

University of Wollongong

## Research Online

---

Faculty of Law, Humanities and the Arts -  
Papers

Faculty of Arts, Social Sciences & Humanities

---

1-1-2014

### Extended cognition and constitution: Re-evaluating the constitutive claim of extended cognition

Michael D. Kirchhoff  
*Macquarie University*, kirchhof@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/lhapapers>



Part of the [Arts and Humanities Commons](#), and the [Law Commons](#)

---

#### Recommended Citation

Kirchhoff, Michael D., "Extended cognition and constitution: Re-evaluating the constitutive claim of extended cognition" (2014). *Faculty of Law, Humanities and the Arts - Papers*. 1244.  
<https://ro.uow.edu.au/lhapapers/1244>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

---

## Extended cognition and constitution: Re-evaluating the constitutive claim of extended cognition

### Abstract

This paper explores several paths by which the extended cognition (EC) thesis may overcome the coupling-constitution fallacy. In so doing, I address a couple of shortcomings in the contemporary literature. First, on the dimension of first-wave EC, I argue that constitutive arguments based on functional parity suffer from either a threat of cognitive bloat or an impasse with respect to determining the correct level of grain in the attribution of causal-functional roles. Second, on the dimension of second-wave EC, I argue that especially the complementarity approach suffers from a similar sort of dilemma as first-wave EC: an inability to justify just what entails the ontological claim of EC over the scaffolding claim of weaker approaches in cognitive science. In this paper I show that two much more promising explanations by which to ground the ontological claim of EC are available, both starting from an exploration of the coordination dynamics between environmental resources and neural resources. On the one hand, I argue that second-wave EC based on cognitive integration, with its focus on bodily manipulations constrained by cognitive norms, is capable of resolving the coupling-constitution fallacy. On the other hand, I argue that the framework of cognitive integration can be supplemented by philosophical accounts of mechanistic explanation, because such accounts enable us to explain the emergence of higher-level cognitive properties due to a system's organization-dependent structure.

### Keywords

extended, cognition, constitution, re, evaluating, constitutive, claim

### Disciplines

Arts and Humanities | Law

### Publication Details

Kirchhoff, M. (2014). Extended cognition and constitution: Re-evaluating the constitutive claim of extended cognition. *Philosophical Psychology*, 27 (2), 258-283.

Please do not cite

**Extended cognition and constitution: Re-evaluating the constitutive  
claim of extended cognition**

Forthcoming in *Philosophical Psychology*

**1. Introduction**

In contemporary philosophy of mind the extended cognition (EC) thesis holds that there are conditions under which cognitive processes and cognitive systems are spatiotemporally spread across brain, body and world in such a way that extra-neuronal elements in close interplay with neuronal elements *constitute* some cognitive process or cognitive system (Clark 2008; Menary 2007, 2010; 2010abc; Rowlands 2010; Sutton 2010; Wheeler 2010; Wilson 2010). Despite the availability of weaker theses of embedded, situated, or scaffolded cognition (Robbins & Aydede 2009; Rupert 2009; Sterelny 2010), proponents of EC argue that extra-neuronal resources do more than aid, sculpt, or augment cognition: extra-neuronal resources are often constitutive elements of cognitive processing or cognitive systems. As Wilson notes: “Thus, if the extended mind thesis is true, it is true in virtue of something implementationally deep about cognition.” (2010, p. 171) Or, as Wheeler argues: “[...] the key issue facing [extended] theorists right now is not how to argue against the received (if that’s what it still is) orthodox view in cognitive science, but rather how to justify the transition from a “merely” embodied-embedded mind to an extended one.” (2010, p. 247)

My aim in this paper is to argue that the philosophical accounts of mechanistic explanation, with a focus on compositional relations, provide a fertile framework for (a) exploring the ontological basis of EC-style explanations, and (b) dissolving the “coupling-constitution fallacy” (Adams & Aizawa 2001, 2008) objection leveled against EC.

*1.1. Scope of the argument and strategy*

In the EC literature, “first-wave EC” (Clark & Chalmers 1998; Clark 2008; Wheeler 2010) is the classical version of EC based on functionalist considerations. The constitution claim of first-wave EC is commonly expressed in terms of partial constitution, where a cognitive process or cognitive system is partly constituted by environmental resources playing functionally isomorphic roles in guiding behavior as those realized by neural resources. With regards to first-wave EC, I aim to show that attempting to base the constitution claim by way of functional isomorphism between the “internal” and the “external” is problematic and therefore unlikely to work as intended (Rowlands 2010; Sprevak 2009). Before looking at “second-wave EC” (Menary 2007, 2010abc; Sutton 2010), I explore other options available to the EC theorist in order to make a case for the constitution claim. That is, the EC theorist could look to metaphysical relations as supervenience, material constitution, composition, or realization. I shall argue that the relata of material constitution and supervenience are problematic due to discrepancies between the properties of supervenience and material constitution and the properties of

Please do not cite

extended cognitive systems. In contrast to these two building-relations, the view I shall defend is that such a lack of fit does not exist between the properties of composition and realization and the properties of extended cognitive systems and processes. These latter relations of determination provide a more solid ground from which to re-evaluate EC-style constitution claims.

In addition to first-wave EC, “second-wave EC” pursues the constitution claim from both complementarity arguments (Sutton 2010) and integrationist arguments (Menary 2007, 2010abc). Even though it shows signs of immunity against the problems hindering first-wave EC, I shall argue that the complementarity view alone is insufficient to substantiate the constitution claim of EC. This is not due to flaws with the complementarity view itself, but rather results from its openness to a stalemate situation into which the debate over causal and constitutive dependence relations has fallen. Hence, it follows that for complementarity to suffice, it requires supplementation. Menary’s cognitive integrationist framework offers a better starting point from which to begin to solve the coupling-constitution fallacy. As a case in point, the cognitive integrationist does not begin from the premise that environmental resources become cognitive simply by being coupled to an already existing cognitive agent (Menary 2006, p. 335). Instead Menary aims to explain how the coordination dynamics of heterogeneously “internal” and “external” components assemble extended cognitive processes and cognitive systems. I follow Menary (see also Hurley 2010) in arguing that EC should focus less on explaining the metaphysics of EC and focus instead on explaining the empirical issue of “[...] why X and Y are so coordinated that they together function as Z [...].” (2006, p. 334)

The strategy of this paper is to argue that it is possible to supplement the cognitive integrationist framework by explaining the kinds of coordination dynamics that Menary has in mind by making use of the philosophical framework of mechanistic explanation and mechanistic composition. Doing so, I submit, generates two different but complementary theoretical models by which to approach the strong ontological claim of EC-style explanations. What is the relationship between these two approaches? On the one hand, the framework of cognitive integration establishes the constitution claim by explaining integration on a dimension of bodily manipulations constrained by socio-cultural cognitive norms, whereas mechanistic explanation targets the emergence of cognitive properties on the systemic dimension of mechanistic organization. On the other hand, both approaches focus on providing sub-personal and mechanistic descriptions of cognitive processes and the formation of cognitive systems.

## **2. Coupling-constitution fallacy:**

I begin by exposing the coupling-constitution fallacy (C-C fallacy). According to several critics of the thesis of EC (Adams & Aizawa 2001, 2008, 2009, 2010; 2010a; Rupert 2004; 2009), the most common and pervasive mistake EC-theorists make (Adams & Aizawa 2010, p. 68) is to tacitly move “from the observation that a process X is in some way causally connected (coupled) to a cognitive process Y to the conclusion that X is part of the cognitive process Y.” (Adams & Aizawa 2009, p. 81) Note that in this passage the C-C fallacy concerns cognitive processes. However, the fallacy is not limited to processes. In the following passage, Adams & Aizawa criticize Clark for making the fallacy on the level of cognitive processes and also the level of entire cognitive systems:

## Please do not cite

“When Clark makes an object cognitive when it is connected to a cognitive agent, he is committing an instance of a “coupling-constitution fallacy.” This is the most common mistake that extended mind theorists make. The fallacious pattern is to draw attention to cases, real or imagined, in which some object or process is coupled in some fashion to some cognitive agent. From this, one slides to the conclusion that the object or process constitutes part of the agent’s cognitive apparatus or cognitive processing. If you are coupled to your pocket notebook in the sense of always having it readily available, use it a lot, trust it implicitly, and so forth, then Clark infers that the pocket notebook constitutes a part of your memory store.” (2010, pp. 67-68)

It follows, according to the critics, that one is committing an instance of the C-C fallacy when one conflates observations about an individual cognitive agent’s causal interactions with extra-neural resources or observations about dynamical coupling relations between an individual cognitive agent and parts of the extra-neural world, with those extra-neural elements being constitutive parts of an individual’s cognitive process or cognitive system. Before looking at the different constitution arguments in first- and second-wave EC, it is important to be clear about what role the appeal to causal interaction or causal coupling is intended to play in the argument for EC. Both Clark (2008, p. 87) and Menary (2010c, p. 608) are particularly clear about this. As Clark says, the “appeal to coupling is not intended to make any external object cognitive [...]. Rather, it is intended to make some object, which in and of itself is not usefully [...] thought of as *either cognitive or noncognitive*, into a *proper part of some cognitive routine*.” (2008, p. 87; italics in original) So it immediately appears as if the EC theorist can counter at least one aspect of the C-C fallacy, namely the charge that EC is guilty of making an object cognitive just by the object being coupled to a cognitive agent. However, the EC theorist has yet to counter Adams & Aizawa’s claim that any inference from coupling to the conclusion that “external” elements are constitutive parts of an extended cognitive system or cognitive process is an instance of the C-C fallacy. It is to these arguments that I now turn.

### 3. First-wave EC & extended functionalism

First-wave EC aims at grounding the constitution claim on functionalist considerations. In “In Defense of Extended Functionalism” (2010), Wheeler provides the following argument concerning whether “external” elements qualify as proper parts of an extended cognitive process or system:

1. “If psychological phenomena are constituted by their causal-functional role, then our terms for mental states, mental processes, and so on pick out equivalence classes of different material substrates, any one of which might in principle realize the type-identified states or process in question.” (Wheeler 2010, p. 248; compare Clark 2005, p. 2)
2. “If there is functional equality with respect to governing behavior, between the causal contribution of certain internal elements and the causal contribution of certain external elements, and if the internal elements concerned qualify as the proper parts of a cognitive trait, then there is no good

Please do not cite

reason to deny equivalent status - that is, cognitive status - to the relevant external elements.” (Wheeler 2010, p. 248; compare Clark 2008, p. 50)

3. If parity of causal contribution mandates parity of status, and if mental states and processes are multiply realizable, then “it is possible for the very same type-identified cognitive state or process to be available in two different generic formats - one non-extended and one extended.” (Wheeler 2010, p. 248)
4. *Therefore*: Cognitive processes and states are realizable (partly at least) by external states and processes (Wheeler 2010, p. 249; compare Clark 2008, p. 39).

The logical form of this argument follows from the functionalist parity principle (PP) originally introduced by Clark & Chalmers (1998):

“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.” (1998, p. 2)

To make sense of Wheeler’s argument and the PP, I shall revisit Clark & Chalmers’ original case of Otto & Inga (1998, pp. 12-14). In a similar vein as Wheeler, Clark & Chalmers invite us to consider the idea that a cognitive process, namely a standing belief about the location of the Museum of Modern Art (MoMA) in New York, could be instantiated in two different generic formats. In the example, Otto & Inga hear about an art exhibition at MoMA. After hearing the news about the exhibition, Inga thinks about the location of MoMA, remembers that it is on 53<sup>rd</sup> St., and sets off. Otto, on the other hand, suffers from a mild form of Alzheimer’s and is therefore unable to reliably use his biological memory to recall the address. Fortunately for Otto, he always – as a compensatory strategy – writes down useful information in a notebook. Just like Inga, Otto hears about the exhibition, but quite unlike Inga, Otto consults his notebook, retrieves the address, and sets off. Moving from the philosophical theory of functionalism, which holds that psychological states like beliefs are constituted not by way of their material makeup but rather in terms of their causal-functional role in generating appropriate behavior, the result that the proponents of first-wave EC want us to accept is that type-identified mental states are available in both a non-extended and an extended format.

### *3.1. Problems with first-wave EC & extended functionalism*

Unlike the standard anti-EC arguments (Rupert 2004), where it is commonly argued that the constitution claim of EC fails because of the obvious distinctiveness between the fine-grained profile of “internal,” neural operations and the profiles of “external,” sociocultural operations, the particular kind of arguments that I shall explore in this section do not revolve around this kind of skepticism. I agree with Menary (2010, p. 5) and Wheeler (2010, p. 248) that it is incorrect to interpret the PP as entailing any fine-grained functional similarity between the properties of “internal” and “external” mechanisms and their operations. Instead a satisfactory response to the kind of objections that I shall deal with begins by acknowledging that functionalism was engineered, in part, so as to avoid chauvinistic exclusion of cognition or mentality from individuals (or organisms) with differences in fine-grained

Please do not cite

causal-functional profile or in fine-grained physiological makeup (Block 1980).

### 3.1.1. Sprevak's Martian intuition

Sprevak (2009) presents first-wave EC with an intriguing dilemma: (a) accept functionalism and radical EC; or, (b) give up EC entirely (2009, p. 503). The point that Sprevak wishes to make is that if one accepts functionalism, then one is committed to a radical and implausible version of EC, where the price of taking EC on board is "rampant expansion of the mind into the world [...]" (2009, p. 503) Here is the form of Sprevak's argument:

1. Functionalism entails the Martian intuition (P).
2. If P, then radical EC (Q).
3. P.
4. *Therefore*, Q.

So, if functionalism entails the Martian intuition and if EC builds its constitutive claim on functionalism, then first-wave EC is committed to a radically implausible version of EC where the following scenario holds: simply by picking up a book, one comes to believe all the information contained in that book. In other words, everything stored in that token book is necessarily a part of my cognition (or mind). Following Sprevak's argument, the justification for this (wild) claim is as follows: (a) a Martian might "internally" encode memories in ink-marks; (b) in addition to gaining its beliefs via sense modalities, such a Martian might equally be born with innate beliefs; (c) moreover, it is possible that the Martian might have such innate beliefs that it has not yet examined, *viz.*, that the Martian has a library of data phylogenetically hardwired into its cognitive system; and (d) finally it is possible to imagine that this Martian has such a stock of innate beliefs stored in an ink-based memory system, most of which it has not yet had any reason (or cause) to employ. Sprevak's point is that it quite plausible to think that such a creature could exist. As Sprevak says, the:

"Martian has ink-marks inside its head that, if it were sufficiently diligent, would guide its action in appropriate ways. The difference between the Martian and me is that it has the ink-marks inside its head, while I have the ink-marks outside. By the fair-treatment principle, if the Martian has beliefs, then so do I." (2009, p. 518)

If Sprevak is correct, basing the constitution claim on the PP would seem to license an overly permissive and implausible attribution of cognitive processes or states. Hence, it is not clear that the PP can be used as a principle of demarcation between what is and what is not a proper part of an extended cognitive system or process. One might attempt to augment the PP by reference to the so-called "glue-and-trust conditions" (Clark & Chalmers 1998). The glue-and-trust conditions state that external resources might count as part of cognition only if such resources are (a) reliably available and typically invoked, (b) more-or-less automatically endorsed, and (c) that the information carried in the resource be easily accessible when required (1998, p. 17). According to Sprevak, these conditions will not save

Please do not cite

first-wave EC, since it is not necessary for an environmental resource to be part of a cognitive system or process that such a resource fulfill (a)-(c). Here I restrict myself to condition (a). On Sprevak's view, it is possible to imagine a Martian, whose cognitive resources are only available after a good night's sleep, and that this Martian does not reliably or often get a good night's sleep. However, "that does not stop, on those occasions when the Martian does get a good night's sleep, those resources from counting as genuinely cognitive." (Sprevak 2009, pp. 515-16) So it would seem that the glue-and-trust conditions would not save first-wave EC from the negative ramifications entailed by the Martian intuition.

### 3.1.2. Rowland's impasse

But should we just accept Sprevak's argument? One response that is available to first-wave EC is to go between the horns of the dilemma, by arguing that the relevant level of grain by which we "individuate functional roles should be set neither too high (so as not to entail radical EC) nor too low (so as to block the critics' difference-argument)." (Anonymous referee) Wheeler (2010) takes up this line of response. Note that this is not the only way the EC theorist could attempt to deflate Sprevak's argument. Second-wave EC seems immune to the Martian intuition. For instance, on Menary's version of cognitive integration, it is not functional similarity that matters, but rather how the coordination between functionally *dissimilar* "inner" and "outer" elements are integrated to establish deeply *hybrid minds* (Menary 2006, p. 333). As Theiner has recently said about second-wave EC, it is one "which is completely devoid of this questionable presupposition [functional similarity between "inner" and "outer], and thus affords a logically independent path to the claim that cognition is extended." (2011, p. 52) However, the question that shall entertain us here is whether Wheeler's case holds water?

Rowlands has argued that establishing the relevant level of granularity with respect to functional roles – between those who accept EC and those who do not – leads to an impasse, the ramifications having a paralyzing affect on the entire debate over cognitive distribution:

"If Rupert's arguments against the extended mind are question-begging because they presuppose a chauvinistic form of functionalism, it is difficult to see why arguments for the extended mind are not question-begging given their predication on a liberal form of functionalism." (Rowlands, unpublished ms, pp. 6-7; quoted from Wheeler 2010, p. 255; see also Rowlands 2010, pp. 209-10)

In what we might call Rupert-style anti-EC (2004, 2009; see also Adam & Aizawa 2001, 2008), the most common way of criticizing extended functionalism consists in noting a set of psychological properties found in human cognitive systems but not socio-cultural systems, and then inferring that there is no parity at the level of fine-grained functional operations between "inner" and "outer". Hence, EC must be false. The proponents of first-wave EC commonly respond to this line of argument by charging Rupert-style anti-EC of advocating a chauvinistic form of functionalism; one that privileges the neuronal innards solely on their fine-grained psychological profile. As Wheeler says, "it seems that Rupert's [...] argument continues to beg the question against extended functionalism. [...], extended functionalism looks to be predicated on the more liberal form of functionalism that generates a locationally uncommitted account of the cognitive." (2010, p. 255) It is from the problem of identifying



Please do not cite

the appropriate level of functional grain that Rowlands concludes that the debate over EC has fallen into an impasse, with both sides potentially begging the question against one another

### 3.1.3. Theiner's rule of extended fairness

The trouble confronting Wheeler's reply to the Martian intuition is not only limited to the debate over extended functionalism. As Rowlands points out (2010, p. 210), the problem of identifying the correct level of grain is an old problem for functionalist theories of mind, the solution being far from readily available. One could argue, with the aim of saving Wheeler's strategy, that just because a "compromise-solution" on how to set the benchmarks for functional grain have plagued philosophical functionalism in the past (and still does), this does not present a lethal problem for extended functionalism. Theiner endorses such a position. In his (2011), Theiner introduces what he calls the rule of "extended fairness": "Don't burden the theory of the extended mind with problems that equally affect theories of the non-extended mind." (2011, p. 34) According to Theiner, both Sprevak (2009) and Rowlands (unpublished ms) disregard this rule precisely because they are pointing to flaws that not only affect EC but equally non-extended versions of functionalism. I agree with Theiner on this issue: the fact that this problem poses "an equal challenge for internalist and extended versions of functionalism shows that the problem cannot be specific to the HEC (rather than functionalism *simpliciter*)." (2011, p. 35; italic in original) However, this response is not completely satisfying because the proponent of Rupert-style anti-EC can force the issue, "and demand a further reason for why we should consider the external resources as part of a single, transcranially extended *cognitive process*." (Theiner 2011, p. 56) Hence, even if "it would be premature to declare that the debate over extended functionalism suffers from a lethal kind of indeterminacy [...]," (Theiner 2011, p. 49) it leaves the debate between Rupert-style anti-EC and EC unresolved.

## 4. Causation, material constitution, supervenience, composition and realization!

The point of the preceding section has been to present evidence suggesting that basing the strong ontological claim of EC by appealing to the functionalist PP of first-wave EC is problematic. In this section, I consider a variety of other options available to the EC theorist in making a case for the constitution claim. I begin by giving a brief sketch of the causation relation. This is done to accommodate Menary's claim that an "account of the difference between causation and constitution would be helpful here, but there is none forthcoming [...]" (2010, p. 607)

### 4.1. Causation

!

Causation is a temporal and asymmetric relation between cause and effect: if  $y$  is caused by  $x$  at  $t$ , then  $x$  do not spatiotemporally overlap with  $y$ . In a slogan: causes precede their effects (Craver & Bechtel 2007, p. 552; Shapiro 2010, pp. 159-61). As such, a causal relation between  $y$  and  $x$  is such that the space-time path of  $y$  is distinct from the space-time path of  $x$ . For instance, to make sense of why the

Please do not cite

constitution relation does not entail a relation of causation, Rowlands introduces the following example (originally due to Davidson 1987):

“*Dependence*, even *essential* dependence, does not add up to *constitution* – not without a lot of argument. Sunburn is (essentially) dependent on solar radiation in the sense that any skin discoloration not produced by solar radiation is not sunburn. But this does not mean that solar radiation is literally part of – a component of – sunburn.” (2010, p. 56; italics in original)

It follows that solar radiation causes sunburn but does not constitute sunburn. At no point in time does solar radiation either spatially coincide or materially overlap with sunburn. As Rowlands intends us to understand the sunburn-example, sunburn is individuation-dependent on its causal etiology, since being overexposed to incident UV-radiation can cause sunburn. But it does not follow that solar radiation is literally part of the sunburnt skin. Contrast this initially with material constitution: if  $x$  takes place prior to  $y$ , or if  $x$  takes place apart from  $y$ , or if  $x$  and  $y$  reciprocally influence each other, then  $x$  causes, but does not constitute,  $y$  (Shapiro 2010).

#### 4.2. Material Constitution

Material constitution is a non-causal and synchronic (atemporal) relation of ontological determination between spatially and materially co-located objects, processes or properties of different or similar kind (Bennett 2011). The constitution relation can be framed in terms of how to fill out the following schema:  $x$  constitutes  $y$  at  $t$  if and only if \_\_\_\_\_? (Wasserman 2004, p. 694). In analytical metaphysics, the philosophical home of constitution debates, there is some dispute about how to adequately fill out this schema. However, it is widely accepted that  $x$  constitutes  $y$  if and only if the constitution relation upholds the following coincidence and formal conditions. There are two coincidence conditions. First, constitution requires *spatial coincidence*:  $x$  constitutes  $y$  at  $t$  only if  $x$  and  $y$  have the same spatial location at  $t$ . Second, constitution requires *material coincidence*:  $x$  constitutes  $y$  at  $t$  only if  $x$  and  $y$  share all the same parts at  $t$  (Wasserman 2004, p. 694; Wilson 2007, p. 5). In conjunction with these conditions of coincidence, it is widely accepted that the relation of constitution respects the following three formal conditions. First, the constitution relation is *transitive*: if  $x$  constitutes  $y$ , and if  $x$  (or the  $x$ s) is (are) constituted by  $z$  (or the  $z$ s), then  $y$  is constituted by  $z$  (or the  $z$ s). Second, the constitution relation is *irreflexive*:  $x$  constitutes  $y$  and neither  $y$  nor  $x$  constitute themselves. Finally, the constitution relation is *asymmetric*:  $x$  constitutes  $y$ ,  $y$  does not constitute  $x$  (or the  $x$ s). According to Wasserman (2004, p. 695; see also Baker 2000; Lowe 1989), these conditions are minimal requirements offered by constitution theorists.

We are now able to highlight a few important features distinguishing causation from material constitution. First, both material constitution and causation are asymmetric relations of dependence. However, unlike the asymmetry of causation, the property of asymmetry on material constitution relations is a “one-to-one” relation of determination. That is to say, when  $y$  is caused by  $x$  and time  $t_1$ ,  $x$  necessarily precedes  $y$  on a temporal dimension. So on a causal relation, the cause ( $x$ ) and the effect ( $y$ )

Please do not cite

do not spatiotemporally overlap. This is different on the relation of material constitution, since here the asymmetry is to be understood as follows: if a piece of clay constitutes a token statue, the piece of clay and the statue overlap in a spatiotemporal sense. This is why relations of material constitution – and not causal relations – are one-to-one relations of determination. Second, the synchronic (or atemporal) nature of material constitution sets it apart from causation, for while cause and effect are independent events in time, the relation between constituent(s) and constituted is simultaneous.

There are several problems with this relation of determination at least insofar as it is employed in helping EC justify its constitution claim. The first problem is ontological in the sense that material constitution requires material coincidence between parts and whole:  $x$  constitutes  $y$  at  $t$  only if  $x$  and  $y$  share all the same parts at  $t$ . However, this condition does not fit with the nature of an extended cognitive system. One objection to this claim is to argue that the structured collection “of the relevant parts of” brain, body, and world which are presumed to materially constitute an extended cognitive system at time  $t$  have all the same parts of the extended cognitive system at time  $t$  – viz., the extended cognitive system exists at time  $t$  wherever its physically constitutive parts are at time  $t$ . I am sympathetic to this claim, since it is indeed the case that an extended cognitive system at time  $t$  is wherever its component parts are at time  $t$ . Moreover, it likely does not follow that all (mereological) parts,  $P$ , of the physical objects partly constituting, at time  $t$ , an extended cognitive system  $S$ , are equally parts of  $S$ . The point that I wish to highlight, though, is different: the material coincidence condition does not mesh with the *hybrid* character of extended cognitive systems, since even if  $y$  shares all the parts of the  $x$ s, this does not hold at the level of the constituents. The  $x$ s do not share the same parts. In addition to this, the  $x$ s are temporally and spatially distributed from one another. Hence, unlike the relation of clay and statue, the dynamics and characteristics between the larger extended cognitive system and its parts do not correspond with material constitution. The second problem is formal. Material constitution is a non-causal, asymmetrical relation of ontological determination. But, the empirical fundament of EC is inherently causal and symmetrical, in the sense that the empirical basis of an extended cognitive system is formulated in terms of a dynamically coupled, co-determining relation between agent and ecological niche (Beer 2000; Clark 1997; Van Gelder 1995).

#### 4.3. Supervenience

I include supervenience here because proponents of EC refer to this kind of relation in order to ground the constitution claim. For instance, Wilson & Clark claim: “we should treat the nonbiological augmentation as part of the material supervenience base for some of Otto’s long-term, nonoccurrent, dispositional beliefs [...]” (2009, p. 66) Supervenience, just as material constitution, is a non-causal and atemporal relation between, e.g., mental properties and properties of physical processes. But unlike material constitution, there is a general consensus that the supervenience relation does not guarantee a relation of ontological dependence (McLaughlin & Bennett 2005). As such, a supervenience relation does not necessarily guarantee a relation of constitution. This fact in itself makes the appeal to supervenience problematic as an attempt to underpin the constitution claim of EC. In all fairness, Wilson & Clark mention other forms of ontological dependence relations – in particular, Wilson’s

Please do not cite

(2001, 2004, 2005) theory of “wide-realization.” I discuss this in the next section. Here I shall keep focusing on the relation of supervenience. According to McLaughlin & Bennett (2005), all a relation of supervenience establishes *prima facie* is a relation of *covariance* between, e.g., mental properties and physical properties. Moreover, unlike material constitution, supervenience is *reflexive*. There can be no variation on mental properties (M) unless there is variation in physical properties (P), and vice versa, given that  $M = P$ . Finally, in contrast to material constitution, the supervenience relation is *non-symmetric*, i.e., the supervenience relation alters between symmetric and asymmetric instantiations (McLaughlin & Bennett 2005).

#### 14.4. Composition

Composition is a non-causal and synchronic relation of ontological dependence and it holds between heterogeneous elements, e.g., between a mechanism and its complementing parts (Craver & Bechtel 2007, p. 547). Again we can frame the composition relation in terms of how to fill out the following schema:  $x$  (or the  $x$ s) compose(s)  $y$  at  $t$  if and only if \_\_\_\_\_? (Hawley 2006, p. 483). Unlike material constitution, where something is a material constitution relation if and only if it upholds certain *coincidence* condition, compositional relations are less demanding in that they require only *coexistence* conditions. First, composition requires only *spatial coexistence*: the  $x$ s compose  $y$  at  $t$  only if  $y$  as a whole share the same space-time path as the  $x$ s and no two of the  $x$ s occupy an overlapping space-time path (Hawley 2006, p. 483). In the case of extended cognitive systems, the system as a whole coexists spatially with the mereological sum of the parts, and none of the parts occupy overlapping regions of space. On a first pass, then, an extended cognitive system exhibits the property of spatial coexistence. Second, composition requires *material coexistence*: the  $x$ s compose  $y$  at  $t$  only if  $y$  as a whole is composed of the material parts of the  $x$ s and no two of the  $x$ s materially overlap in terms of their parts. This fits the nature of an extended cognitive system, in the sense that EC theorists often conceive of such systems as collectively composed of disparate parts, none of which materially overlap. Hence, the component parts of an extended cognitive system occupy non-identical spatial locations within *one* distributed cognitive system.

To make it clear which kind of compositional relation I have in mind here, I shall adopt Craver’s (2007) taxonomy distinguishing between four different notions of composition: mereological, aggregative, spatial/material, and mechanistic composition. The notion of composition I shall support, and the one lending support to an explanation of the emergence of higher-level cognitive properties is *mechanistic composition* (Theiner 2009). First, consider the “extensionality theorem” inherent to the mereological relation of composition: “an object is completely determined by the set of its parts.” (Craver 2007, p. 185) In contrast to mechanistic composition, where emphasis is on explaining emergent higher-level phenomena by reference to the organization-dependence of the enabling system, the extensionality theorem implies the structural organization of the systemic components does not matter as an ontological principle. Second, on the aggregativity relation of composition, “the relata are properties of wholes and the properties of parts, and the relation between them is that higher-level properties are sums of lower-level properties.” (Craver 2007, p. 186) However following Wimsatt (1974, 1986), Theiner (2009, 2011) has recently argued that the emergence of complex properties in

Please do not cite

extended cognitive systems is a result of that system failing to exhibit aggregativity in its structural composition (Theiner & O'Connor 2010, p. 84). Hence, aggregate systems lack the required integration between its component parts to be explanatorily useful in addressing EC's constitution claim. Third, the relation of material and spatial composition equally fails to capture the nature of extended cognitive systems. For instance, as Craver notes, "thinking of levels in this way does not allow one to distinguish between mere *pieces* and its *components*." (2007, p. 187; italics in original) Decomposing a system into mere spatial and material pieces does not (and will be highly unlikely to) guarantee decomposition into the systemic components, i.e., elements "that make identifiable contributions to the behavior of a mechanism." (Craver 2007, p. 188) Fourth, Craver defines the relation of mechanistic composition as follows: "X's  $\Phi$ -ing is at a lower mechanistic level than S's  $\Psi$ -ing. Lower-level components are *organized together* to form higher-level components." (2007, p. 189) Consider, for instance, the compositional layout of a synapse. A synapse is composed of part of a pre-synaptic cell, part of a post-synaptic cell and of the gap between the pre- and post-synaptic cells. According to Craver, "what unifies these items into a component is their organized behavior: the pre-synaptic cell releases transmitters that traverse the cleft and act on the post-synaptic cell." (2007, p. 190) Mechanistic composition, to be spelled out in further detail in the section "Mechanistic compositional organization," supplements an explanation of the organization of extended cognitive systems, underpinning the claim of deeply *integrated* and *hybrid* cognitive systems.

#### 4.5. Realization

The relation of realization has played some role in the debate over EC, predominantly in Wilson's theory of "wide realization" (1994, 1995, 2004, 2005). According to Wilson, realization is usually taken to imply that physical realizers are "determinative" of the properties they realize and "physically constitutive" of the intrinsic states of the individuals with those properties (2001, p. 1). The determinative part of realization, Wilson calls the *metaphysical sufficiency thesis*: "Realizers are metaphysically sufficient for the properties or states that they realize." (2004, p. 103) The physically constitutive part of the standard view, Wilson refers to as the *physical constitutivity thesis*: "Realizers of states and properties are exhaustively physically constituted by the intrinsic, physical states of the individual whose states or properties they are." (2004, p. 104) Central to Wilson's argument is that the conjunction of the metaphysical sufficiency thesis and the physical constitutivity thesis is false, since a "metaphysically robust notion of realization is ineliminably *context-sensitive*." (2001, p. 1; italics in original) As a case in point, consider Rumelhart & McClelland's (1986) classical example of using pen and paper to complete a mathematical algorithm. According to Wilson & Clark (2009), what initially starts as causal interactions between neural representational or informational processes and cultural representational or informational processes become "incorporated as part of the computational process itself." (Wilson & Clark 2009, p. 10) This case shows that on some occasions cognitive properties "have realizers that extend beyond the individual instantiating them." (Wilson 2004, p. 107)

Generally there is much to like about the idea of wide realization, since it fits with the ontology of extended cognitive systems and processes. But, it also seems evident that the project of wide

Please do not cite

realization turns on its *empirical* substantiation before its metaphysical one. Hence, to avoid begging the question about where to place the causal-constitutive boundary (Hurley 2010), we can launch a wide realization argument for EC *only after* having empirically based the validity of EC. It is this project that the philosophical account of mechanistic explanation can help with.

## 5. Second-wave EC and Complementarity

Kirchhoff (2011) has recently shown – in his exploration of a distinctive third-wave version of EC – that second-wave EC is quite unlike EC based on the PP and extended functionalism. The big question of this section is: does second-wave EC provide the fodder required to justify its constitution claim? The argument that I shall pursue is as follows: complementarity alone is insufficient to ground the constitution claim of EC essentially because complementarity is open to a kind of stalemate scenario into which the debate over causal relations and constitutive relations has fallen. I am not the only one skeptical of the complementarity framework. Theiner (2011) has recently argued that second-wave EC – which he refers to as cognitive integration – does not imply EC unless second-wave EC is augmented with some version of the PP (see also Rowlands 2010). However, unlike the original formulation of the PP, which is pitched at the level of resources, Theiner argues that the PP should be applied to the level of capacities from which one can then individuate entire systems or processes as cognitive. I agree with this, since my conclusion points in a similar direction, namely that the second-wave EC proponent still needs an argument for why this particular system as a whole is a system of type Z (where Z = cognitive). Having discussed the complementarity framework, I shall then explore the main argument of this paper.

### 5.1. The Complementarity Argument

Psychologists (Donald 1991) and philosophers of cognitive science (Sutton 2010) inform us that standard external memory records, or exograms, like those found in Otto's notebook, are stored in a discrete fashion (e.g., like the words 'milk', 'meat', and 'fruit' on a grocery list). That is, their representational schemes are either linguistic or pictorial (or maybe both). They exhibit no dynamics or activity (for instance, they are not prone to primacy or recency effects). Nor are they intrinsically integrated with other memory records (e.g., they are not subject to negative transfer effects). Inga's biological memory, in contrast, may well blend and interfere (Sutton 2010, p. 197), and is vulnerable to degradation (Donald 1998, p. 15). Moreover, unlike encodings in exograms, biological memory exhibits such dynamic features such as informational holism (Clark 1989, p. 107). Additionally, the representational formats of biological memory are composed and stored in the connection strengths between the units of neural nets as exemplified in artificial neural networks, which is very different from symbolic and pictorial representations. In accordance with complementarity, it is precisely this disparate nature – this deep mechanistic dissimilarity – between neural and extra-neural elements that explains why agents accomplish cognitive tasks that the 'naked brain' alone could not achieve. As Sutton explains: "Brains like ours need media, objects, and other people to function fully as minds

Please do not cite

[...]: biological traces are typically integrative, active, and reconstructive, but in using them we hook up with more enduring and transmissible exograms, mostly of our own making, with supplement and extend our powers.” (2010, p. 205) However, the complementarity framework still faces a worry that we now ought to expose.

!

### 5.2. *The stalemate problem*

!!! !

Here is the argument in its conceptual form (for a similar argument see Rowlands 2010, pp. 90-91): An inherent assumption of the C-C fallacy is that even if second-wave EC requires essential differences between “internal” and “external” processes, this by itself is insufficient to justify the inference that the latter has cognitive status – the reason being that it remains an open empirical question whether the “external” parts merely make up an incredibly important array of extra-neuronal scaffolding in which what constitutes real cognition is internally, neurobiologically generated (Rupert 2010, p. 348). Hence, even if Otto and his notebook differ importantly in their fine-grained dynamics, it does not follow that the notebook in virtue of being coupled to or used by Otto-the-agent becomes part of Otto’s cognitive processing rather than merely augmenting Otto’s capacity to perform a cognitive task, e.g., remembering. However, as Theiner mentions, this will not scare the complementarity theorist, since she/he will typically respond to this charge by “claiming that what makes the external parts of the transcranial process as cognitive as its internal parts is the fact that they all make *some* distinctive, functionally indispensable causal contribution to the exercise of a cognitive capacity.” (2011, p. 56) It is this move that makes the (potential) stalemate emerge, since now there seems to be no way of resolving the question whether “external” elements merely causally contribute – perhaps in some deep way – to cognitive processes and systems or whether these elements in part constitute cognitive processes and systems. Given that both anti-EC and EC can accommodate the view that there are uniquely and important differences between “internal” and “external” processes (Adams & Aizawa 2001; Clark 1998; Rupert 2004; Sutton 2006, 2010), why should we prefer a “constitutive” over a “scaffolding” view of external, cognition-enhancing resources? If this kind of indeterminacy poses a real problem for second-wave EC, then it would seem that complementarity itself does not suffice in order to substantiate the constitution claim. This does not entail that the EC theorist should reject a commitment to complementarity; rather, it indicates that complementarity needs to be supplemented further. This I take to be the job of Menary’s cognitive integrationist model.

## 6. **Second-wave EC and cognitive integration**

Following Menary, I shall argue, firstly, that the focus of cognitive integration on the dimension of manipulation and cognitive norms establishes one way that EC can ground its ontological claim. In his *Dimensions of Mind*, Menary analyzes “epistemic actions” – involving a reciprocal coordination between the embodied organism and its cognitive niche – so as to support the strong ontological claim of EC. Epistemic actions are examples of Menary’s “manipulation thesis,” (MT) which he defines as:

## Please do not cite

“The manipulation thesis as a constituent thesis of cognitive integration is first understood to be an embodied engagement with the world, [...]. Secondly it is not simply a causal relation, bodily manipulations are also normative – they are embodied practices developed through habit and training and governed by cognitive norms.” (Menary 2007, p. 84)

As a constitutive thesis of cognitive integration, the MT establishes just why Menary is not pursuing a metaphysical conception of constitution but rather an empirically sensitive understanding of the ontological basis of extended cognitive systems and processes. The MT explains how cognitive processes and systems are assembled by both habituated bodily patterns of actions, *viz.*, by way of body schemas (Gallagher 2005), and the cultural practices that govern such actions (Hutchins 2008). The coordination of brain, body and environmental resources is governed by cultural and social norms and practices. Hence, the MT begins from the perspective that “we are actively embodied in a socially constructed cognitive niche [...] and this has led to the development of hybrid cognitive systems where the bodily manipulations of vehicles [...] in the niche involves the coordination of neural, bodily, and environmental vehicles.” (2010c, p. 611) To make this a bit more concrete, let us consider how Menary defends the MT by reference to Kirsh & Maglio’s (1994; see also Kirsh 1995; 2009) work on epistemic actions in the case of Tetris.

Epistemic actions are a class of actions, which make mental computation (processing) “easier, faster, or more reliable.” (Kirsh & Maglio 1994, p. 513) In accordance with the MT, epistemic actions have this cognitive consequence, because the underlying components of epistemic actions are spread across neural, bodily, and socio-cultural resources. Instead of relying purely on “internal” processing, Kirsh & Maglio argue that actively manipulating structures in the environment alters the informational structure of the immediate environment, thereby reducing the need for too much “internal” processing. In their study of Tetris, it turns out that expert players, by physically manipulating the falling zoids on the screen rather than rotating the zoids mentally, can rotate a zoid 90° in only 100 ms, where the process of mental rotation takes circa 800-1200 ms (Kirsh & Maglio 1994, p. 514). Epistemic actions thus allow a player (or in the more general case, a cognizer) to reap important time-related benefits. As Menary argues, the actions that Tetris players perform directly transform the informational structure of the environment (Menary 2010b, p. 566). On this view, epistemic actions are literally part of the problem solving process. As Kirsh & Maglio state about their approach – “its chief novelty lies in allowing individual functional units inside the agent to be in closed-loop interactions with the outside world.” (1994, p. 542) Or, as Menary claims: “Epistemic actions and computations take place within the same state space. If this state space is the problem-solving state space, then we have difficulty in pulling apart mind, action and world [...]” (2007, p. 90) One might object that without further argument the notion of “within the same state-space” is somewhat ambiguous. However, this would be to misinterpret the explanation given by Menary (2010b). By “state-space” it is clear that Menary means “information processing space,” which “includes both processes in the head, and processes outside of the head.” (2010b, p. 567) Hence, the hypothesis of an extended state space entails the hypothesis of an extended information processing space in which problem solving occurs. Moreover, and this is an aspect unique to the approach offered by Menary, public vehicles manipulated during epistemic actions have their own *norms* “governing how we are to manipulate token representational



Please do not cite

vehicles.” (2010b, p. 570) On Menary’s view, human cognitive capacities are shaped by the normative and social structure embodied in our recurrent socio-cultural practices, e.g., through participation in the practice of Tetris playing, the players have developed motor programmes for fast and fluent actions so as to transform the zoids on the screen. But, “the manipulation of the buttons is itself a normative practice, something that is learned and habituated.” (2010b, p. 575) Cognitive norms, therefore, are socio-culturally constructed, and these environmental vehicles take precedence over their sensorimotor or higher-order instantiations – or are first in “public space” before they become internalized into “neural space.” What this case study shows is that cognitive integrationism avoids the C-C fallacy, since the capacity for engaging in epistemic actions are not first “internal” to the organism and then somehow are extended to include elements in the environment. Rather, epistemic actions are “in the problem solving state space, not just as a clever strategy for off-loading complexity onto the environment [...], (2010b, p. 568) but as “part of our cognitive economy.” (2010b, p. 568)

## 7. Mechanistic compositional organization

The second dimension of the ontological claim supplements the MT by explaining the emergence of cognitive properties from the relation of composition and mechanistic explanation<sup>1</sup>. In their seminal paper, Machamer and colleagues give this definition of a mechanism:

“[...] mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination.” (2000, p. 3)

Mechanisms are typically composed of different kinds of entities ( $x$ 's) performing different kinds of activities ( $\Phi$ 's), with the  $x$ 's and  $\Phi$ 's organized such that they cooperatively produce a specific kind of behavior or performance of an overall mechanism ( $S$ 's  $\Psi$ -ing). Generally, if we understand a cognitive system as a mechanism, extended or non-extended, a cognitive *system* has properties that are different from the properties of its individual parts. For instance, a system distributed across brain-body-world has different properties in regards to either the brain-body system or the brain or body (Hutchins 2011, p. 425). Complex properties of a mechanism, then, exhibit what Theiner calls a strong form of

---

<sup>1</sup> There is a slight disagreement in the mechanistic literature about the extent to which one can apply the mechanistic framework to cases of EC. In particular, Craver argues that “many cognitive mechanisms draw upon resources outside of the brain and outside of the body to such an extent that it is not fruitful to see the skin, or surface of the CNS, as a useful boundary [...]” (2007, p. 141). However, Bechtel (2009) thinks that mechanism-style explanations are more compatible with weaker embodied and situated approaches to cognition. Note that Bechtel does not think that this necessarily must be this way. As he mentions (personal communication), what is critical for a mechanistic explanation is first to identify the phenomenon for which one is seeking an explanation. If the system accomplishing the phenomenon functions as a result of causal interdependence between an individual and artifacts, then that is the place to locate the explanation. But, “my bet is that there will remain the task of explaining how the organism engages their environment, and that this phenomenon is to be explained inside the head and these will resemble the traditional explanations in cognitive psychology” (personal communication). This emerging picture between proponents of mechanistic explanation as to whether the framework can be applied to extended and distributed cognitive science is evidence, it seems to me, of an important fact, namely that the ideas of EC is being discussed outside the mainstream debates over EC. Moreover, as I have already mentioned, Theiner (2009) as well as Theiner et al. (2010) apply mechanistic explanation to group cognition, and Theiner (2011) appeals to the framework of mechanistic explanation to resist the C-C fallacy in the context of his discussion of extended reasoning.

Please do not cite

organization-dependence (2009, p. 335). Wimsatt (1976, 1986) has argued that one way of determining the relevancy of the parts in a mechanism with respect to *S*'s  $\Psi$ -ing is to contrast aggregative and non-aggregative modes of composition. The key in this distinction is between systems *whose parts are interchangeable* (i.e., aggregative) and systems *whose parts are not interchangeable* (i.e., non-aggregative or mechanistic). Craver (2001), following Wimsatt, defines an aggregate system as follows (2001, pp. 58-59): "Suppose that a property  $\Psi$  of the whole *S* is a function of the properties  $\{\Phi_1, \Phi_2, \dots, \Phi_n\}$  of the parts  $\{X_1, X_2, \dots, X_n\}$ . Then a  $\Psi$  property of *S* is an aggregate of the  $\Phi$  properties of *Xs* when:

1. (W1)  $\Psi$  is invariant under the *rearrangement and intersubstitution* of *Xs*;
2. (W2)  $\Psi$  remains *qualitatively similar* (if quantitative, differing only in value) with the addition or subtraction of *Xs*;
3. (W3)  $\Psi$  remains invariant under the *disaggregation and reaggregation* of *Xs*; and
4. (W4) There are no *cooperative or inhibitory* interactions among the *Xs* that are relevant to  $\Psi$ ."
- 5.

A *failure* of a system to exhibit aggregativity, then, entails that the properties of the system as a whole are highly organization-dependent and that it is possible to explain the overall behavior of the system in terms of its *x*'s and the interactions of its  $\Phi$ 's in generating *S*'s  $\Psi$ -ing. It follows from Wimsatt's account of non-aggregativity that the emergence of higher-level properties in a system is a result of the system's compositional organization. First, an extended system is hierarchically organized such that its parts are organized at multiple different levels – some at the level of the brain, some at the level of the non-neural body, and some at the socio-cultural level. Second, an extended system exhibits a many-to-one compositional organization, where multiple different parts as well as the interactions among these parts determine *S*'s  $\Psi$ -ing. Third, an extended system (although it also applies to non-extended systems) is composed of heterogeneous but complementary parts working together in the assembly process of *S*'s  $\Psi$ -ing. Fourth, an extended system (and again not limited to extended systems) is, according to Craver (2007, pp. 134-139; compare Craver & Bechtel 2007) not only composed of *x*'s and  $\Phi$ 's but also *organized* from various *spatial, temporal* and *causal* conditions and constraints.

### 7.1. Case study #1: Tetris revisited

In this section, I shall apply the mechanistic framework to the Tetris case in order to supplement the dimension focusing on bodily manipulations and cognitive norms. I shall argue that we can explain the emergence of skillful behavior in Tetris (*S*'s  $\Psi$ -ing) as a result of the enabling system's organization-dependent structure violating the criteria of aggregativity (W1-W4). First, in violation of condition (W4), the *x*'s and their  $\Phi$ -ing cooperate in ways deeply relevant to *S*'s  $\Psi$ -ing. Without a particular part playing its role, the overall performance of *S*'s  $\Psi$ -ing would break down or be incomplete. This holds regardless of whether there is a malfunction in the parts located at the socio-cultural level or at various biological levels. Second, this implies that the distributed mechanism underlying skillful performance in Tetris is organized in a way violating condition (W2): subtracting or even adding component parts

Please do not cite

cannot be achieved without disturbing the mechanisms overall ability to  $\Psi$ . Hence, subtracting or adding parts to the mechanism is likely to result on qualitative dissimilarity with respect to  $\Psi$ . Third, in violation of condition (W1), the parts of the mechanism cannot be rearranged arbitrarily, since what binds  $S$ 's  $\Psi$ -ing is the *timing* of the parts and their interactions (Kirsh 1995, p. 62). The idea of timing of processing marks an aspect of mechanistically organized systems – *temporal organization*. As Craver says:

“The order, rate, and duration of successive component activities are crucial for the [mechanism], [and] it is not possible to change their order without interfering with how the mechanism works (or making it a different mechanism entirely).” (2007, p. 138)

In the Tetris study, the physical action of rotating a zoid generates neurobiological processes – Kirsh & Maglio term these ‘buffer icons’ – which, in turn, are reciprocally interacting with multiple different functional processes such as motor, attention, generation, and matching processes, while constantly modulated and constrained by higher-level sociocultural processes (Kirsh & Maglio 1994, p. 542). Furthermore, mechanisms are also *spatially organized*. As Craver says: The spatial organization of a mechanism includes, [...], the sizes, shapes, structures, locations, orientations, directions, connections, and compartments of mechanisms.” (2007, p. 137) The importance of space in cognition has long been recognized as an essential part of the acquisition of abstract problem-solving abilities (Kirsh 1995), of developing abstract concepts (Sheets-Johnstone 2011, p. 167), and so on. In Tetris, spatial organization takes on a similar role, in the sense that the effects of time-efficiency and transformation of informational structure is due to the location of the zoids, their shapes, and the causal connectivity established by manipulation. If this is the case then we can understand how extended cognitive systems include not only neural processes but are compositionally organized so as to include bodily manipulations of socio-cultural processes as well. This is what supplements Menary’s claim: “Epistemic actions and computations take place within the same state space. If this state space is the problem-solving state space, then we have difficulty in pulling apart mind, action and world [...]” (2007, p. 90)

## 7.2. Case study #2: Emergence of socially distributed remembering

In their study of socially distributed remembering (or collaborative recall), Sutton et al. (2010) argue that on some occasions, long-term couples have evolved interactive, dynamically integrated systems for collective memory of past events. Here is one of the dialogues, where a couple is acting in concert when remembering the beginning of their relationship (2010, p. 551):

*Husband*: No, I asked her out that night, but she said she couldn’t go.

*Wife*: No, that’s right.

*H*: So then I started to pester her the next week.

*W*: You did, you turned up after my [classes].

*H*: [Cooking classes].

Please do not cite

W: On Monday night.

H: That'd be it.

W: And took me for coffee.

H: Yes, the next Monday night.

W: And impressed me.

H: Yes.

Drawing on Wegner's theory of transactive memory (1987), the authors indicate that it is quite likely that the successful remembering in this long-term couple is an organization-dependent feature of a "socially coupled dynamical system with emergent properties, which in certain cases can be highly integrated and enduring and exhibit high levels of continuous reciprocal causation." (2010, p. 547) It is of course true, as the authors point out, that not all instances of people remembering together will involve interactive and dynamical processes, thereby exhibiting aggregative outcomes (2010, p. 552). But "in other cases each partner offers distinct but complementary contributions to a shared emergent product." (Sutton et al. 2010, p. 552) According to Sutton et al. (2010), remembering in this case is not the operation of the naked brain on its own, with the other person merely an external influence. As Wegner puts it: "A transactive memory structure thus can be said to reside in the memories of both individuals – when they are considered as a combined system." (Wegner 1985, p. 257; quoted in Sutton et al. 2010, p. 547) Cases of socially distributed cognition are equally taken seriously by Theiner (2009, 2011), Theiner & O'Connor (2010), and Theiner et al., (2010) who argue in favor of a revised version of the group mind thesis. Influenced by anthropologically sensitive cognitive science (Hollan et al. 2000; Hutchins 1995), the philosophical framework of mechanistic explanation and the transactive memory approach, Theiner and colleagues argue that recent studies of group cognition – especially problem solving cases – reveal that cognitive capacities that we normally would not hesitate to ascribe individual cognizers are also ascribable to groups. What these case studies show, and as the one by Sutton et al., shows, is how continuous reciprocal causation between people in a group results in emergent cognitive properties at the group level. The point of contact between my argument in this paper and the work done by Theiner and colleagues is that we think that extended cognitive systems – including at the group level – have organization-dependent cognitive capacities. As Theiner et al., says: "Group cognition is thus emergent phenomenon in the sense of Wimsatt (1986)." (2010, p. 378) In their study of socio-culturally distributed cognition, Theiner and colleagues also report on the phenomenon of a transactive memory system (TMS). They distinguish between two components of a TMS. First a *representational* component, which "consists of the organized stock of memories that are retained by individual members, including higher-order memories about who knows what." (2010, p. 388) The second component is a *procedural* component, whose function is the continuous creation and constant maintenance of the representations, including "all direct and indirect communication processes [...] by which individuals cooperatively allocate, encode, retrieve, share, and elaborate memories." (2010, p. 388) Among the many areas of socially distributed remembering examined by Theiner et al., they look at how TMS's have been used in small group research. In these studies, TMS's are used as latent group-level variables. According to Theiner et al., TMS's in small groups working on assembly tasks

Please do not cite

become manifested in several different ways. First as *memory differentiation*: “i.e. the tendency of group members to specialize in recalling distinct aspects of the assembly process.” (2010, p. 389) The second is what they call *task credibility*: “i.e. how much members trusted one another’s expertise.” (2010, p. 389) And finally in terms of *task coordination*: “i.e. the effective use of transactive retrieval strategies [...]” (2010, p. 389) From this, Theiner et al. report several findings. Groups perform much better when their members undergo training on the task together, and worse if each member of the group went through the training individually. An important aspect of coordination and collaboration is that the process of undergoing training together establishes “more elaborate, accurate, and mutually shared beliefs about the distribution of know-how in groups.” (Theiner et al. 2010, p. 389)

## 8. Dissolving the coupling-constitution fallacy

I have already hinted at why Menary’s version of second-wave EC avoids the C-C fallacy. One horn of the C-C fallacy entails that the EC theorist commits an instance of the fallacy when an environmental resource deemed cognitive by being causally coupled to an individual cognitive agent. But this is not what Menary is arguing. As Menary schematizes the idea: “X is the manipulation of the notebook *reciprocally* coupled to Y – the brain processes – which together constitute Z, the process of remembering.” (2006, p. 333). So, the MT as a constitutive thesis does not infer cognitive status to an environmental resource simply because that resource is causally coupled to a pre-existing cognitive agent. Menary is clear about this when he says that the goal of cognitive integration is “to explain why X and Y are so coordinated that they together function as Z, which causes further behavior.” (2006, p. 334) Hence, an extended cognitive property, Z, is an emergent property of a mechanistically composed system. It is not the case – as exemplified in the Tetris study – that cognition is first “inside” the head and then “extends outwards”! This claim can be further underpinned by Theiner et al’s use of the mechanistic framework to support the idea of group cognition or socially distributed cognition. As Theiner et al., argues:

“To begin with, when we claim that an individual cognitive system X is on some principled way coupled to another individual cognitive system Y, we do not mean to imply that X is thus part of Y. *Instead, what we assert is rather that the individuals who instantiate X and Y can engage in structured interactions so as to constitute an organized group-cognitive system Z that encompasses those individuals among its proper parts.*” (2010, p. 390; italics added)

Whereas Menary’s position deals primarily with the “process” version of the C-C fallacy, Theiner et al. set up an argument against the “systemic” version. The systemic version of the C-C fallacy holds that even if a cognitive system X is dynamically coupled to a cognitive system Y, it does not entail that X is part of Y. But this is not what Theiner et al. are arguing. What these authors are stating is that the persons who “instantiate X and Y can engage in structured interactions so as to constitute an organized group-cognitive system Z that encompasses those individuals among its proper parts.” (2010, p. 390) We can follow Theiner et al. (2010), so as to argue that the case studies discussed here compose an organized cognitive system or process Z that encompasses one or more individuals. The Tetris study

Please do not cite

and the studies on socially distributed remembering “form an *integrated system with functional gains*.” (2010, p. 390) Theiner et al., following Wilson (in press), distinguish between three elements of what it means to be an integrated system with functional gains. First, “two (or more) elements are *coupled* just in case they exchange information by means of reliable, two-way causal connections between them [...]” (Theiner et al. 2010, p. 390) so that the elements “are interdependent in their cognitive and behavioral activities.” (Theiner et al. 2010, p. 390) In Sutton et al’s study of socially distributed remembering, successful remembering is an organization-dependent feature of such continuous reciprocal causation. Second, “two (or more) coupled elements form an *integrated system* in situations in which they operate as a single causal whole [...] – with causes affecting the resultant system as a whole, and the activities of that system as a whole producing certain effects.” (Theiner et al. 2010, p. 390) In both Menary’s redeployment of the Tetris case and in my explanation of the Tetris study by applying the mechanistic framework, successful performance in Tetris is an emergent property of an organization-dependent system failing to exhibit aggregativity. Because of this, there is an integrated system underlying *S’s  $\Psi$ -ing*. Third, “an integratively coupled system shows *functional gain* just when it either enhances the existing functions of its coupled elements, or manifests novel functions relative to those possessed by any of its elements.” (Theiner et al. 2010, p. 390-91) Both the Tetris case and the Sutton et al. case of socially distributed remembering satisfy this condition. Also, we see this condition fulfilled in Theiner et al’s research on small group cognition, since the individual members perform better when working in joint collaboration rather than working alone.

## Conclusion

In this paper, I have argued that the proponent of EC is better off explaining the ontological claim of EC by exploring the coordination dynamics between “internal” and “external” resources rather than opting for a metaphysical approach to the constitutive claim. In particular, of all the alternative ways of construing the constitutive claim in the EC literature, I have shown that only Menary’s manipulation thesis provides an adequate fertile ground for making EC’s ontological claim. In supplementing Menary’s position, I have argued that the philosophical accounts of mechanistic explanations can shed additional light on EC’s ontological claim. I have shown this by reference to a few empirical case studies across the dimensions of “individual-artifact EC” and “collective, socio-culturally distributed EC.” I hope to have shown that by approaching the constitutive claim from an empirically oriented perspective, it is possible to resolve the C-C fallacy. I end this paper by briefly positioning my argument that EC’s constitution claim can be re-evaluated by adopting the relation of composition in mechanistic explanation within the EC literature. In order to do so, I shall apply Kirchhoff’s (2011) taxonomy of first-, second-, and third-wave EC. As we have seen, first-wave EC attempts to ground constitution by way of functionalist parity arguments and second-wave EC in terms of complementarity and cognitive integration. Hence, my attempt to justify EC’s ontological foundation based on the philosophical framework of mechanistic explanation arguably fits better with Kirchhoff’s third-wave version of EC.

Please do not cite

## References

- Adams, F., & Aizawa, K. (2010a). The value of cognitivism in thinking about extended cognition. *Phenomenology and the Cognitive Sciences*, 9, 579-603.
- Adams, F., & Aizawa, K. (2010). Defending the bounds of cognition. In R. Menary (ed.), *The Extended Mind* (pp. 67-80). Cambridge, MA: The MIT Press.
- 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: Cambridge University Press.
- Adams, F., & Aizawa, K. (2008). *The Bounds of Cognition*. New York: Blackwell.
- Adams, F., & Aizawa, K. (2001). The bonds of cognition. *Philosophical Psychology*, 14(1), 43-64.
- Baker, L. R. (2000). *Persons and Bodies: a Constitution View*. Cambridge: Cambridge University Press.
- Beer, R. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4, 91-99.
- Bechtel, W. (2009). Explanation: Mechanism, modularity, and situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 155-170). Cambridge: Cambridge University Press.
- Bennett, K. (2011). Construction area (no hard hat required). *Philosophical Studies*, 154, 79-104.
- Block, N. (1980). Introduction: What is functionalism? In N. Block (ed.), *Reading in Philosophy of Psychology* (Vol. 1, pp. 171-184). Cambridge, MA: Harvard University Press.
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University.
- Clark, A. (2005). Intrinsic content, active memory, and the extended mind. *Analysis*, 65, 1-11.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis*, 58, 7-19.
- Clark, A. (1998). Author's response: review symposium on *being there*. *Metascience*, 7, 95-103.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: The MIT Press.
- Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: The MIT Press.
- Craver, C.F., and Bechtel, W. (2007). Top-down causation without top-down causes. *Biology and Philosophy*, 22, 547-563.
- Craver, C.F. (2007). *Explaining the Brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
- Craver, C.F. (2001). Role function, mechanisms and hierarchy. *Philosophy of Science*, 68, 31-55.
- Davidson, D. (1987). Knowing one's mind. *Proceedings and Addresses of the American Philosophical Association*, 60, 441-458.
- Donald, M. (1998). Material culture and cognition: concluding thoughts. In C. Renfrew., & C. Scarre (eds.), *Cognition and Material Culture: the Archaeology of Symbolic Storage* (pp. 181-187). Cambridge: The McDonald Institute of Archaeology Monographs.
- Donald, M. (1991). *Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Hawley, K. (2006). Principles of composition and criteria of identity. *Australasian Journal of*

Please do not cite

*Philosophy*, 84(4), 481-493.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *Transactions on Computer-Human Interaction*, 7(2), 174–196. Hurley, S. (2010). Varieties of externalism. In R. Menary (ed.), *The Extended Mind* (pp. 101-153). Cambridge, MA: The MIT Press.

Hutchins, E. (2008). The role of cultural practices in the emergence of modern human intelligence. *Philosophical Transaction of the Royal Society*, 363, 2011–2019.

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

Kirchhoff, M. D. (2011). Extended cognition and fixed properties: Steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences*, DOI: 10.1007/s11097-011-9237-8.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, 73, 31-68.

Kirsh, D., and Maglio, P. (1994). On distinguishing epistemic from pragmatic actions. *Cognitive Science*, 18, 513-549.

Lowe, E. J. (1989). *Kinds of Being*. Oxford: Basil Blackwell.

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

McLaughlin, B., & Bennett, K. (2005). Supervenience. *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/supervenience/>

Menary, R. (2010c). The holy grail of cognitivism: A response to Adams and Aizawa. *Phenomenology and the Cognitive Sciences*, 9, 605-618.

Menary, R. (2010b). Dimensions of Mind. *Phenomenology and the Cognitive Sciences*, 9(4), 561-578.

Menary, R. (2010a). Cognitive integration and the extended mind. In R. Menary (ed.), *The Extended Mind* (pp. 227-243). Cambridge, MA: The MIT Press.

Menary, R. (2010). Introduction. In R. Menary (ed.), *The Extended Mind* (pp. 1-25). Cambridge, MA: The MIT Press.

Menary, R. (2007). *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave Macmillan.

Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19(3), 329-344.

Robbins, P., & Aydede, M. (Eds.). (2009). *The Cambridge Handbook of Situated Cognition*. New York: Cambridge University Press.

Rowlands, M. (2010). *The New Science of Mind*. Cambridge, MA: The MIT Press.

Rowlands, M. (unpublished, ms). Perception: From extended mind to embodied phenomenology. Unpublished paper given at a symposium on Embodied Perception held at the spring 2008 meeting of the Pacific Division of the American Philosophical Association, March 18-23, 2008, Pasadena, California.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: The MIT Press.

Rupert, R. (2010). Extended cognition and the priority of cognitive systems. *Cognitive Systems Research*, 11, 343-356.

Rupert, R. (2009). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.



Please do not cite

- Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 9, 625-636.
- Shapiro, L. (2010). *Embodied Cognition*. London and New York: Routledge.
- Sheets-Johnstone, M. (2011). Thinking in movement: Further analysis and validations. In J. Stewart, O. Gapenne, and E.A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science* (pp. 165-182). Cambridge, MA: The MIT Press.
- Sprevak, M. (2009). Extended cognition and functionalism. *The Journal of Philosophy*, 9, 503-527.
- Sterelny, K. (2010). Minds: Extended or Scaffolded? *Phenomenology and the Cognitive Sciences*, 9(4), 465-481.
- Sutton, J. (2010). Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. In R. Menary (ed.), *The Extended Mind* (pp. 189-225). Cambridge, MA: The MIT Press.
- Sutton, J., Harris, C. B., Keil, P. G., & Barnier, A. J. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, 9(4), 521-560.
- Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics & Cognition*, 14(2), 235-247.
- Theiner, G. (2011). *Res Cogitans Extense: A Philosophical Defense of the Extended Mind Thesis*. Frankfurt: Peter Lang.
- Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, 11, 378-395.
- Theiner, G. & O'Connor, T. (2010). The emergence of Group Cognition. In A. Corradini & T. O'Connor (Eds.), *Emergence of Science and Philosophy* (pp. 78-117). New York: Routledge.
- Theiner, G. (2009). Making sense of group cognition: the curious case of transactive memory systems. *ASCS09: Proceedings of the 9<sup>th</sup> Conference of the Australasian Society for Cognitive Science*. DOI: 10.5096/ASCS200951.
- Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, 92, 345-381).
- Wasserman, R. (2004). The constitution question. *Nous*, 38(4), 693-710.
- Wegner, D. (1987). Transactive memory: A contemporary analysis of group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of Group Behavior* (pp. 185-208). New York: Springer-Verlag.
- Wheeler, M. (2010). In defense of extended functionalism. In R. Menary (ed.), *The Extended Mind* (pp. 245-270). Cambridge, MA: The MIT Press.
- Wilson, R. (2010). Meaning making and the mind of the externalist. In R. Menary (ed.), *The Extended Mind* (pp. 167-188). Cambridge, MA: The MIT Press.
- Wilson, R. A., & Clark, A. (2009). How to situate cognition: letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 55-77). Cambridge: Cambridge University Press
- Wilson, R. A. (2007). A puzzle about material constitution & how to solve it: Enriching constitution views in metaphysics. *Philosophers' Imprint*, 7(5), 1-20.
- Wilson, R. A. (2005). *Genes and the Agents of Life*. New York: Cambridge University Press.

Please do not cite

Wilson, R. A. (2004). *Boundaries of the Mind*. New York: Cambridge University Press.

Wilson, R. A. (2001). Two views of realization. *Philosophical Studies*, 104, 1-31.

Wimsatt, W. (1986). Forms of aggregativity. In A. Donagan., A.N. Perovich., and M.V. Wedin (eds.), *Human Nature and Natural Knowledge* (pp. 259-291). Boston: D. Reidel.

Wimsatt, W. (1976). Reductionism, levels of organization and the mind-body problem. In G. Globus et al. (eds.), *Consciousness and the brain* (pp. 199-267). New York: Plenum.