account of cognition as information processing that has the proper function of making representationally available information hitherto unavailable is an example for a nominal kind view according to which extended processes count as cognitive. A nominal kind account may, however, also be offered by someone convinced that cognition is an entirely intracranial affair. Any account according to which cognition sunders into a variety of causally heterogenous processes that lack a scientifically discoverable essence and yet are characterized in terms of their form, function, or origin in such a way that only intracranial processes fulfill the mark will be an intracranialist nominal kind view. In both cases, cognitive processes form a kind—provided that their form, function, or origin that is used to characterize their nominal essence supplies them with enough higher level unification. Whether it does is an open scientific question, but one should not—as the examples of other sciences show—assume without further argument that extended processes are going to exhibit less higher level unification than intracranial processes, simply because the former, but not the latter, involve artifacts.

Apart from a nominal kind view, there are, as far as we can see, two further accounts of the cognitive according to which “cognition” is neither a full-fledged natural kind concept nor a mere umbrella concept, but picks out a class of items characterized by at least some sort of higher level unification.

\textsuperscript{23} See, e.g., Schwartz (1980). The distinction between “real” and “nominal” essences is, of course, John Locke’s.
First, “cognition” may be a cluster concept. Cluster concepts are characterized by a set of conditions such that something falls under the concept if and only if it satisfies a goodly number of them, while several possible—and not necessarily overlapping—subsets of these conditions can be sufficient and it may be vague just how many conditions must be fulfilled. Dutton (2009), e.g., has recently argued that “art” is a cluster concept in this sense: there are a dozen features shared by the most unequivocal art, where some pieces of art no doubt satisfy fewer than the full set and there is no fixed number of conditions that is sufficient. “Cognition,” one may hold, is just such a cluster concept. If “cognition” works the way “art” does, then the worry that there can be no science of cognition because extended processes form a nasty mess not amenable to scientific scrutiny seems to dissipate, even though there may not even be a single joint characterization in terms of form, function, or origin. Although in a sense “art” picks out a causally heterogeneous motley, artworks can nevertheless be the subject matter of science, and although not a full-fledged (natural or nominal) kind, they are by no means a mere “foul hybrid” either. Therefore—although a cluster account of cognition is per se once again neutral with regard to the debate between the intracranialist and the advocate of cognitive extension—it may seem like a particularly good idea from the latter’s point of view.  

24 Poirier & Chicoisne (2008, p. 30), e.g., argue that “cognition” is a cluster concept and that cognition is extended.
and intracranialists may be resolved in favor of the former by finding an extended process which definitely is cognitive.

However, “cognition” just does not seem to work like a cluster concept. Is it really plausible to assume that memory processes count as cognitive in virtue of fulfilling one set of conditions, while spatial reasoning processes count as cognitive in virtue of fulfilling a different, albeit overlapping, set of conditions, and language comprehension processes count as cognitive in virtue of fulfilling yet another overlapping set of conditions, and so on? It is not impossible that this is the correct way of thinking about cognitive processes, but one would at least want to hear something more about the different sets of conditions that make it the case that one or the other paradigm case of a cognitive process counts as cognitive. The problem seems to be that advocates of extended cognition and intracranialists are unlikely to agree on even a “pool of conditions” from which some vague number of conditions has to be fulfilled. Therefore, the cluster view only postpones the problem.

Second, “cognition” may be a family concept. “Game,” e.g., is often said to be a family concept. The conditions something has to fulfill in order to count as a game, Wittgenstein (1953) argued, include “a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities” (§66), but there is nothing that is common to all, or even all and only, games. Although two arbitrary games are always similar in some non-trivial respect, *game* has neither a natural nor a nominal essence in the form of a set of necessary and sufficient conditions, nor is there something like a “pool of conditions” from which some vague number of conditions has to be fulfilled.\(^{25}\) The same may be true of “cognition.” Like games, any two

\(^{25}\) The terms “cluster concept” and “family concept” are often used interchangeably. There are, however, important differences. First, for a cluster concept there is a fixed set of conditions such
cognitive processes may resemble one another in at least some significant respect, although there is neither something in common to all and only cognitive processes nor a set of conditions such that a process is cognitive if and only if it satisfies a goodly number of them. Once again, a family account may seem attractive to the advocate of cognitive extension. It is comparatively liberal with regard to the conditions a process must fulfill in order to count as cognitive, and if “cognition” functions like “game,” a science of cognition would not seem to be out of reach, because cognitive processes, if they indeed are characterized by a set of family relationships, are more than a mere motley, just as games form a more homogeneous set than the things in our living room or the things that can be used as a paper weight. Wheeler (manuscript), e.g., argues for cognitive extension and suggests that cognitive processes can be (i) noncomputational and nonrepresentational, (ii) noncomputational and representational, and (iii) computational and representational, where “there is clearly a fairly straightforward family resemblance structure to this space of mechanisms” (manuscript, ch. 3, p. 18). The problem, however, is that if all that is required for a process to be cognitive is a loose family resemblance with some other process clearly regarded as cognitive, then the account is too liberal: The Java applet running on our computer, e.g., contains computational and representational processes that resemble our own cognitive processing of the items on the screen in some interesting respect or other, although we that anything falling under it must satisfy a sufficient number of them, while the conditions characteristic of family concepts are open-ended. Second, while the set of things falling under a cluster concept may share a single common characteristic, the set of things falling under a family concept do have nothing relevant in common. Third, while two things that fall under the same cluster concept may not have a common characteristic, any two things falling under a family concept will share some relevant characteristic or other.
should—on pain of “cognitive bloat”, i.e. the problem of not knowing where to stop once we enter into cognitive extension (Rowlands 2009b, p. 634)—not take the processes in the Java applet to be genuinely cognitive. This is not a problem for the advocate of cognitive extension only, but also for the intracranialist who opts for a family view of cognition. Some processes within the bounds of our skull that clearly should not be counted as cognitive—just think of e.g. fluctuations of cerebro-spinal fluid—will arguably bear some sort of family resemblance to typical cases of intracranial cognitive processes (if only because they both are intracranial), thereby leading to cognitive bloat again.

5. Conclusion

Regarding the question “What kind of concept could “cognition” be?” one answer would be that it is a natural kind concept. In that case, intracranialism is likely going to be true, for on any interesting version of cognitive extension, cognitive processes do not form a natural kind. Another answer would be that “cognition” is a non-natural kind concept, i.e., a concept that picks out a class of things which are causally heterogeneous at a lower (physical, in Fodor’s, or mechanistic, in Clark’s terms) level, but for which there is sufficient higher level unification for them to be worthwhile objects of scientific study. In that case, “cognition” may be a nominal kind, a cluster, or a family concept, and neither of these options is per se biased with regard to the Where of cognition. For in either of these cases cognitive processes may extend beyond the skull into the world or still be an intracranial affair only (although all three options seem particularly promising for the advocate of extended cognition). “Cognition” may also be an umbrella concept. But those who treat “cognition” as an umbrella concept have to face the Motley Crew Argument, and so an umbrella view would not be a wise choice for either the
intracranialist or the advocate of extended cognition. Given that a family view of “cognition” seems to be excessively liberal, thereby leading to cognitive bloat, and given that a cluster view is drifting dangerously close towards cognitive agnosticism, the safest bet for the advocate of cognitive extension seems to be a nominal kind view, and thus an account of cognition in terms of its form, function, or origin. As said above, there is no reason why such an account should be more likely to render cognitive processes a “foul hybrid not amenable to scientific scrutiny” than a natural kind view.

Acknowledgements

removed for anonymous review

References


AUTHOR-a (2010). REMOVED.

Barker, M. From cognition’s location to the epistemology of its nature. *Cognitive Systems Research*, ??, ??–??.


Coleman, S. (in press). There is no argument that the mind extends. *Journal of Philosophy*.


Wheeler, M. (manuscript). Extended X. Recarving the biological and cognitive joints of nature.

