



# In search of an ontology for 4E theories: from new mechanism to causal powers realism

Charles Lassiter<sup>1</sup> · Joseph Vukov<sup>2</sup>

Received: 3 February 2020 / Accepted: 19 May 2021  
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

Embodied, embedded, enactive, and extended (4E) theorists do not typically focus on the ontological frameworks in which they develop their theories. One exception is 4E theories that embrace New Mechanism. In this paper, we endorse the New Mechanist's general turn to ontology, but argue that their ontology is not the best on the market for 4E theories. Instead, we advocate for a different ontology: causal powers realism. Causal powers realism posits that psychological manifestations are the product of mental powers, and that mental powers are empirically-discoverable features of individuals that account for the causal work those individuals do. We contend that causal powers realism provides a unifying framework for the central commitments of 4E theories, as well as additional resources for theorizing in a 4E framework. And while New Mechanism offers some of these resources as well, we argue that causal powers realism is ultimately the better of the two.

**Keywords** 4E theories of cognition · Causal powers · Metaphysics of mind · Philosophy of psychology · New mechanism

## 1 Introduction

Saying our minds include bits of the extracranial world doesn't start fights like it used to.<sup>1</sup> Since Clark and Chalmers (1998) made 4E approaches to mind mainstream, otherwise rival bands of philosophers and cognitive scientists have declared themselves part of the embodied, embedded, enactive, or extended (4E)

<sup>1</sup> Author order was determined by who can eat an apple pie faster.

---

✉ Joseph Vukov  
jvukov@luc.edu

Charles Lassiter  
lassiter@gonzaga.edu

<sup>1</sup> Department of Philosophy, Gonzaga University, Spokane, WA, USA

<sup>2</sup> Department of Philosophy, Loyola University Chicago, Chicago, IL, USA

mind movement. Current membership ranges from Clark-style functionalists to the autopoietic enactivism inspired by Maturana and Varela (1979). 4E theorists have staked out their positions, agreeing that cognition is not merely a function of disembodied brains even while disagreeing about the details. Since the explosion of 4E theories in the 1990s and early 2000s, however, there has been a lull in the development of genuinely novel positions. Progress for a 4E theory is now measured in the smallest of gains: an intuitive pull here; an elegant experimental interpretation there; a theoretically parsimonious explanation still here. The initial gung-ho charge of 4E cognition has devolved into trench warfare.

Why think this? Two pieces of evidence: co-citation and keyword networks. Consider Figs. 1 and 2 (more details about the methodology are given in the “Appendix”).

The co-citation network maps 30 papers that were co-cited at least ten times in the dataset of 556 papers. The larger the node, the more often it shows up in co-cites and the more central it is in the network. The first thing to see is that there are two distinct clusters: an enactivist cluster and an extended mind (very broadly construed) cluster. There’s some connective tissue between the clusters, but by and large each keeps to their own. What does this tell us? Just that more people, when publishing, tend to co-cite enactivist authors together or extended mind authors together; but philosophers tend not to cite both in the same documents as much. This is initial evidence of 4E’s siloing.

Further evidence is found in the keyword network. While citation practices can be more or less promiscuous, keywords tend to focus on capturing the main ideas of the paper in a few searchable words. The dataset had 2500 keywords, roughly 1200 of them unique. Figure 2 shows the relative importance and connection of the 47 most-used keywords. Notice that there are three major clusters (indicated by the separate colors): an extended mind cluster, an embodied cognition cluster, and an enactivist cluster, each characterized by their own constellation of keywords.

We could draw various conclusions from these figures, but there’s one in particular to highlight. While 4E theorists are all working on ideas in the same neighborhood, it’s clear that there are divisions among the theories, each with their own set of keywords and topics. There are many reasons for this. Enactivists and Clark-style extended mind theorists, for example, differ on methodology and criteria for adequate scientific explanation, the former appealing to phenomenological characterization of cognitive phenomena and the latter to functional description. In fact, it’s likely these differences that drive much of the balkanization of 4E theorizing.<sup>2</sup> But another roadblock is the lack of a shared ontological framework. Among proponents of embodied cognition alone, some, like Goldman, are dyed-in-the-wool cognitive internalists. Others, like Gibson, are most definitely not.

One way forward is the ‘New Mechanism’ movement sweeping the philosophy of science.<sup>3</sup> New Mechanism says that causal explanations cite the underlying

<sup>2</sup> Thanks to an anonymous reviewer for bringing this point to our attention.

<sup>3</sup> Machamer et al. (2000), Bechtel and Richardson (2010), Glennan (2002), Bechtel and Abrahamsen (2005), Craver (2007).

mechanisms of an explanandum.<sup>4</sup> If you want to know how circadian rhythms (Bechtel & Abrahamson, 2010) or spatial memory (Craver, 2007) work, it is enough to identify and describe the mechanisms underlying these processes. And what *is* a mechanism? It's an entity, often made of other parts and organized, that performs a function in virtue of its constituent parts. It also plays well with other mechanisms, usually (Machamer et al., 2000; Bechtel & Abrahamsen, 2005). The New Mechanism is a broad position, one that can be developed in many directions. Bechtel (2009) sets the stage for 4E theorizing in a New Mechanist framework. Explanations should not only look down to the target mechanism's constituents but also around and up to see how the mechanism is connected with other mechanisms and its larger environment. 4E theorists have enjoyed success by following Bechtel's advice. Krickel (2020), for example, proposes an account that can distinguish cases of cognitive extension from cognitive situatedness—a refreshed version of issues raised by Adams and Aizawa's (2001, 2008) coupling-constitution fallacy—using resources native to New Mechanism. We're happy to admit that the New Mechanism provides one viable path for navigating out of the theoretical doldrums we have identified above.

In this paper, however, we propose an alternative ontology for 4E theories, one we believe is ultimately a better one. Specifically, we look to the rise of a literature devoted to ontologies of causal powers, ontologies that often explicitly reject modern ontological commitments: proponents of powers have rejected the metaphysical contingency of natural laws (e.g. Bird, 2007), the distinction between categorical and dispositional properties (e.g., Heil, 2003), a commitment to abundant conceptions of properties (e.g. Jaworski, 2016), and Humean accounts of causation (e.g. Ellis, 2001). In what follows, we bring causal powers realism to bear on 4E theories, and in doing so, advance two central contentions. First, that causal powers realism can breathe fresh air into 4E debates by offering a unifying framework, one that accommodates core 4E commitments and provides an array of valuable theoretical resources. Second, we offer a limited version of an inference to the best explanation, contending that causal powers realism does this work better than its most obvious contemporary rival—the New Mechanism. 4E theorists who are looking for a framing ontology—or simply hoping to avoid the slump facing the 4E literature—should therefore give serious consideration to causal powers realism.

We proceed as follows: we begin by providing a gloss of the core commitments of 4E theories, and the ways in which the New Mechanism has been used to move 4E conversations forward. Then, we introduce causal powers realism—first, we outline the theory; second, we show how causal powers realism can also unify the core commitments of 4E theories. We then offer our inference to the

---

<sup>4</sup> One point of contention is whether mechanistic explanations have any ontological purchase in New Mechanism. It's possible for the New Mechanist to advocate for mechanistic descriptions but be agnostic on whether those descriptions map onto the world they're supposed to describe. In this paper, we're attributing to New Mechanism realism about mechanistic descriptions: when neuroscientists describe mechanisms for vision, those mechanisms are part of the furniture of the world (at least until otherwise notified).

4E Co-Citation Network

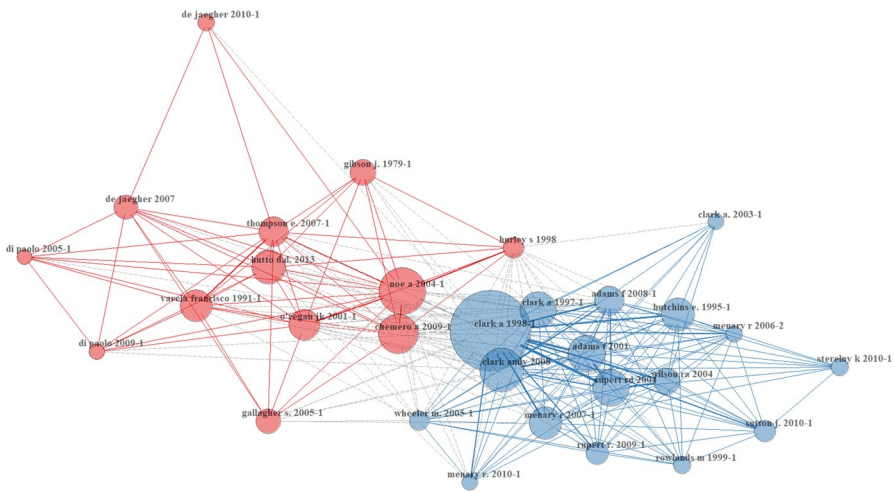


Fig. 1 Co-citation network

4E Keyword Co-occurrence Network

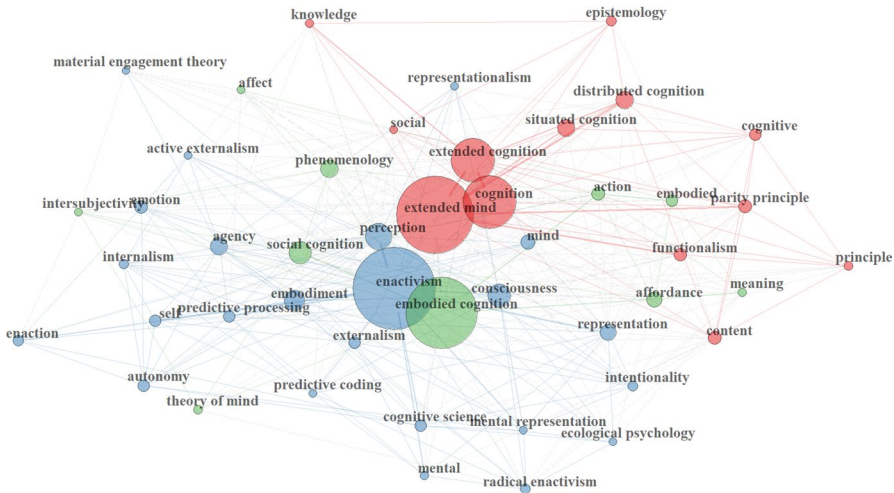


Fig. 2 Keyword network

best explanation, arguing that the explanatory fit between causal powers realism and 4E theories is better than that between the New Mechanism and 4E theories. Our conclusion? Theories of 4E cognition are better served by clarifying their metaphysical commitments with an old ontology instead of the New Mechanism.

## 2 The core commitments of 4E theories<sup>5</sup>

The late twentieth century saw a proliferation of 4E theories of cognition. Minds, on the 4E view, are embodied,<sup>6</sup> extended,<sup>7</sup> embedded,<sup>8</sup> and enactive.<sup>9</sup> In what follows, we'll need to capture a wide array of 4E theories under a handful of principles. The first step is to observe that 4E cognitive science is a response to “mainstream” cognitive science. It doesn't help that “mainstream” cognitive science is a diverse field including mutually exclusive theories. Still, one can get the flavor of mainstream cognitive science by considering titans of the field in the mid- to late-twentieth century: Jerry Fodor, Marvin Minsky, Donald Broadbent, George Miller, Noam Chomsky, David Marr, Zenon Pylyshyn, Oliver Selfridge. What ties these figures together is a willingness to take the computer metaphor of mind literally along with the implicit assumption that the realizing substance has unambiguous spatial boundaries. 4E theories are likewise united in their *resistance* to this understanding of the mind.<sup>10</sup>

That's not to say 4E theories are themselves unified, either in terms of scientific explanation or ontology. Far from it. For example, while all enactive theories are also embodied, not all embodied theories are enactive. Other 4E theories actively exclude each other: some theories of extended cognition reject the idea that minds are embedded.<sup>11</sup> And Clark's functionalist vision of the extended mind is hardly sympathetic to the radical enactivism of Hutto and Myin (2013, 2017). Even so, we can provide a handful of conditions that most 4E theories satisfy. We mentioned before that 4E theories are united in their rejection of some components of mainstream cognitive science, so we'll use Fodor and Dretske as reality-checks to ensure our broad-strokes painting isn't too sketchy. If our 4E conditions cause an allergic

<sup>5</sup> It might be that ‘research tradition’ is a better moniker for collections of claims about enactivism, embodied cognition, etc. (Miłkowski ms). For us, nothing hinges on whether such collections of claims are theories or research traditions. We're happy to call them either. For the sake of consistency—and to avoid repetitions of clunky phrases—we will stick with ‘theory’, with the understanding that some philosophers of science demur to such usage.

<sup>6</sup> Thompson et al. (1991), Goldman (2012), Clark (1997), Gallagher (2005).

<sup>7</sup> Clark and Chalmers (1998), Clark (2011), Menary (2007), Sutton (2006), Kirchoff (2012).

<sup>8</sup> This also goes by the name ‘situated’ cognition. Millikan (2009), Brooks (1991), Vera and Simon (1993).

<sup>9</sup> Gallagher (2018), Thompson (2007), Thompson et al. (1991), di Paolo et al. (2018).

<sup>10</sup> The pushback against mainstream cognitive science was largely about scientific explanations rather than metaphysics, though there are, plausibly, some objections to this. For example, Thompson et al. (1991) and Gallagher (2005) are committed to Buddhist and/or phenomenological ontologies. While they may have been dissatisfied with the scientific explanations proffered by mainstream cognitive science, it's conceivable that a principled philosophical objection initially drove the opposition. When there's a history of 4E cognitive science to rival Maggie Boden's *Mind as Machine*, maybe we'll get some insight. Thanks to an anonymous reviewer for suggesting we clarify these points.

<sup>11</sup> The reason for this is straightforward: theories of extended cognition treat the environment as constitutive of mind while theories of embedded cognition treat the environment as causally related to mind. This is what the debate over the coupling-constitution fallacy was all about, cf. Adams and Aizawa (2001, 2008, 2010) and responses by Clark (2008, 2010), Menary (2010), Rowlands (2010), and Kagan and Lassiter (2013).

reaction in hardcore Fodorians or Dretsians, then our conditions probably aren't *too* broad. We intend them to capture family resemblances that exist among 4E theories, rather than necessary and sufficient conditions. That way, we can use them to suggest that causal powers realism can provide a unifying ontological framework.

Consider, then, the following<sup>12</sup>:

1. **Embodied Cognition:** Bodily states are non-trivially incorporated into cognitive processing.
2. **Environmental Cognition:** Environmental conditions outside the body figure non-trivially into explanations of cognitive processing beyond correlation with internal entities.
3. **Enactive Cognition:** Cognitive agents are autonomous organism-environment systems, whose self-organizing and self-controlling dynamics include undecomposable feedback loops between perception, thought, and action; that is, cognition is enacted.

Each of these characterizations benefits from a little nuancing.

Embodied Cognition captures the idea that the body is not merely a site of cognition; it's part of the cognitive processing. The strategically vague term is 'non-trivially.' We can get a sense of what's involved by considering *trivial* involvement of the body in cognition. Here are a few examples: you can't do mental math without lungs (or some other way to move oxygen to the bloodstream), it takes eyes to read print, and there's no doing the Stroop task without kidneys or a functional equivalent. These are intuitively trivial because the activity of the mentioned organs are, at best, only indirectly involved in the cognitive task. Even better: systematic interventions in the functionings of these organs don't result in changes to particular cognitive functions. Tinkering with the activity of my heart or pumping carbon dioxide into a mask for me to breathe inhibits my ability to be alive, not just do mental math. An example of non-trivial involvement, by contrast, may include a 4-year-old's use of fingers to complete simple addition problems: the fingers are directly involved in the cognitive task, and systematically intervening with their functioning (say, by asking the child to put their hands in their pockets) derails the specific cognitive function. Now, the extent to which the body is non-trivially involved in cognitive functioning depends on the details of each 4E theory.<sup>13</sup> On the shallow end, Goldman's bodily

<sup>12</sup> Readers will observe that there are only *three* E's and not *four*. The reason for this is because two of the E's—'extended' and 'embedded'—are subsumed under our one 'E' of environment. We make this choice for two reasons. First, as we mention in note 14, debates between embedded and extended theorists don't turn on whether the environment relevant for an understanding of cognition but whether the world beyond the skull is part of the mind (cf. Clark, 2008). So bringing both 'extended' and 'embedded' under the principle **Environmental Cognition** doesn't miss any of the positions currently in the literature. Second, it's theoretically and argumentatively parsimonious to capture the wide array of 4E positions under three principles rather than four. Doing so lets us streamline our argument. Thanks to an anonymous review for pressing us to clarify this.

<sup>13</sup> Wilson (2002) and Shapiro (2011) are two taxonomies for theories of embodied and extended cognition. See also Newen et al. (2018) for another recent overview.

formats holds that the body gets into cognition by way of being internally represented (Goldman, 2012). On the deep end, ecological psychologists inspired by J.J. Gibson argue that the body achieves cognition directly, rather than via representations (1979). Goldman's and Gibson's accounts of embodied cognition, on their faces, are incompatible: information in cognitive processing for Goldman is internally realized, which is not the case for Gibson; however, they both fall under the umbrella of Embodied Cognition. The umbrella is large but not so large that it covers anything. Fodor's account of cognition and mental content (1975, 1987, 2008), for example, rejects Embodied Cognition: on Fodor's account, the body is merely the place where cognitive processing happens. There is no sense in which sentences in the Language of Thought are shaped by bodily states and processes. This is a consequence of mainstream cognitive science's implicit commitment to functionalism.<sup>14</sup>

Next consider Environmental Cognition. This claims that the world beyond the body is more important for thinking than just correlation of internal contents and external states. In other words, the environment non-trivially figures into cognition. As we've worded it, Environmental Cognition is thus set up as an explicit foil to Fodorian and Dretsikian accounts of mental contents as well as methodological solipsism (Fodor, 1980). Those depict the surrounding world as a source of inputs, but the real cognitive action is internal to the subject (cf. Dretske, 1981, Fodor, 1987). 4E accounts of cognition, however, often reject the Fodorian and Dretsikian picture (and methodological solipsism) in favor of Environmental Cognition. Consider, for example, accounts of *embedded* cognition (e.g. Simon, 1996, Rupert, 2009), as well as its more radical cousin *extended* cognition. Embedded cognitive scientists hold that the mind is a relatively stable and enduring system well-ensconced inside the skull. However, a complete theory about mind requires looking outside the skull to environmental forces that causally impact the internal cognitive processing. Extended cognition pushes the boundaries, maintaining that the cognitive system loops beyond the bounds of the skull. But no matter whether the mind is causally related to or is constituted by the environment, both theories take Environmental Cognition on board.

Finally, consider Enactive Cognition. A central commitment of Enactive Cognition is the rejection of the idea that perception, thought, and action are separate, context-independent processes in need of interfacing. Rather, proponents of Enactive Cognition hold that cognitive systems are "self-organizing" and "self-controlling" (cf. Thompson, 2007) with cognition consisting in undecomposable feedforward and feedback loops between organisms and their environments. Organism-environment units are self-organizing and have self-controlling dynamics by which they define and maintain their own boundaries (which may not line up with where the skin is) and the behavior of the system is determined by its internal elements. For cognitive agents, organism-environment loops of perception, cognition, and action are tangled up with one another and are accompanied by a phenomenology about perception that guides us in our interactions with the world.

<sup>14</sup> See Piccinini (2004) for an excellent discussion of functionalism and computationalism (and their conflation) in cognitive science. Thanks to an anonymous reviewer for bringing this to our attention.



An illustration may help. O'Regan and Noë (2001) cite visual inversion experiments in which participants wear goggles that vertically invert visual stimuli: what is actually in the lower half of the visual field appears in the top half (and vice versa). Perhaps surprisingly, wearers eventually adjust to the goggles. All told, it takes about two weeks to adapt and move about normally. It's the *process* of the adjustment, however, that is of greatest interest for enactivists. During the adaptation, participants will report inconsistencies that wouldn't be expected given a mainstream cognitive science view of perception and action. For example, reporting Kohler's (1951) observations, a subject perceived a car driving on the correct side of the road (something that was not previously true in the early days of adapting to the goggles) but the numbers on the car's license plate appeared as if in a mirror. Piecemeal adaptation is odd unless perception, action, and thought are entangled and context-dependent processes. It's vital that one avoids moving cars but not necessarily accurately perceiving license plates; so, perception of where cars are coming from adapts more quickly than veridical perception of license plates. So it is not particularly surprising for the proponent of Enactive Cognition—who claims that action and perception are mutually constituting—that the timeline for restored perception should be distinctive as well. For the proponent of Enactive Cognition, locations and trajectories of cars correct sooner because subjects act in a way to constitute those veridical perceptions sooner. Generally, proponents of Enactive Cognition point to cases in which action, perception, and thought are tied closely together to motivate their claim that these are not in fact separate processes at all.<sup>15</sup>

It's not so much that mainstream cognitive science gets off the bus at a certain point with Enactive Cognition inasmuch as it takes an entirely different mode of transportation. But one point of difference between proponents of Enactive Cognition and mainstream cognitive science is that the latter keep perception, cognition, and action separate. Modules do the theoretical lifting here: perception modules input and output a small range of relevant information. Visual perception, then, is an object of study independent of its relations with action and cognition.<sup>16</sup>

There's more to say about Embodied Cognition, Environmental Cognition, and Enactive Cognition, but this will do for now. Recall that we do not intend these theses to capture necessary or sufficient conditions for a theory to count as 4E, but rather to spell out certain family-resemblances among contemporary 4E theories. If a theory is committed to at least one of these theses, there is good reason to think 4E theorists will recognize it as one of their own.

<sup>15</sup> Some enactivists further claim that this tangled skein of processes is part of a larger process of sense-making, a process of adaptive coping to an uncertain world.

<sup>16</sup> Given this quick sketch of the difference between mainstream cognitive science and Enactive Cognition, one might wonder if a weaker view of the information encapsulation of perception that accommodates cognitive penetration might be a way to bridge the gap. Pursuing this possibility would take us too far afield, but it suffices to note that acknowledging that perception is penetrable by cognition or behavioral processes doesn't entail that one is on board with Enactive Cognition.



### 3 New mechanism and 4E theories

We saw the main idea behind the New Mechanism in the introduction: causal explanations cite mechanisms underlying phenomena of interest.<sup>17</sup> We also suggested that the New Mechanism is capable of providing a unifying ontology for 4E theories. In what follows, we explore both claims. While the literature in this genre is rich with insights, we focus here on three claims about the concept of mechanisms.

First, *mechanistic explanation invokes hierarchical levels at which mechanisms are located*. This is the price of admission for New Mechanism. Bechtel's (2009) suggestion that philosophers and scientists look down, around, and up embodies this insight. This strategy of explanation takes to heart that no mechanism is an island. In that spirit, New Mechanists often advocate for integrative pluralism: explanation of some target phenomenon will require integrating mechanisms across levels and fields (e.g. Mitchell, 2003, Brigandt, 2010). Such integrative accounts will often involve both mechanistic explanations and mathematical modelling to provide insight into some complex phenomenon (cf. Deulofeu et al., 2019).

Second, *higher-level mechanisms do not reduce to lower-level mechanisms*. Craver (2007), for instance, offers a way of understanding how higher-level properties are relevant for explaining a particular phenomenon. In brief: change the differently-leveled properties, see what happens. Glennan (2010) offers as an example a key opening a lock. The key has its microstructure (arrangement of its constituent atoms) and its macrostructure (the shape of the key). There are limits on changes to the microstructure that will allow the lock to open. Keys made of ice typically won't open locks but brass keys will. If you keep the microstructure the same but change the shape of the key, then your key is useless. This suggests that the macrostructure is causally relevant for opening the lock and not just the microstructure. So the causal relevance of the macrostructure is independent of the microstructure. This lesson goes for all manner of phenomena: higher-level mechanisms are causally efficacious and not obviously reducible to lower-level ones.

Third, *phenomena have definite spatio-temporal boundaries and so do their underlying mechanisms*. Andersen (2014a) makes this point about mechanisms uncovered by the sciences explicit. Consider, by analogy, the process of fuel injection in a car. It has definite spatio-temporal boundaries and so do the mechanisms that underlie it—I can point to (or remove and clean) the carburetor. Importantly, the idea that mechanisms have locations differs from the idea that the mechanism is *localized*. It might turn out that mechanisms underlying target phenomena are distributed. Language comprehension, for example, involves different, non-contiguous regions of the brain. Even so, the parts involved, though distributed, have definite spatio-temporal boundaries. Issues of vagueness in metaphysics aside, there is a place where Wernicke's area begins and ends. It has an address.

We're happy to admit that New Mechanism as an explanatory and ontological framework can help advance debates in 4E cognition. *Prima facie*, they get on like

<sup>17</sup> Cf. Andersen (2014a, b) for discussion of the many ways in which 'mechanism' is used in contemporary philosophy of science. We're using her mechanism<sub>1</sub>: mechanisms as integral to scientific practices.

gangbusters. 4E accounts often describe causal relationships among different levels, like the “looping” causal relationships of enactivism or even Clark’s predictive processing accounts of perception, action, and cognition (Clark, 2013). 4E accounts likewise resist reduction of higher-level mechanisms to lower-level ones. Krickel (2020) goes further. She uses Kaplan’s (2012) account of mutual manipulability as a tool to settle debates about the extension of cognition. Krickel ultimately finds the framework wanting, concluding that a mutual manipulability account of integration into a cognitive system would include too much. She then amends Kaplan’s account in a more promising direction for refining theories of embedded and extended cognition. The details of Krickel’s account here are not crucial to our purposes. Instead, we want to emphasize that Krickel demonstrates how an explanatory and metaphysical framework of mechanisms can help move the 4E debate forward. By looking to ontology, Krickel can both diagnose problems in a 4E theory and also provide a remedy.

This isn’t the only way in which New Mechanism can underwrite 4E theories. For example, Miłkowski et al. (2018) holds that “going wide” for understanding cognitive mechanisms to include what’s beyond the body is crucial for the development of a mature theory of cognition. The next step forward in understanding mind, on their account, is to leave 4E theories behind and adopt a wide mechanistic view. For Miłkowski et al., the metaphysical thickets of 4E theories need clearing to make way for a metaphysics of mechanisms.

Both Krickel and Miłkowski et al. therefore point to ways in which New Mechanism helps move 4E discussions forward. New Mechanism provides a common conception of the metaphysical underpinnings of 4E theories. In doing so, it provides common currency by which to unify a range of 4E theories by means of a shared ontology of (New) mechanisms. Generally, we are sympathetic to the strategy adopted by the New Mechanists—looking to ontology when theorizing gets stale. But the New Mechanism isn’t the only way forward. An ontology grounded in causal powers—what we call causal powers realism—provides another route. And ultimately, we’ll argue, a better one.

## 4 From causal powers realism to mental powers

We have written elsewhere about causal powers realism (Vukov & Lassiter 2020). Here, we’ll limit our discussion of it to four central claims, and include citations for readers interested in going further:

First, according to causal powers realists, powers cannot be reduced to counterfactuals (e.g. Martin, 1994; see also our discussion in Vukov & Lassiter, 2020). For example, while salt’s solubility can certainly be described counterfactually (‘if you were to add the salt to water, it would dissolve’), causal powers realists deny that the salt’s solubility is reducible to this or any other counterfactual. According to causal powers realists, salt rather dissolves in water because of the powers salt and water have.

This brings us to the second claim of causal powers realism: powers are manifested in conjunction with manifestation partners (cf. Martin, 2007: Chapter 3;

Vukov & Lassiter, 2020).<sup>18</sup> According to causal powers realists, salt does not and *cannot* manifest its solubility on its own. Rather, salt manifests its solubility when it is conjoined with an appropriate partner—say, water’s power to dissolve salt—in the right conditions. What those conditions are and the precise nature of relevant partners is not merely a function of logic and definitions, but rather something that must be determined empirically.<sup>19</sup>

The third claim of causal powers realism follows on the heels of this insight. According to causal powers realists, causal powers do real causal work. They help account for how individuals behave (e.g. Jacobs, 2011). According to causal powers realists, causal processes must therefore be understood at least partly in reference to the powers individuals have. When salt dissolves in water, any complete description of this process must therefore reference the powers that salt and water have.

Finally, causal powers realists are committed to the idea that the manifestation of every power is itself empowering (Vukov & Lassiter, 2020). This claim stems from the idea that if powers are not mere counterfactuals, then powers are *directed towards* potential future manifestations: soluble individuals are directed towards dissolving; flammable individuals are directed towards igniting, and so on. It’s perhaps helpful to compare the directedness of powers with the directedness of mental states (Martin & Pfeifer, 1986; Place, 1996, Molnar, 2003; Jaworski, 2016). Intentional mental states are directed at things: your hope for a better life is *for* a better life, my hunger for an apple pie is *for* an apple pie, and so on. Intentional mental states, however, can remain unfulfilled: your hope for a better life may never transpire, my hunger for an apple pie may go unsatisfied, and so on. Similarly, then, in claiming that powers are directed towards their manifestations, causal powers realists are claiming that powers are *for* their various manifestations, even if those manifestations are never realized. So suppose some salt dissolves in water. According to the lexicon of causal powers realists, when salt dissolves, it manifests a power it has: its solubility. The manifestation of this power, however, is also a power itself, and so is itself directed towards potential future manifestations. Salinated water is empowered to properly make pasta in a way undissolved salt is not. Generally, according to causal power realists, every manifestation of a power is itself empowering (cf. Jaworski, 2016, p. 54).

The mechanics of causal powers realism have typically been developed in the context of chemical or physical powers: salt’s power of solubility, a vase’s power of fragility, and so on (e.g. Prior et al., 1982, Johnston, 1992). However, the applicability of causal powers to psychological phenomena has resulted in a growing literature devoted to exploring distinctively mental powers. From the perspective of a causal powers realist, psychological phenomena correspond to the manifestation of mental powers much as chemical phenomena correspond to the manifestation of chemical powers. Psychological phenomena can therefore be folded into the causal powers framework under the guise of mental powers. Recent work has seen proponents of mental powers applying the concept to topics in philosophy of mind ranging from

<sup>18</sup> Martin calls them “reciprocal disposition partners.”

<sup>19</sup> Thanks to an anonymous reviewer for the wording here.

emergence (Carruth, 2018) to phenomenal properties (Gozzano, 2018; Yates, 2018) to affordances (Vetter, 2018) to the problem of mental causation (Jaworski, 2016). Here, we set aside specific applications of mental powers to focus on the idea of mental powers generally—the idea that psychological phenomena are manifestations of powers, and can therefore be understood using the basic commitments of causal powers realism.

For a working example, take a token mental state of belief. Most proponents of mental powers will understand the tokening of belief as an individual manifesting a power it has, much as salt manifests its powers of solubility when it dissolves in water. In this respect, belief is most accurately thought of as a power people manifest rather than a state that people are in. Of course, proponents of mental powers can disagree about the precise nature of this power: for example, they may disagree about whether a belief is an emergent power (Paoletti, 2018) or if it can ultimately be reduced to fundamental physical powers (Bird, 2018; Heil, 2003). But almost all proponents of mental powers will agree a psychological manifestation such as belief is no inert property, but must rather be analyzed in terms of what it empowers an individual to do—the causal relations it enables the individual to enter into.<sup>20</sup>

Consider, then, how this understanding meshes with the commitments of causal powers realism as we've introduced it above. First, just as any power cannot be reduced to a counterfactual, so too belief cannot be counterfactually reduced, though it can be described that way. Some proponents of mental powers may insist that belief can be reduced to fundamental physical powers,<sup>21</sup> yet even for these causal powers realists, belief is no *mere* counterfactual—causal powers realists are decidedly not behaviorists.

Second, causal powers realists are committed to saying that psychological manifestations such as belief are manifested in conjunction with partners. Which partners? We'll discuss that below. Most obviously, though, belief is manifested in conjunction with things in the environment that have relevant manifestation partners. For example, my belief that there is a snake in my path is manifested in conjunction with the snake's power to induce that belief in me. As with all powers, the precise nature of these partner powers and what conditions are necessary for those powers to manifest themselves are determined empirically.<sup>22</sup>

Third, causal powers realists are committed to saying that psychological manifestations do real causal work. Belief is not only best discovered and studied by empirical disciplines; it is best understood in reference to the causal profile these disciplines uncover.

And finally, causal powers realists are committed to saying that belief, being a power, is itself empowering. My belief is not some inert state I find myself in. Rather, any token belief I have—for example, my belief that there is a snake in the path—empowers me to interact with my environment in ways I wouldn't had I not tokened that belief.

<sup>20</sup> Thanks to an anonymous reviewer for help with the phrasing here.

<sup>21</sup> Not us, though.

<sup>22</sup> Thanks again to an anonymous reviewer for help with the phrasing here.

That's not a complete picture of causal powers realism, nor is it a particularly detailed snapshot of how a causal powers framework can be applied to psychological phenomena. We have no doubt omitted details that some causal powers realists will find crucial and included details others will find controversial. Still, we believe both causal powers realism and mental powers as we have described them are true to the central commitments of the causal powers framework. Moreover, we have laid enough of a foundation to turn to one of our central purposes: to consider how causal powers realism can play a unifying role for 4E theories of cognition.

## 5 Why 4E theorists should embrace causal powers realism

We have two central contentions in this paper: that causal powers realism can unify the commitments of 4E theories, and that the framework causal powers realism provides does this work better than the New Mechanism. We argue for the first point in the current section, and for the latter point later. If our argument is sound, the upshot is that proponents of 4E theories should recognize an ally in the ancient ontology of causal powers. Before diving in, though, we should be clear: our argument is not that causal powers realism *entails* a 4E theory but rather that causal powers realism provides an elegant ontological background for the commitments 4E theorists already embrace.

Begin with Embodied Cognition, according to which bodily states are non-trivially incorporated into cognitive processing. As we have seen, causal powers realists understand psychological phenomena as the manifestation of powers, and further claim powers are best discovered and studied by the empirical disciplines. When causal powers realists are tasked with providing an account of a cognitive process, they will therefore look to the ways in which that process is studied empirically. They will ask: how do the empirical disciplines in fact study phenomena like fear or memory? What powers do the empirical disciplines postulate? Considered against the backdrop of causal powers realism, this amounts to the claim that bodily powers help constitute psychological manifestations. For proponents of Embodied Cognition, the empirical claim is that biological powers will non-trivially feature into our understanding of cognition. Importantly, causal powers realism does not *entail* Embodied Cognition. It is itself agnostic in regards to how these questions ultimately get answered. It is possible, for example, that all the causal powers necessary for understanding cognition are “in the head”—that all the powers, partners, and manifestation conditions pertaining to mental powers could be understood completely in terms of the contents of someone's skull.<sup>23</sup> In short: the claims made by Embodied Cognition are not entailed by causal powers realism. And because causal powers realists believe an account for cognitive processing must be empirical, whether or not *they* should accept Embodied Cognition is ultimately an empirical question. But for 4E theorists who already accept Embodied Cognition, causal powers realism

<sup>23</sup> Thanks to an anonymous reviewer for the phrasing here.

provides an elegant background in which their commitment to Embodied Cognition fits naturally.

Consider Environmental Cognition next, according to which conditions outside the body non-trivially figure into explanations of cognitive processing. For those who *already* accept Environmental Cognition, causal powers realism provides an intuitive framework. Recall that powers are manifested in conjunction with partners. Salt cannot manifest its solubility in a vacuum, but only in conjunction with the powers of water. Likewise, we manifest our powers—mental powers included—only in conjunction with relevant partners. Fear, anger, pain, desire, hunger: we do not manifest these powers on our own, any more than salt dissolves on its own. From within the framework of causal powers realism, Environmental Cognition means that the fear I manifest in the presence of a snake is not merely constituted by my internal processing, but also by the snake's power to produce the fear; that when I reflect on Christmas upon seeing a pine tree, the reflection is constituted by my perceptual and memorial powers *and* the pine tree's power to be seen; that when I listen to a Beethoven symphony, my perception is a manifestation of my powers to hear together with the symphony's power to be heard. Causal powers realists remain agnostic as to precisely what those partners are; the relevant sciences will fill in those details.<sup>24</sup>

When Environmental Cognition is hitched with causal powers realism, however, there is not merely empirical but ontological reason for incorporating the extrabodily world into explanations of cognition: causal processes must be understood at least partly in reference to the powers individuals have. According to causal powers realists, manifestation partners are *never* merely correlated, but are rather causal co-constituents of manifestations. When salt dissolves in water, the water's presence is not merely correlated with the salt's dissolving—rather, the water's powers help constitute the manifestation. Likewise, from the perspective of causal powers realism, to claim Environmental Cognition is true is to claim the presence of a snake is not merely correlated with my manifesting fear—rather, the snake's powers help constitute this manifestation. For the causal powers realist, manifestation partners are indeed correlated, but never *merely* correlated. When proponents of Environmental Cognition take causal powers realism on board, they therefore gain an ontological framework that undergirds their commitments.

Finally, consider Enactive Cognition. According to Enactive Cognition, there are undecomposable, co-constitutive, context-dependent feedback loops between perception, thought, and action. For those already committed to Enactive Cognition, causal powers realism also offers an elegant ontological background. For example, perception, on an enactivist view, doesn't happen in the brain but is rather about the brain and body embedded in an environment characterized by regularities and affordances (cf. Gallagher, 2018, p. 119). Putting the matter in terms of causal powers,

<sup>24</sup> That's not to say proponents of Environmental Cognition who take causal powers realism on board must be committed to the idea that environmental conditions are themselves constitutive of mental states. Environmental Cognition is an intentionally broad thesis, one compatible with both extended and embedded views of cognition: that conditions outside the body figure into explanations of cognitive processing beyond correlation with internal entities.

we would say that organisms and their subsystems have an array of powers; likewise, the environment is characterized by an array of powers. Enactivist perception, then, is a matter of the organism's powers partnering with the environment's powers to constitute a particular manifestation. But crucial for the supporter of Enactive Cognition, the power to perceive isn't unidirectional with information flowing from the world to the organism. Rather, it's bidirectional, with the organism's and the environment's powers bolstering and constraining one another. Consider a toy example: apple pie has the power to make us stop whatever we're doing and eat it.<sup>25</sup> My powers of vision and locomotion conjoin with the powers of the pie to be seen and be eaten to manifest my power to desire and eat the pie. That's how powers theorists who accept Enactive Cognition will understand the interaction. Causal powers realism thus provides an elegant background ontology for Enactive Cognition because causal powers realism is capable of taking on board the specific relationships among cognitive processes and environments that proponents of Enactive Cognition claim obtain.<sup>26</sup>

As with Environmental Cognition, causal powers realism does not *entail* Enactive Cognition. We have already seen the reason: causal powers realists believe powers are best discovered and studied empirically. For the causal powers realist, whether perception, thought, and action are constitutive of each other is thus an empirical question. Insofar as Enactive Cognition makes a claim about the empirical nature of cognitive processes, it could turn out that causal powers realism is true while Enactive Cognition is false. The empirical jury's still out. But if the jury rules in favor of Enactive Cognition, causal powers realism will be there to support the verdict.

Together with the rest of our discussion, these observations suggest that causal powers realism fits naturally with each of the central commitments of 4E theories, and in doing so, provides a unifying framework for them. This should be an interesting observation in itself: an ancient ontology provides an elegant framework for a contemporary research agenda. But that might not be enough to persuade 4E theorists to adopt causal powers realism. In what follows, we therefore turn back to the New Mechanism, arguing that causal powers realism can do more for 4E theorists. Our argument in what follows is thus an inference to the best explanation: those 4E theorists who are convinced of the need for a unifying ontology would do well to look past the New Mechanism to the ancient framework offered by causal powers realism.

<sup>25</sup> For one of us, this isn't far from the truth.

<sup>26</sup> We can make this more concrete. Proponents of Enactive Cognition often take dynamic systems theory to provide a better mathematical toolkit for describing cognitive processes than computational theory: minded organisms are more like water eddies than Turing machines. A very rough characterization of dynamic systems theory is that cognitive activity is constituted by interactions of organismic and environmental processes over time. If dynamic systems theory is true, however, causal powers realists ought to describe the jointly activated processes in terms of powers and as having a Gestalt effect, with the whole being greater than the individual components. From the perspective of dynamic systems theory, whatever metric we use to measure efficaciousness of powers, the contribution of each power to the behavior is not merely summative, but exponential. The lesson? Enactive Cognition fits especially well with causal powers realism.



## 6 Why 4E theorists should embrace causal powers realism instead of new mechanism

We have seen that causal powers realism satisfies our three 4E principles. And we also saw that New Mechanism can help push 4E debates forward. So why pick causal powers realism over New Mechanism?

In three sentences: New Mechanism is an adequate metaphysic for some but not all 4E theories. Causal powers realism is an adequate metaphysic for all 4E theories. Considerations of ecumenicism thus push us away from New Mechanism and towards causal powers realism.

Let's set the stage for this argument: in the wide spectrum of 4E theories, some aim to nudge the cognitive science status quo and others try to burn it to the ground. They are 'reformers' and 'radicals' (Lassiter, 2015). Goldman's bodily formats approach to embodied cognition is a reformer. He holds on to the tools of representations but works bodily representations into the story. Some of Andy Clark's work also counts as reformist: while looping Otto's notebook into his cognitive processes was a significant departure from mainstream thinking about cognition, his views about the effects of language on thinking are as radical as the Thatcher administration. Malafouris's (2014) material engagement theory, on the other hand, is a radical theory. It pulls together resources from anthropology and phenomenology to characterize how the culture materials we use shape our thinking. There's nary a representation to be found. Also found at the radical end is work by Maturana and Varela (1979) and di Paolo et al. (2018) among others.

Bring to mind one of New Mechanism's claims: mechanisms have definite spatio-temporal boundaries. This makes New Mechanism incline particularly towards the reformer end of the 4E spectrum. Goldman's embodied cognition is amenable to the ontology native to New Mechanism: representations have realizers with definite spatio-temporal boundaries. Even if Andy Clark's functionalist take on extended mind theory turns out to be wrong (e.g. Krickel, 2020), it's shown to be wrong using the tools of New Mechanism.

But New Mechanism's commitment to mechanisms with definite spatio-temporal boundaries doesn't fit well with many radical 4E theories. Consider, for example, recent work in cultural psychology, illustrating the effects of culture on cognition. Such work enables generalizations like, "members of tight cultures tend to follow rules even when there is no cost to breaking them" and "collectivist cultures prioritize the well-being of the group over that of the individual."<sup>27</sup> Not only are these generalizations well-validated—they also guide radical 4E theorizing. But a question remains: *how* exactly can cultures affect cognition? The New Mechanist perspective requires identifying mechanisms with definite spatio-temporal boundaries, but it's not obvious that that is the best way to characterize how culture affects cognition. Compare, for example, Bechtel and Abrahamsen's (2010) account of the mechanisms underwriting circadian rhythms with how tight cultures affect

<sup>27</sup> Gelfand (2018), Heine (2015).

judgment. Cultures affect judgment by means of agent-internal and agent-external bearers of cultural information which are dispersed across space and time (Vukov & Lassiter, 2020). And while New Mechanists are comfortable with mechanisms distributed across the brain, it's not clear that they are ok with mechanisms distributed across countries and centuries. Or rather, if New Mechanists *are* comfy with mechanisms distributed across space and time—as cultural forces seem to be—then they should let go of the commitment to mechanisms having definite spatio-temporal boundaries.

The lesson? New Mechanism is a satisfying metaphysic for reformist, but not radical, 4E theories. Causal powers realism, by contrast, is far more ecumenical. As argued above, it is a metaphysic that is suitable for all theories satisfying Embodied Cognition, Environmental Cognition, and Enactive Cognition. Causal powers realism remains agnostic about the definite spatio-temporal boundaries of powers, and so does not clash with a perspective that takes seriously the cultural effects on psychology.

Of course, New Mechanism has matured in ways that causal powers realism has not. For example, New Mechanism provides working scientists ways to identify and individuate mechanisms. That aspect of causal powers realism is still in the development phase. New Mechanism has resources to distinguish between mechanisms that are and aren't part of some phenomenon. Causal powers realism, for better or worse, leaves that in the hands of the empirical scientists. These are important issues that will eventually have to be developed. For now, though, we leave things at this: despite causal powers realism's need to develop in some ways, it has superior resources for describing culture's far-flung effects on human judgment. And this, in turn, provides 4E theorists reason to prefer causal powers realism to the New Mechanism.

## 7 Additional resources that causal powers realism provides to 4E theories

We opened this paper by suggesting 4E theorists have hunkered down into rival camps, and that causal powers realism can change things. We have offered our argument for causal powers realism largely as an alternative to the New Mechanism—if you are convinced that 4E theories need a unifying metaphysical theory and are tempted by the New Mechanism, our argument should give you serious reason to consider causal powers realism as an alternative. However, for those who are not already tempted by the New Mechanism or the general move to ontology, our argument may fall flat. We have argued that causal powers realism provides a better solution to a literature gone stale, but if you are neither motivated by the problem nor tempted by the general solution, you will have little reason to follow us to causal powers realism. In this section, we therefore argue that causal powers realism provides additional resources for 4E theorists, resources that may be either unavailable for, or else obscured from, the perspective of rival ontological frameworks. And this, we believe, gives 4E-minded scholars further reason to pay attention to the ontology.

First, consider that 4E theories tend to resist a strict mental-physical divide—this is especially true for proponents of Enactive Cognition but also holds for

Clark-style extended cognition and (to a lesser extent) lite-embodiment theorists like Goldman. Causal powers realism should be attractive to 4E theorists in providing a framework that, for purely ontological reasons, also resists this way of carving up nature. According to proponents of powers, mental powers correspond to psychological phenomena in exactly the same way that physical powers correspond to physical phenomena. For the causal powers realist, psychological manifestations are the product of mental powers just as chemical manifestations are the product of chemical powers. Generally, mental and physical vocabularies can be helpful for the causal powers realist, just as it can be helpful for causal powers realists to refer to distinctive physical, chemical, and neurobiological powers. But these distinctions are only skin-deep. For the causal powers realist, powers are powers before they are chemical or biological, mental, or physical.<sup>28</sup> Strict demarcations are largely for university administrators. Jaworski (2018) argues for this view generally, and Marmodoro and Grasso (2020) apply it specifically to color perception.<sup>29</sup> Causal powers realism thus provides a ready-made framework of psychology and perception that resists a strict mental-physical divide. By taking on board causal powers realism, contemporary 4E theories can gain an ontological framework capable of supporting a move many of them already make. That's not to say rival ontologies cannot also reject the mental-physical divide. Some can. But causal powers realism is especially sympathetic to and provides recent resources for grounding such a move. And that, we believe, should count in its favor for 4E theorists.

A second advantage that causal powers realism provides contemporary 4E theories is its exceptional openness to empirical inquiry. In short: causal powers realism is empirically friendly. By 'empirical friendliness,' we don't mean that causal powers realism entails any specific empirical results, but rather that the framework is compatible with the empirical results, no matter how they turn out. Contemporary 4E theories are almost without exception reliant on empirical inquiry. One need only think of the modern titans of 4E theorizing—Varela, Thompson, and Rosch; Clark; Brooks; Dreyfus; de Jaegher; Gibson; Merleau-Ponty; Smith and Thelen; Dewey; Gallagher. All are either psychologists with philosophical interests or philosophers with psychological interests. Causal powers realism provides a framework for their empirical openness. In punting to the sciences to understand cognitive powers, the only empirical bias a causal powers realist is allowed is a bias towards the best available science. By adopting the ontology, 4E theorists can therefore gain a framework for the empirical openness they already embrace. Again, rival ontologies do not necessarily reject empirical-friendliness of the kind with which we are concerned. Many are even sympathetic to it. Our point here is simply that causal powers realism ought to be counted among the most vocal sympathizers of empirically-friendliness.

---

<sup>28</sup> An example of this is increasing recognition of the influence of culture on evolution. Culture-gene coevolution suggests that distinct biological and cultural vocabularies to explain evolutionary processes is unhelpful. Rather, what's needed is a way to talk and think about genes as shaping cultural processes over time and cultures shaping genes and heritability over time. Cf. Boyd and Richerson (1985), Richerson and Boyd (2005).

<sup>29</sup> Thanks to an anonymous reviewer for directing us to these articles.

The third sort of resources causal powers realism offers to 4E theorists is through its sympathy to a kind of causal pluralism, one that can integrate different kinds of forces constituting psychological phenomena. In general, 4E theories appeal to a variety of causes in accounting for cognition—biological causes (Thompson, 2007), psycho-social causes (Menary, 2007), neurological causes (Anderson, 2014a, b), cultural causes (Malafouris, 2014), and so on. But it isn't always clear how these causal contributors are related to each other, how biology, culture, and neurochemistry 'hook up' to produce psychological phenomena. Causal pluralism, we believe, can provide this clarity.

The kind of causal pluralism we have in mind here is a relatively simple idea—that it may not always be appropriate to talk about *the* cause of a manifestation, since genuinely causally-efficacious influences are relevant only in certain contexts. Consider a non-psychological example: suppose two groups of investigators—a group of engineers and a group of political theorists—are tasked with explaining the sinking of the *Luisitania*, the Cunard liner whose torpedoing helped push the US into WWI. When the political theorists investigate why it sank, they will invoke social-political causes: the intentions of Germany and England, the wartime strategy of targeting civilian vessels, and so on. The engineers, by contrast, will cite very different causes: they may appeal to the hull's construction, the torpedo's blast, the cargo's placement, and so on. Start talking politics with the engineers, and you'll get a blank stare. Not because the engineers doubt the relevance of politics to the ship's sinking but rather because political causes are not relevant to their investigation. *Qua* engineers, they are not aiming to grasp the subtleties of US-England-German relations in the early twentieth century. They are instead aiming to determine how the liner's construction contributed to its sinking. How to interpret this? According to the causal pluralist, there are mechanical *and* political causes that help constitute the sinking of the *Luisitania*, and only some of these causes are relevant to certain investigations, depending on the context. So it doesn't always make sense to talk about *the* cause of its sinking.

Likewise with psychological phenomena. For example, suppose you react in fear to a clown, and that several groups of investigators—social psychologists, biologists, and neuroscientists—are tasked with explaining your reaction. Each will inevitably invoke different kinds of causes in their explanation: the social psychologists will perhaps mention socio-cultural associations that clowns carry in twenty-first century America; the biologists may invoke the fact that you haven't been sleeping well and are currently hungry; the neuroscientists may invoke patterns of activation in your amygdala and prefrontal cortex; and so on. This causal cacophony reflects the host of causes invoked by the 4E literature. For the causal pluralist, it stems from the fact that there are many causes of human psychology and that these causes are relevant to only certain investigations, depending on the context.

Back to causal powers realism. Causal pluralism does not *follow from* causal powers realism. But causal pluralism sits comfortably with it. As we have seen, causal powers realists are eager to allow the sciences to account for manifestations—they let empirical disciplines take the lead in discerning what causes salt to dissolve in water, and what causes your fear of the clown. So whether causal pluralism is true depends on how the empirical work turns out. Maybe your coulrophobia

is exhaustively explained by activity in the prefrontal cortex, but maybe not. As we have seen, however, 4E theorists are often committed to a kind of causal pluralism. And for those who are *already* committed to causal pluralism, causal powers realism offers an especially elegant framework for their commitment: from this perspective, the seemingly rival causal explanations offered by social psychology, biology, and neuroscience are not in competition with each other. Instead, the interests of investigators uncover a wide range of causal powers involved in the manifestation, all of which are partners that help constitute it and none of which needs to vye for the status of being *the* cause. Causal powers realism thus allows for the integration of different causes contributing to manifestations of psychological phenomena. That is precisely what the causal pluralist wants. And insofar as 4E theorists embrace causal pluralism, it is what they want too.

Causal powers realism, therefore, may not entail causal pluralism, but fits comfortably with it. As with other points we have made above, this does not mean that rival ontologies cannot account for causal pluralism. Some can. But insofar as causal powers realism can be counted among these ontologies—and given everything else that causal powers realism can do for 4E theories—this gives 4E theorists further reason to take the ontology seriously.

## 8 Conclusion

We will close our discussion here. We have suggested that causal powers realism can unify the core commitments of 4E theories, and that it can do this work better than its most obvious contemporary rival—the New Mechanism. 4E theories do not entail causal powers realism; nor does causal powers realism entail any 4E theory. Nonetheless, given antecedent commitment to any or all of the central 4E claims we've identified, causal powers realism offers a compelling background ontology for developing and refining 4E theories.

There is one final, but more speculative, advantage we would like to discuss before finishing. As a unifier of 4E theories, causal powers realism can highlight not only points of agreement between 4E theories—through a new lingua franca, it can also highlight particular points of disagreement among them, disagreements that can lead to 4E siloing of the kind we highlighted in the introduction. Consider just one example: Goldman-style lite embodiment theorists and enactivists agree the body is involved in cognition, but disagree about the nature of this involvement. Causal powers realism helps clarify the terms of their disagreement: the groups disagree about the nature of the relevant powers involved in cognition. Lite embodiment theorists hold that perceptual powers integrate with bodily powers to create bodily formats for representations. Bodily powers, on this view, affect the structure of the representation and not necessarily its contents. Enactivists, by contrast, think perceptual powers integrate with bodily powers to integrate further with environmental powers. One way to understand the difference between lite embodiment and enactivism is therefore in reference to the way in which they understand perceptual powers: do these powers have the power to alter the structure of perceptual representations or

do they integrate with other bodily and environmental powers? Armchair philosophy will not settle the question, but theorizing with powers helps sharpen the issues.<sup>30</sup>

## Appendix

The data were downloaded from Web of Science on January 15, 2020. The search term used was TS=("extended mind\*" OR "extended cognition" OR "enactive mind\*" OR "enactive cognition" OR "enactivism" OR "embodied mind\*" OR "embodied cognition" OR "embedded mind\*" OR "situated mind\*" OR "embedded cognition" OR "situated cognition"). This returned a file of 3,867 entries for the years 1960–2019. We included explicitly philosophical journals in our analysis and excluded journals that were primarily dedicated to another discipline. E.g. *Philosophical Psychology*, *Phenomenology and the Cognitive Sciences*, and *Philosophical Studies* (among others) were included, but *Frontiers in Psychology*, *Cognitive Science*, and *Cognitive Processing* (among others) were excluded. This returned a list of 556 papers, whose combined bibliographies contain over 14,000 items.

For the co-citation network, only the top 30 were plotted (for reasons of space, readability, and computational power). Two sources are linked iff they were co-cited at least 10 times among the 556 papers. The size of the node is scaled to its centrality: the bigger the node, the more central it is to the network.

For the keyword network, analysis shows that there were 35 cliques in the entire network, using the fast and greedy community detection algorithm in the package *igraph*. Clearly, that would render the plot unreadable were we to try to account for all these. Instead, we plotted the three largest cliques, each of which has over 250 (non-unique) members. Trial and error shows that plotting the top 47 keywords maximizes the number of members per clique while keeping the number of cliques to three. Trying to do the same for the top four cliques would require plotting 99 keywords, which obscures the information the plot is trying to convey: namely, that enactivism, embodied cognitive science, and extended cognitive sciences constitute three distinct (but overlapping) areas of research. Finally, we chose to “play the keywords as they lie” as much as was possible, correcting misspellings and homogenizing British and American spellings. For example, “computation” and “computational” are distinct keywords in the dataset, but we can imagine them being used differently, so we didn’t hazard a guess and which would be lumped with which. Since we’re trying to get the lay of the land rather than develop (say) a predictive model, it is best if we don’t alter the terrain too significantly.

Bibliometric data was analyzed using the *bibliometrix* package for R 4.0.2 (Aria & Cuccurullo, 2017).

---

<sup>30</sup> Thanks to anonymous reviewers for their invaluable feedback on earlier drafts. In particular, we thank Marcin Miłkowski for a number of suggestions that helped frame our discussion, and a special thanks to our (volunteer!) research assistant Julia Taylor for all of her help. Thanks, too, to Abram Capone for comments on an earlier draft. Finally, CL would like to thank Michele Lassiter and JV would like to thank Kelsey Vukov for their years of support and encouragement.

## References

- Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14, 43–64.
- Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell.
- 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.
- Andersen, H. (2014a). A field guide to mechanisms: Part I. *Philosophy Compass*, 9, 274–283.
- Andersen, H. (2014b). A field guide to mechanisms: Part II. *Philosophy Compass*, 9, 284–293.
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975.
- Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology*, 22(5), 543–564.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441.
- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321–333.
- Bechtel, W., & Richardson, R. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research* (2nd ed.). Cambridge, MA: MIT Press/Bradford Books.
- Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford: Oxford University Press.
- Bird, A. (2018). I—fundamental powers, evolved powers, and mental powers. *Aristotelian Society Supplementary*, 92, 247–275.
- Boyd, R., & Richerson, P. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Brigandt, I. (2010). Beyond reduction and pluralism: Toward an epistemology of explanatory integration in biology. *Erkenntnis*, 73(3), 295–311.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Carruth, A. D. (2018). Emergence, reduction and the identity and individuation of powers. *Topoi*. <https://doi.org/10.1007/s11245-018-9621-x>
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 7–19.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Clark, A. (2010). Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa. In R. Menary (Ed.), *The extended mind* (pp. 81–100). Cambridge, MA: The MIT Press.
- Clark, A. (2011). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–253.
- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. New York: Oxford University Press.
- Deulofeu, R., Suárez, J., & Pérez-Cervera, A. (2019). Explaining the behaviour of random ecological networks: The stability of the microbiome as a case of integrative pluralism. *Synthese*, 1–23.
- di Paolo, E., De Jaeger, H., & Cuffari, E. C. (2018). *Linguistic bodies: The continuity between life and language*. Cambridge, MA: MIT Press.
- Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.
- Ellis, B. (2001). *Scientific essentialism*. Cambridge: Cambridge University Press.
- Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 3(1), 63–73.
- Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
- Fodor, J. (2008). *LOT 2: The language of thought revisited*. New York: Oxford University Press.
- Gallagher, S. (2005). *How the body shapes the mind*. New York: Oxford University Press.
- Gallagher, S. (2018). *Enactivist interventions*. New York: Oxford University Press.



- Gelfand, M. (2018). *Rule makers, rule breakers: How tight and loose cultures wire our world*. New York: Scribner.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. New York: Psychology Press.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342–S353.
- Glennan, S. (2010). Mechanisms, causes, and the layered model of the world. *Philosophy and Phenomenological Research*, 81(2), 362–381.
- Goldman, A. I. (2012). A moderate approach to embodied cognitive science. *Review of Philosophy and Psychology*, 3, 71–88.
- Gozzano, S. (2018). The dispositional nature of phenomenal properties. *Topoi*. <https://doi.org/10.1007/s11245-018-9597-6>
- Heil, J. (2003). *From an ontological point of view*. Oxford: Oxford University Press.
- Heine, S. (2015). *Cultural psychology* (3rd ed.). New York: WW Norton & Co.
- Hutto, D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge, MA: The MIT Press.
- Hutto, D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. Cambridge, MA: The MIT Press.
- Kagan, A., & Lassiter, C. (2013). Much ado about nothing: The coupling-constitution fallacy. *Pragmatics & Cognition*, 21, 178–192.
- Kirchoff, M. (2012). Extended cognition and fixed properties: Steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences*, 11, 287–308.
- Kohler, I. (1951). Über Aufbau und Wandlungen der Wahrnehmungswelt. *Österreichische Akademie der Wissenschaften. Sitzungsberichte, philosophisch-historische Klasse*, 227, 1–118.
- Newen, A., de Bruin, L., & Gallagher, S. (2018). 4E cognition: Historical roots, key concepts, and central issues. In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 3–16). New York: Oxford University Press.
- O'Regan, K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 939–973.
- Jacobs, J. D. (2011). Powerful qualities, not pure powers. *The Monist*, 94(1), 81–102.
- Jaworski, W. (2016). *Structure and the metaphysics of mind: How hylomorphism solves the mind-body problem*. Oxford: Oxford University Press.
- Jaworski, W. (2018). Psychology without a mental-physical dichotomy. In W. M. R. Simpson, R. C. Koons, & N. J. Teh (Eds.), *Neo-Aristotelian perspectives on contemporary science* (pp. 261–292). New York: Routledge.
- Johnston, M. (1992). How to speak of the colors. *Philosophical Studies*, 68, 221–263.
- Kaplan, D. (2012). How to demarcate the boundaries of cognition. *Biology and Philosophy*, 27(4), 545–570.
- Krickel, B. (2020). Extended cognition, the new mechanists' mutual manipulability criterion, and the challenge of trivial extendedness. *Mind & Language*, 35(4), 539–561.
- Lassiter, C. (2015). Review of *cognition beyond the brain*. *Philosophical Psychology*, 8, 1245–1249.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1–25.
- Malafouris, L. (2014). *How things shape the mind: A theory of material engagement*. Cambridge, MA: MIT Press.
- Marmodoro, A., & Grasso, M. (2020). The power of color. *American Philosophical Quarterly*, 57(1), 65–78.
- Martin, C. B. (1994). Dispositions and conditionals. *The Philosophical Quarterly*, 44(174), 1–8.
- Martin, C. B. (2007). *The mind in nature*. Oxford: Oxford University Press.
- Martin, C. B., & Pfeifer, K. (1986). Intentionality and the non-psychological. *Philosophy and Phenomenological Research*, 46, 531–554.
- Maturana, H. R., & Varela, F. J. (1979). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel Publishing Company.
- Menary, R. (2007). *Cognitive integration: Mind and cognition unbound*. New York: Palgrave Macmillan.
- Menary, R. (2010). Cognitive integration and the extended mind. In R. Menary (Ed.), *The extended mind* (pp. 227–244). Cambridge, MA: The MIT Press.
- Millikan, R. G. (2009). Embedded rationality. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 171–181). New York: Cambridge University Press.
- Mitchell, S. (2003). *Biological complexity and integrative pluralism*. Cambridge: Cambridge University Press.

- Molnar, G. (2003). In S. Mumford (Ed.), *Powers: A study in metaphysics*. Oxford: Oxford University Press.
- Paoletti, M. P. (2018). Emergent powers. *Topoi*. <https://doi.org/10.1007/s11245-018-9585-x>
- Place, U. T. (1996). Dispositions as Intentional States. In T. Crane (Ed.), *Dispositions: A Debate* (pp. 19–32). London: Routledge.
- Prior, E., Pargetter, R., & Jackson, F. (1982). Three theses about dispositions. *American Philosophical Quarterly*, 19, 251–257.
- Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. Cambridge, MA: MIT Press.
- Rupert, R. (2009). *Cognitive systems and the extended mind*. New York: Oxford University Press.
- Shapiro, L. (2011). *Embodied cognition*. New York: Routledge.
- Simon, H. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
- Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics & Cognition*, 14, 235–247.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the science of the mind*. Cambridge, MA: Harvard University Press.
- Thompson, E., Varela, F., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science*, 17, 7–48.
- Vetter, B. (2018). Perceiving potentiality: A metaphysics of affordances. *Topoi*. <https://doi.org/10.1007/s11245-018-9618-5>
- Vukov, J., & Lassiter, C. (2020). How to power encultured minds. *Synthese*, 197, 3507–3524.
- Yates, D. (2018). Neural synchrony and the causal efficacy of consciousness. *Topoi*. <https://doi.org/10.1007/s11245-018-9596-7>
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.
- Miłkowski, M., Clowes, R., Rucińska, Z., Przegalińska, A., Zawidzki, T., Kreuger, J., Gies, A., McGann, M., Afeltowica, Ł., Wachowski, W., Stjernber, F., Loughlin, V., & Hohol, M. (2018). From wide cognition to mechanisms: A silent revolution. *Frontiers in Psychology*, 9, 2393. <https://doi.org/10.3389/fpsyg.2018.02393>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.