

## In which sense should we be radical? On two radical alternatives to cognitivism

José Carlos Camillo

Université de Genève

[josecarloscamillo@gmail.com](mailto:josecarloscamillo@gmail.com)

### Abstract

Cognitivism, a framework strongly shaped by developments in computer and communication sciences, has dominated the study of cognition since the 1960s. From the 1990s onward, however, alternative frameworks have gained prominence. Among these, anti-cognitivism designates approaches that deliberately reject cognitivist assumptions, with some examples found in Radical Embodied Cognition, Radical Enactivism, and Autopoietic Enactivism. In this paper, a distinction between pluralist and eliminativist anti-cognitivism is proposed. Pluralists regard cognitivist and non-cognitivist approaches as, at least in principle, complementary, while eliminativists argue that cognitivism is misguided and should be discarded. The paper starts by clarifying cognitivism to avoid common misunderstandings. After that, focusing on enactive and embodied versions of anti-cognitivism, it analyzes its most influential arguments against cognitivism and concludes that eliminativism fails to provide a convincing case against cognitivism, and that pluralism better reflects the epistemological commitments of at least some versions of anti-cognitivism, namely those that adopt an ecological perspective.

Keywords: Radical Embodied Cognition; Radical Enactivism; Cognitivism; Representations; Scientific explanations; Scientific concepts.

### 1 Introduction

Around the 1960s, cognitivist frameworks became dominant in the cognitive sciences. *Cognitivism* refers to a set of approaches to studying cognition strongly influenced by developments in computer and communication sciences (Gallistel & King, 2009; Colombo & Piccinini, 2024). Since the 1990s, however, alternative frameworks have gained prominence.<sup>1</sup> Some of these aim to address limitations of cognitivism while preserving its core commitments. Others, often labeled “radical,” reject those commitments entirely. I will refer to this latter

---

<sup>1</sup> This is not to suggest these alternative views emerged solely in opposition to cognitivism or are exclusively recent developments. The intent is to highlight that, following the dominance of cognitivism post-1960, other frameworks were often treated with caution and viewed as peripheral. After 1990, however, these perspectives gained more significant influence within the field.

stance as an *anti-cognitivist commitment*, meaning a deliberate break from cognitivist assumptions. The most prominent anti-cognitivist views include *Radical Embodied Cognition*, *Radical Enactivism*, and *Autopoietic Enactivism*. Recent work has sought to clarify the aims and scope of some of these anti-cognitivist frameworks. For instance, some argue that these frameworks are not frameworks for the sciences of the mind, but rather for a philosophy of nature (Meyer & Brancazio, 2022), and others distinguish between scientific and utopian forms of anti-cognitivism (Meyer & Brancazio, 2023). In this paper, I propose an additional distinction. Namely, I suggest we can differentiate two forms of anti-cognitivism: pluralist and eliminativist. Pluralist anti-cognitivism acknowledges the value of non-cognitivist approaches while also recognizing the contributions of cognitivism. It maintains that both anti-cognitivism and cognitivism highlight different, yet complementary, aspects of cognition. Eliminativist anti-cognitivism, by contrast, views cognitivism as flawed and argues that it should be abandoned altogether as a framework for studying cognition. I argue that, at present, the case for eliminativism is unconvincing, and that pluralism better aligns with the epistemological commitments of at least some of anti-cognitivist frameworks.

Some clarifications are necessary regarding the target of my analysis, which I call *anti-cognitivism*. Firstly, the prefix “*anti*” should not be read too strongly. If it were, every form of anti-cognitivism would amount to eliminativism. One might instead use *acognitivism* to indicate the mere absence of cognitivism. However, even pluralist anti-cognitivists make a deliberate effort to avoid cognitivist assumptions, which is why I adopt the term *anti-cognitivism* (see Section 4). Secondly, there are multiple forms of anti-cognitivism. In this paper, I focus only on those that endorse embodied or enactivist commitments. One could say that my real target is *enactivism* more broadly. I avoid this label, however, because some versions of enactivism do not adopt anti-cognitivism, whether in pluralist or eliminativist forms (e.g., Noë, 2004). For this reason, I will use the term *enactive/embodied anti-cognitivism* (eAC) to designate my specific target. Thirdly, since its revival in the 1990s, eAC has advanced many arguments against cognitivism. It is impossible to address them all in a single paper. My aim is instead to examine the main arguments most frequently cited and relied upon in the literature. Finally, cognitivism itself is not a unified framework. As Meyer and Brancazio (2023) note, many anti-cognitivist arguments fail because they overlook this fact. Still, following Raja (2024) and others, I take most cognitivist approaches to share a set of core assumptions despite their internal diversity. I therefore interpret anti-cognitivism as the rejection of these shared assumptions. Importantly, since not all cognitivist frameworks endorse all of these assumptions,

I treat cognitivism here as a property cluster (similar to Boyd’s 1991 proposal) rather than as a list of necessary and sufficient conditions.

The paper is structured as follows. In Section 2, I introduce cognitivism and clarify several of its core tenets, some of which are often overlooked or misunderstood by anti-cognitivism advocates. In Section 3, I present a brief example illustrating the success of cognitivism. With these foundations in place, Section 4 outlines two distinct attitudes when defending radical alternatives to cognitivism: pluralist and eliminativist. Recognizing that the debate between cognitivism and anti-cognitivism may hinge on conceptual issues, Sections 5 through 7 examine the arguments eliminativists offer for discarding cognitivism. Finally, Section 8 concludes by reaffirming that pluralism is more consistent with the epistemological outlook of at least some versions of anti-cognitivism.

## **2 Clarifying cognitivism**

As discussed, anti-cognitivism rejects cognitivism, the dominant framework—or set of frameworks—in the cognitive sciences.<sup>2</sup> However, anti-cognitivist portrayal of cognitivism is often too narrow and superficial, making it an easy target for criticism. While many do acknowledge cognitivism’s main tenets and its empirical successes, their critiques frequently rely on misunderstandings or oversimplifications. In this section, I offer a more charitable presentation of cognitivism to help avoid common misconceptions.

Some of eAC’s advocates often stress their framework’s continuity with the physical sciences, at times labeling opposing views as “magical” (e.g., van Gelder, 1998; Chemero, 2009; Raja, 2018; Favela, 2024). Especially in versions influenced by ecological psychology, some eAC frameworks seek to explain behavior through laws, much like Newtonian physics. Although cognitivism does not pursue laws in the same strict sense, it does aim to identify stable generalizations about cognitive phenomena. Many of its proponents even aspire to exceptionless generalizations, which, in the philosophy of science, qualify as laws (see Pietroski & Rey, 1995; Woodward & Hitchcock, 2003; Andonovski, 2022; De Brigard, 2024). Some eAC theorists might insist that cognitivism remains distinct because its generalizations are not reducible to physical laws or to the explanatory tools used in physics. For example, Raja (2018)

---

<sup>2</sup> As Raja (2024) notes, cognitivism comprises a range of distinct frameworks that share certain common approaches to questions in the cognitive sciences. He refers to these shared elements as research *motifs*.

criticizes cognitivism as “magical” for explaining visual perception without relying on the laws or tools of optics.<sup>3</sup>

Yet, even if cognitivism is not continuous with physics in the way eAC proponents prefer—i.e., through shared laws or explanatory tools—it still aligns with physics in a different, often overlooked, sense. Each scientific field relies on a set of methods and underlying principles that justify how those methods are used to investigate the world (Massimi, 2022). The proponents of anti-cognitivism often overlook a key methodological principle at the heart of cognitivism. As a result, they portray cognitivism as lacking proper justification, which makes it an easier target for critique. The principle in question is a methodological reductionism that is well exemplified by what Chomsky (1978, 1980) called the *Galilean explanatory style* (see Allott et al., 2021). To understand this style, it is helpful to revisit Galileo’s classic example involving free fall.

In his theory of free fall, Galileo proposed that all objects fall to Earth with the same acceleration, regardless of mass or volume. The well-known thought experiment of dropping two objects of different masses from the top of the Tower of Pisa, predicting that they would hit the ground simultaneously, illustrates this idea. Of course, everyday experience shows that falling objects are influenced by many factors: air resistance, wind, and even electromagnetic fields.<sup>4</sup> Because of this complexity, the real-world behavior of falling objects is often modeled using dynamical systems (e.g., Xie & Miyata, 2014).

However, Galileo’s goal was not to account for the full behavior of falling objects. Instead, he focused on a specific explanatory regularity common to all such cases—the uniform acceleration now known as due to gravity. He isolated this factor by using thought experiments and by designing scenarios that deliberately excluded complicating factors such as air resistance and friction. One might say, then, that he adopted a *ceteris paribus* generalization: all else being equal, falling objects accelerate at the same rate (Pietroski & Rey, 1995; Allott et al., 2021). His model was thus highly abstract and idealized, intentionally stripping away other variables in order to theorize about one key factor (Andonovski, 2023). In other words, his theory has not as object the behavior of falling objects “in the wild”. Rather, it has as object a particular explanatory factor that explain the fall of objects, namely the uniform acceleration.

In sum, the Galilean explanatory style can be summarized in four steps:

---

<sup>3</sup> It is important to acknowledge that this kind of theoretical reductionism is highly controversial (see Craver, 2007; Bechtel, 2008). Moreover, even these anti-cognitivism’s proponents who criticize cognitivism for failing to reduce its explanations to physical laws (labeling it “magical”) do not necessarily aim to reduce cognitive science—or ecological psychology—to physics themselves.

<sup>4</sup> Throughout this paper, I use the terms factors and variables interchangeably.

1. The target behavior occurs within a complex system and is shaped by many interacting factors.
2. One universal explanatory factor is identified across all instances of the behavior.
3. Theory is developed around this single factor, not the overall complex behavior.
4. Strong abstraction and idealization are used to exclude other explanatory factors from the model.

Here, universal (in point 2) simply means that the factor is present in all instances of the phenomenon. The history of physics, which ecological versions of anti-cognitivism often celebrate, demonstrates that the Galilean style has been both epistemically and pragmatically successful (Chomsky, 1980; Allott et al., 2021). For instance, Chang (2012) argues that modern chemistry relies almost entirely on a comparable form of methodological reductionism. Even physics—despite the exponential growth of dynamical models in recent decades—still makes extensive use of, and depends heavily on, similar methodological principles (Massimi, 2022; Suárez, 2024). That is, it has revealed truths about the natural world and has guided the development of new experiments and hypotheses.<sup>5</sup>

This methodological principle is central to cognitivism.<sup>6</sup> For example, Andonovski (2023) recently argued that research on memory systems follows the Galilean style. Cognitivists recognize that behavior results from a complex, dynamic interplay of many factors (point 1). Yet, they isolate a single factor, which they label *cognition*, that appears across all human behaviors (point 2), and they theorize about this factor rather than behavior in its full complexity (point 3). Consequently, although cognition is posited to help explain behavior, cognitivists do not aim to explain behavior as such, but rather the specific cognitive processes that underlie it. To do so, they, like Galileo, rely on heavy abstraction and idealization (point 4).<sup>7</sup>

As with Galileo's free fall theory, cognitivist models typically provide accurate predictions only under highly controlled laboratory conditions if we disregard its highly abstracted and idealized nature. The prediction that two objects of different mass and volume

---

<sup>5</sup> "Revealing truths about the world" is a realist reading. An instrumentalist, however, would emphasize the hypothesis's success in generating accurate predictions or measurements, rather than its claim to truth.

<sup>6</sup> In fact, the Galilean style of explanation **is behind** many—if not all—reductionist and localized approaches, which includes cognitivism. Here, "reductionism" is used in a weak sense. It does not imply that all explanations must be found at the most fundamental level, but rather that investigating a phenomenon's lower levels is one valid explanatory strategy among others (cf. Silberstein & Chemero, 2013; Chirimuuta, 2024).

<sup>7</sup> It is also important to anticipate a key point discussed in Section 4: when cognitivists argue that external factors are not relevant for cognition (e.g., Rupert, 2005; Adams & Aizawa, 2010), they are working with a specific, internalist definition of cognition. **However, they would never disagree that such factors are explanatory for behavior.** In contrast, anti-cognitivism theorists typically adopt a broader definition, one that incorporates a wider range of explanatory factors within a single model (Potochnik & Sanches de Oliveira, 2020; Favela, 2024).

will reach the ground simultaneously holds true only in a vacuum chamber—not in everyday environments. Similarly, cognitivist theories are not primarily designed to predict behavior “in the wild,” but rather to isolate and investigate specific factors under idealized conditions. This limitation does not undermine their scientific value. Both free fall theory and cognitivism are successful, fruitful, and in many cases accurate—they are simply focused on narrow, carefully disentangled aspects of reality.

One implication of this methodological approach is that applying theories like cognitivism or free fall theory directly to real-world contexts requires caution. Cognitivists themselves recognize this. For example, in memory research, Tulving (2005; cf. Andonovski, 2023) acknowledged that real-life remembering often does not fit neatly into his division of memory systems. Likewise, Schacter (2002) urged caution when applying findings about memory distortions—derived from laboratory experiments—to real-world situations, particularly in legal contexts.

Understanding cognitivism as a Galilean science, in the sense outlined above (points 1–4), helps clarify why two common criticisms often miss the mark. The first is illustrated by Gallagher’s (2018a, 2018b) defense of strong embodiment and his proposal to decenter the brain in cognitive science. Gallagher argues against weak embodiment—the view that the body matters for cognition only insofar as it is represented in the brain—by noting that the brain co-evolved with the body and the environment and is influenced by both. On this basis, he claims that behavior cannot be properly explained without reference to bodily and environmental factors (2018b). Similarly, he provides extensive evidence that human behavior results from the interaction of environmental and bodily processes, advocating for a broader view that shifts attention away from a brain-centered model (2018a).

However, as the Galilean principle makes clear, cognitivists do not deny the role of these other factors for behavior. They fully acknowledge that behavior arises from complex interactions involving the brain, body, and environment. Their aim, however, is not to model the full complexity of behavior, but to theorize about one particular factor—cognition—which they locate primarily in the brain. Thus, Gallagher’s critique does not, in principle, threaten cognitivism. It simply reflects a different methodological focus.<sup>8</sup>

---

<sup>8</sup> As will be discussed in Section 4, cognitivists and proponents of anti-cognitivism use the term “cognition” to label distinct sets of factors, which may lead to a verbal dispute. For cognitivists, “cognition” refers to factors limited to the brain, whereas for eAC theorists, it is almost conflated with “behavior”—a broader category requiring additional explanatory factors. Crucially, cognitivism does not take behavior itself as its object of study or theorizing; rather, it focuses solely on brainly/computational factors that explain behavior.

The second common criticism of cognitivism is that its theories often fail to predict directly real-world behaviors (e.g., Baggs & Sanches de Oliveira, 2024). As explained above, however, this is not the goal of cognitivism. Just as Galileo’s free fall hypothesis holds accurate predictions only in vacuum conditions—not “in the wild”—yet remains scientifically valuable, so too cognitivist theories aim to isolate and explain specific aspects of cognition which results in limiting its predictive capacity in the wild. The limited scope of their predictive application is not, in itself, a flaw. Interestingly, dynamical models—often proposed as alternatives—also fail to fully predict real-world behavior (Raja, 2018, 2021). In their case, however, this failure is more serious because their stated aim is to model behavior *in the wild*. If they fall short in doing so, the theory fails on its own terms. Cognitivism, by contrast, is not committed to capturing the full complexity of real-world behavior but rather to theorizing about a specific factor, which is only possible under idealized conditions.

The reader will have noticed that, in defending cognitivism from these two criticisms, I noted they are not problems *in principle*. This qualification matters. These criticisms do not undermine cognitivism as a Galilean science, understood according to the four steps outlined above. Of course, other aspects of cognitivism may still be incompatible with these criticisms—but anti-cognitivist theorists have not yet demonstrated this.

With this key methodological point clarified, I now offer a general definition of cognitivism. Before proceeding, it is crucial to remind the reader that this definition is not a set of necessary and sufficient conditions. As established in the introduction, it is a property cluster: it describes a collection of properties that cognitivist frameworks typically, but not universally, share. Cognitivism maintains that cognition:

- i) is (primarily) located in the central nervous system;
- ii) can be studied in functionalist terms;
- iii) involves, as a crucial functionalist feature to be studied, computational operations over representations.

This characterization aligns both with definitions offered by eAC proponents (e.g., Favela, 2024; Raja, 2024) and with traditional cognitivist works (e.g., Marr, 1982). Importantly, I present points ii) and iii) as scientific claims rather than metaphysical ones. While many cognitivists endorse both the metaphysical and scientific views (i.e., that cognition *is* computational *and* can be studied as such), others—such as Chirimuuta (2024) and Schöner et al. (2015)—accept them purely as methodological commitments. In other words, they provide

a (set of) framework(s) for the scientific study of cognition. That is the view adopted in this paper.

A few clarifications are in order. In (i), I add “primarily” because some people interpret extended functionalist approaches as compatible with cognitivism rather than with embodied frameworks (see Gallagher, 2018c; Rolla, 2021; Telakivi, 2023). In (ii), functionalism should be understood as a decompositional analysis that maps entities from one domain (e.g., input) to another (e.g., representation), typically via causal relations in concrete systems (Gallistel & King, 2009). This mapping is usually assumed to be shaped by evolutionary and developmental constraints (Neander, 2017; Piccinini, 2020). In (iii), computations refer to computable manipulations of symbols: given an input  $x$ , the system produces an output  $y$  (Gallistel & King, 2009).

The notion of representation requires further clarification. Many anti-representationalist arguments target specific types of representations, which often differ from how the term is used in contemporary cognitive science. As Meyer and Brancazio (2023, p. 5) point out, these critiques address “... but one of many examples of how representations have been used in the sciences...”. For instance, Van Gelder (1998) only argues against digital representations, i.e., representations whose vehicles are discrete entities. While it is true that traditional cognitivism theorized about cognition in terms of digital representations and computations, this is merely a contingent aspect of cognitivism. For example, Gallistel and King (2009, p. 24) argue in their classic textbook that we should study cognition based on digital representations. However, they acknowledge that this is because, at the time, there were much better mathematical models for digital representations and computations than for analog ones. They also recognize that this could change in the future, as it already has, at least partially (cf. Piccinini, 2015, 2020). Similarly, Chemero (2000, 2009) targets Millikan’s teleological notion of representation—a view influential in philosophy, but not mainstream in cognitive science (Shea, 2018).

Finally, Hutto and Myin (2013, 2017, 2018) reject a semantic notion of representation, specifically one that involves truth-conditions. Although such views were common before 2000, they have since declined in influence. For example, Hutto and Myin critique structural representations on the grounds that they lack truth-conditions (Hutto & Myin, 2018; Segundo-Ortin & Hutto, 2021). But structural representations are not intended to be truth-evaluable. Like maps, they *stand in for* their referents and guide action or reasoning toward them (Shagrir, 2012; Gładziejewski & Miłkowski, 2017; Piccinini, 2022). A map of Berlin is neither true nor false, but it can misrepresent Berlin by causing false beliefs or poor navigation if key features are

missing or inaccurate. In practice, truth-conditional representations have become largely irrelevant to cognitive science in recent decades (Ramsey, 2023). For example, Gallistel and King's (2009) widely used textbook defines representation entirely in functional terms, without reference to truth-conditions (cf. Ramsey, 2023).

In summary, the proponents of anti-cognitivism often reject specific notions of representation without engaging with how the concept is broadly used in contemporary cognitive science. To avoid this misunderstanding, and following the accounts of Vilarroya (2017), Kriegeskorte & Diedrichsen (2019), Baker et al. (2022), Cao (2022), and Piccinini (2020, 2022), I adopt the following definition of a neural representation: an identifiable pattern of neural activity or organization that

- a) reliably covaries with specific stimuli;
- b) causally influences the cognitive system's response to that stimulus;
- c) can be dissociated from the presence of the stimulus while still exerting a causal effect directed toward it.

Even though there is no single, unified use of the term representation in the cognitive sciences, I adopt this definition because it results from a survey-based analysis of the notion. Vilarroya (2017) surveyed how neuroscientists use the term and identified its most common and theoretically significant features—the features defined above. This survey has since been supported by further empirical investigation into scientists' usage (Baker et al., 2022) as well as by broader philosophical analyses of representational practices (Cao, 2022; Piccinini, 2020, 2022).

The crucial feature here is c), which grants the neural pattern the status of a *representation*—something that stands in for another thing. The cognitive system uses the representation to guide action or processing directed at the stimulus. This is most evident when the stimulus is absent but the representation remains active and continues to influence cognition or behavior.

My focus on neural representations reflects the example I present in the next section, which centers on a neural case. This focus ensures consistency within the paper. While not all cognitive science research explicitly relies on neural representations, there is a widely accepted assumption that cognitive representations are ultimately implemented in neural structures—even if the two are sometimes treated separately for methodological reasons (Marr, 1982).<sup>9</sup>

---

<sup>9</sup> One may also observe that arguments in favor of structural representations point to features (a–c) to justify their status as representations (Shagrir, 2012; Gładziejewski & Milkowski, 2017; Piccinini, 2022).

Having now clarified cognitivism's key methodological foundation, its core features, and its working concept of representation, I turn to an example that illustrates how cognitivism has been successfully applied in neuroscience.

### **3 An example of success of cognitivism**

A recent and prominent example of cognitivism's success comes from neuroscience, specifically, the study of memory traces or *engrams*. While this case does not emerge from computational cognitive science,<sup>10</sup> it clearly illustrates cognitivist principles being applied within a science focused on explaining behavior. In fact, I am not alone in identifying neuroscience as largely cognitivist in orientation as several eAC proponents acknowledge that neuroscience remains, for the most part, aligned with cognitivist assumptions (Raja, 2018, 2021, 2024; Favela, 2024).

The search for memory traces—physical correlates of memory—has a long history in the science of memory (see Kandel, 2007; Baddeley et al., 2020). However, recent technological advances in transgenics, optogenetics, and chemogenetics have significantly advanced this research by enabling scientists to monitor and manipulate specific neurons (Kunz et al., 2018; Josselyn & Tonegawa, 2020). Optogenetics, for example, involves introducing light-sensitive genes into targeted neurons using modified viruses. Once implanted, these neurons can be monitored and controlled—either activated or inhibited—by projecting light onto them (Deisseroth, 2011). Chemogenetics employs a similar strategy, using synthetic molecules instead of light to influence neural activity (Josselyn & Tonegawa, 2020).

In memory studies, researchers took advantage of the fact that neurons highly active during an experience express immediate early genes. These genes were used to identify and label the neuronal assemblies most active during specific events—these assemblies were putatively designated as engrams (Tonegawa et al., 2018; Choucry et al., 2024). Using optogenetic or chemogenetic tools, researchers could activate or inhibit these assemblies and observe their effects on behavior. Activation reliably produced behavior specific to the event originally associated with the neural activity, satisfying condition a) of neural representation: reliable covariation with stimuli (Josselyn & Frankland, 2018; Nambu et al., 2022). To illustrate, imagine four events: *e1*, receiving a foot shock in chamber 1; *e2* receiving a foot shock in chamber 2; *e3*, receiving a shock after hearing a specific tone (*f*); and, *e4*, receiving food after pressing a button (*b*) (for real studies, see Abdou et al, 2018, Sun et al, 2020; see also

---

<sup>10</sup> For examples of success in the computational cognitive sciences, see Quilty-Dunn et al (2023).

reviews in Josselyn & Frankland, 2018, Josselyn & Tonegawa, 2020). Each of these events would allow for tracking different neuronal assemblies. Artificially activating the assembly corresponding to *e1*, *e2*, or *e3* reliably produces fear behavior, regardless of the animal's current environment—demonstrating that the assembly continues to guide behavior in the absence of the original stimulus (satisfying condition *c*). Activation of the *e4* assembly produces reward-seeking behavior, similarly independent of external cues.

Even more compelling is how selectively inhibiting specific assemblies affects behavior, thereby satisfying condition *b*).<sup>11</sup> For instance, inhibiting the *e1* assembly eliminates the fear response to chamber 1, while leaving responses to chamber 2 (*e2*), the tone (*e3*), or the button press (*e4*) unaffected. This selective forgetting persists as long as the assembly is inhibited. The proliferation of such studies—across both optogenetic and chemogenetic methods—has led many researchers to conclude that engrams are both *necessary* and *sufficient* for memory-related behavior (Josselyn & Frankland, 2018; Tonegawa et al., 2018). For our purposes, this is also significant because these engrams fulfill the three conditions (*a–c*) outlined earlier, qualifying them as representations of the associated events.

Building on their ability to track engrams and treat them as representations of specific events, scientists began attempting to manipulate these engrams to expand their control over behavior—aligning with tenet (iii) of cognitivism. In a now widely cited study, researchers succeeded in creating false memories in mice (Ramirez et al., 2013; Robins, 2016). In the experiment, mice were first exposed to chamber 1, where they received foot shocks, thereby forming engrams for fear conditioning. Later, they were placed in a neutral chamber 2. Scientists then simultaneously activated the engrams corresponding to both chambers using optogenetics. As a result, the mice exhibited fear responses in chamber 2, despite never having experienced fear conditioning there. One might notice that this manipulation led to a *misrepresentation*, not in the semantic sense criticized by Hutto and Myin (2013, 2018), but in the functional sense: the mice's fear behavior was directed toward a neutral stimulus, making the response maladaptive. Such behavior is evolutionarily and developmentally inappropriate, as fear of a safe environment could interfere with vital activities like foraging or mating (cf. Piccinini, 2020).

Even more surprising are recent experiments of memory trace implantation (Najenson, forthcoming). Vetere and colleagues (2019) used optogenetics to simultaneously activate

---

<sup>11</sup> Here, attributing causation to satisfy conditions *b* and *c* is based on an interventionist approach to causation (Woodward, 2003). It is worth noting that monitoring and controlling (i.e., inhibiting and exciting particular assemblies) is a means of performing ideal interventions.

sensory neurons that were sensitive to the odor of a specific substance (acetophenone) and neurons that were specific to fear responses (dopamine neurons in the medial ventral tegmental nucleus) in mice that have never felt the smell of that substance before. The result was these mice showed fear responses to the smell of acetophenone even though they never felt that odor before. Similar, but positive results, were found when these sensory neurons were simultaneously activated with neurons associated with reward, which led the mice to show interest responses to the odor when introduced to it for the first time. Other similar results were found using different sensory neurons, from visual, auditory and somatosensory cortices (Luis-Islas et al, 2022).

The present studies demonstrate that these neural assemblies covary with specific stimuli (condition a). These assemblies have causal effects specific to those stimuli. The activation of neurons responsible for certain types of responses results in an association of responses to those stimuli. Conversely, the inhibition of these neurons suppresses the learned behavior only toward the stimulus (condition b). Given that these mice have never experienced that smell before, the causal-effect relation toward the stimulus can be said to be dissociable from the presence of the stimulus (condition c).

Beyond fulfilling the criteria for representation, this research program also aligns with the Galilean principle. Scientists acknowledge that memory behavior arises from a complex interaction of bodily and environmental factors (Iani, 2019, Righetti, 2024). Moreover, engrams depend on prior contact with the world—either through developmental experience (as in the earlier experiments) or likely through inherited structures shaped by evolution (as in implantation studies). Moreover, engrams alone do not suffice for behavior: a functioning nervous system and body are also necessary. Nevertheless, researchers describe engrams as necessary and sufficient for memory behavior (e.g., Josselyn & Frankland, 2018). This reflects the use of *ceteris paribus* generalizations—other things being equal, engram activation is necessary and sufficient for memory behavior—typical of Galilean explanatory strategies (Pietroski & Rey, 1995). In sum, this research exemplifies cognitivism. It adopts the Galilean methodology, focuses on explanatory factors within the central nervous system, employs functionalist reasoning (via causal relations), and manipulates representations to explain and control behavior.

The success of cognitivism in the case of engrams is evident in the explanatory depth and experimental productivity it has enabled.<sup>12</sup> Engrams, as explanatory factors for memory behavior, have allowed researchers to manipulate a wide range of memory-related phenomena. One particularly notable example comes from studies on mice models of Alzheimer's disease. Researchers found that engrams formed during the early stages of the disease were not destroyed but merely silent—that is, they could not be activated by environmental cues, but remained intact. Using optogenetic and chemogenetic techniques, scientists were able to artificially activate these engrams, thereby prolonging memory retention compared to other silent engrams (Josselyn & Tonegawa, 2020). This finding opened new avenues for potential interventions in memory disorders like Alzheimer's, even if only at an experimental stage. Thus, while the engram-based framework abstracts away from many real-world causal factors and is not designed for direct application "in the wild," it has proven remarkably effective in targeting internal, brain-based mechanisms.

Cognitivism's success here is not limited to explanatory power as it has also demonstrated significant fruitfulness. If researchers had not adopted the cognitivist assumption that neural assemblies can represent events, it is unlikely they would have pursued the identification of engrams in the first place. Historically, many researchers were skeptical of this approach and avoided investigating the neural basis of memory representations (e.g., Schacter & Addis, 2007). In fact, earlier research suggesting that there were no physical representational vehicles in the brain for mnemonic information discouraged work on engrams for decades (cf. Kandel, 2007; Baddeley et al., 2020). Moreover, the bold idea that such representations could be manipulated to control behavior with higher precision fostered studies on false memory creation and memory implantation. Without these cognitivist commitments—specifically, the view that internal representations can be isolated and causally manipulated—these discoveries would likely not have occurred. As a matter of comparison, enactive/embodied-inspired approaches to neuroscience (e.g., Raja, 2018, 2021, 2024; Raja & Anderson, 2019; Favela, 2024) do not provide a clear path to such experimental strategies. The focus on real-world interactions, coupled with a general skepticism toward internal representations, makes it

---

<sup>12</sup> Hutto (2023) argues that memory traces are dispensable for explaining memory behavior. However, his critique focuses primarily on traditional causal theories of memory and avoids engaging with the recent body of research on engrams. As a result, he does not contend that this research is mistaken or unsuccessful. In contrast, a growing number of scholars maintain that engram research has made memory traces indispensable for understanding remembering (De Brigard, 2020, forthcoming; Najenson, 2021) and strongly supports the claim that such traces are representational, as discussed in the previous section (Robins, 2023, forthcoming; Andonovski, 2024).

difficult to generate hypotheses or interventions like those described above. As such, eAC frameworks would be unlikely to motivate the same kinds of experimental designs.

To conclude this section, the case of engrams demonstrates both the explanatory success and the fruitfulness of cognitivism. It illustrates how a theory grounded in internalist and representational assumptions, combined with a Galilean methodology, can lead not only to understanding but also to innovative methods of intervention. In the next section, I turn to two different ways in which eAC proponents attempt to reject cognitivism: pluralist and eliminativist approaches.

#### **4 Two radical alternatives to cognitivism**

As stated in the introduction, eAC represents a radical departure from cognitivism as a framework for the cognitive sciences. In practice, this departure has taken two distinct forms. The first is what I call pluralism; the second, eliminativism. Although many eAC proponents appear to lean toward the eliminativist view, I begin by clarifying the pluralist position.

Before proceeding, I will clarify what defines *eAC* in light of the characterization of cognitivism given in Section 2, just as I did there for cognitivism itself. As noted in the introduction, there are many forms of anti-cognitivism. A well-known example is behaviorism. Although less influential today than before the rise of cognitivism, it remains active and rejects cognitivist features i)–iii) (Lazzeri, 2016). For this reason, it can be considered an anti-cognitivist framework. However, this paper does not address that form of anti-cognitivism. My analysis focuses instead on anti-cognitivist frameworks grounded in enactive and embodied research programs, which is why I refer to the target as *eAC* rather than simply *AC*.

Importantly, my target is not enactivism as a whole family of frameworks. Several versions of enactivism and embodied cognition do not reject cognitivist core features i)–iii), thus they cannot be called anti-cognitivist. For example, sensorimotor enactivism (Noë, 2004) does not reject features ii) and iii) (cf. Hutto, 2018), and weak embodied cognition (Goldman, 2012) does not reject features i)–iii) at all (cf. Gallagher, 2018a). By contrast, my analysis focuses on those enactive/embodied frameworks that explicitly reject—and propose models and explanations that do not rely on—cognitivist features i)–iii). This rejection of all features i)–iii) is what I am calling anti-cognitivism.

Despite this relatively minimal definition of anti-cognitivism, I continue to use the label anti-cognitivist because proponents of these views consciously and deliberately avoid features i)–iii). Typically, eAC frameworks are described as “radical,” including Radical Embodied

Cognition and Radical Enactivism. However, Autopoietic Enactivism can also be included within this category, since many have argued that it, too, avoids i)–iii) (cf. Rolla & Hufferman, 2022), and thus qualifies as anti-cognitivist in the sense adopted here.

Philosophers on both sides of the debate, cognitivist and eAC, have recognized that, even if empirical evidence supported both approaches, a disagreement would persist due to their conflicting conceptualizations of cognition. In this sense, the dispute between cognitivism and eAC potentially becomes a conceptual or verbal one (Aizawa, 2015; Sanches de Oliveira & Chemero, 2015). Cognitivists define cognition<sub>1</sub> as a set of internal processes involving symbolic manipulation localized in the brain. eAC proponents, by contrast, define cognition<sub>2</sub> as something that emerges from the dynamic interaction between organism and environment. This contrast is illustrated in Figure 1. Thus, in principle, both frameworks could succeed within their respective domains of inquiry: cognitivism in exploring and theorizing about cognition<sub>1</sub> and eAC in exploring and theorizing about cognition<sub>2</sub>.

Aizawa (2015, 2017) sought to resolve this potential verbal disagreement by claiming that cognitivists have historical precedence over the concept of cognition and that, accordingly, eAC theorists should adopt a different term for what they study. From his perspective—one that some eAC proponents accept (Sanches de Oliveira & Chemero, 2015; Favela & Chemero, 2015)—cognition<sub>2</sub> amounts to what has traditionally been called behavior, the interaction between organism and environment. Thus, cognitivists would continue to study cognition, while eAC theorists would study behavior. However, this proposal neglects the historical record: as Sanches de Oliveira and Chemero (2015) and Crippen (2025) argue, the term cognition has been used in both senses since at least the late nineteenth century. Moreover, although eAC's conception of cognition is closely aligned with behavior, it specifically targets the interactivist regularities behind behavior, not behavior per se. Hence, the conceptual divergence between the two views supports a form of framework pluralism: using eAC to study interactivist explanatory factors and cognitivism to study internalist ones.

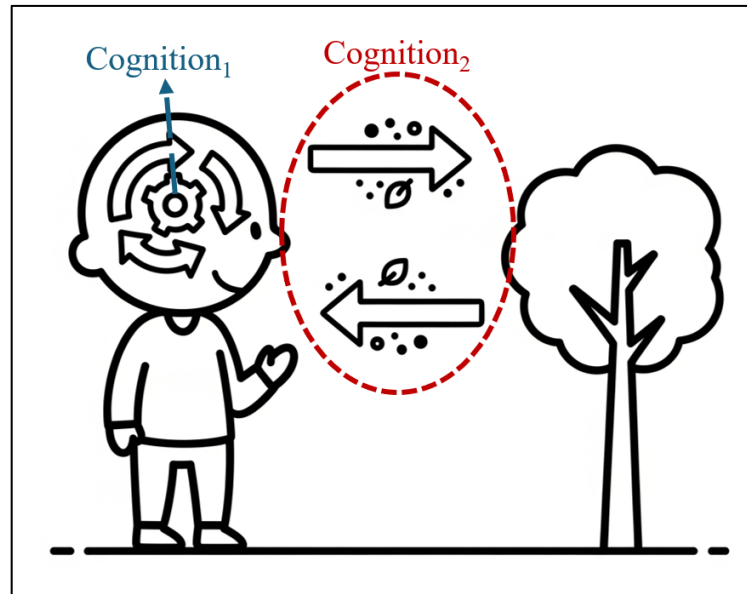


Fig. 1 – A representation of two concepts of cognition. Cognition<sub>1</sub> (cognitivist) selects an internalist set of factors that explain behavior whereas cognition<sub>2</sub> (REC) selects a set of interactive factors between organism and environment to explain behavior. Image generated by Gemini AI.

Indeed, some eAC theorists have explicitly endorsed such pluralism. A clear example is Potochnik and Sanches de Oliveira (2020), who argue that both cognitivist and eAC approaches identify distinct, yet valid, causal patterns underlying behavior. Even when eAC theorists investigate the brain, their interest lies in how brain activity is shaped by organism–environment interactions and on its internal dynamical features (Raja, 2024). In this view, both frameworks can be legitimate because they track different types of explanatory factors. eAC, then, is an alternative to cognitivism not because cognitivism is flawed *per se*, but because eAC is concerned with a different explanatory domain. A useful analogy comes from physics, where different models are used to study the atom depending on the particular features one aims to explain (Massimi, 2022, ch. 4). Other examples of eAC pluralists include Kirchhoff (2025b, ch. 8), who advocates for a pluralist approach, and Raja (2024). Raja states that his defense of radical embodied cognitive neuroscience does not require abandoning cognitivism, but rather acknowledges it as one of several successful and promising frameworks for studying cognition, which is a pluralistic attitude.

Eliminativist eAC, by contrast, holds that cognitivism is, for some reason, fundamentally flawed—or at least inferior to eAC—and should therefore be abandoned in favor of it. Several influential figures within eAC have expressed this non-pluralist stance. For

example, van Gelder (1995, 1998) presents cognitivism and eAC as competing approaches and argues that eAC is the more promising framework, implying that it will ultimately replace cognitivism. In a similar vein, Chemero (2000, p. 626) advances an anti-representationalist argument whose implication is that “... cognitive science must be re-built from the ground up, re-contextualized without reference to representation or computation,” an implication he attributes to van Gelder’s earlier work. This eliminativist tendency is also present in Hutto (2018), who characterizes his version of eAC as a conceptual revolution in the cognitive sciences—one that would reframe the entire theoretical and methodological landscape of the field, much like the Copernican revolution transformed astronomy. As he puts it: “REC aims not merely to adjust certain aspects of the classical cognitivism vision, but to supplant that outlook entirely” (Hutto, 2018, p. 389).

This eliminativist orientation, however, requires some nuance. Chemero, for instance, can be interpreted as adopting a more pluralist stance in later work (e.g., Chemero & Silberstein, 2008; Silberstein & Chemero, 2013). Hutto and Myin (2013, 2017) may also seem more cautious, restricting their criticism to the claim that cognitivism is mistaken about *basic* cognition. Yet this more modest qualification often serves as the foundation for later, stronger claims—such as the passage from Hutto (2018, p. 389) cited above and the further assertions in Hutto and Myin (2018, pp. 104–105):

REC has all the hallmarks of being bona fide revolutionary given that it presses for “the replacement of a whole system of concepts and rules by a new system” (Thagard, 1992, p. 6). Like other major revolutions in thought, REC does not force us to jump from one conceptual branch to another while staying within the same familiar tree—it demands that we switch to a new tree altogether... Going radical the REC way is to abandon the information-processing and representationalist views of cognition in favor of a purely embodied know-how.

Although Hutto and Myin’s (2017) version of eAC, as well as many others (e.g., Di Paolo et al., 2018; Rolla, 2018), holds that high-level cognition involves representations, these representations are not viewed as mental—and certainly not as neural—representations (cf. Rolla, 2021). Instead, high cognition is understood as enacted within a social domain, giving rise to *social* rather than *mental* representations (Hutto & Satne, 2015). This view differs significantly from cognitivism as presented in Sections 2 and 3. A clear example of this

difference appears in Hutto and Myin's (2017) account of memory. They argue that "pure episodic memory" is not representational and becomes representational only through social practices (see also Myin & Dijk, 2022). Thus, even though they acknowledge the existence of representations, they continue to reject cognitivist explanations across all forms of cognition, not only basic cognition. For this reason, they can be accurately described as eliminativists.

Given that cognitivism and eAC employ different concepts of cognition—as noted at the beginning of this section, a point that lends support to pluralism—eliminativists must reject the cognitivist concept of cognition, that is, cognition<sub>1</sub> (Hutto & Myin, 2018, pp. 105–106). Drawing on the tools of conceptual engineering, one can identify two broad categories of reasons for rejecting a concept: ethical–political and epistemic–metaphysical (Dutilh Novaes, 2020a). Some critics have argued that cognitivist frameworks in neuroscience support social injustice and should therefore be reformed or eliminated (e.g., Einstein, 2012). While this type of argument is important, it falls outside the scope of this paper, which focuses on epistemic and metaphysical critiques. Nonetheless, it is worth noting that eAC-based concepts have also been linked to problematic social consequences (Russell, 2023). A pluralist approach, as Brancazio (2018) suggests, may mitigate these issues without requiring the outright rejection of either framework.

Setting aside ethical and political considerations, one may also reject a concept on the grounds that it fails to refer—that is, that nothing in the world corresponds to it. This is a metaphysical argument. On this view, cognitivism would be fundamentally mistaken because it theorizes about something that does not exist, much like past theories about the ether or phlogiston. A different group of arguments is epistemic. These claim either that the concept lacks genuine explanatory value (or can be replaced by a better one), or that the methods used to investigate the regularities associated with the concept are flawed. According to these arguments, the posits of cognitivism are either insufficient for explanation or rely on internal assumptions and methods that undermine the study of the very regularities they aim to capture.

In the remainder of this paper, I will assess the main epistemic and metaphysical arguments for eliminating cognitivism. These include the claims that cognitivism is not (or cannot be) explanatory (Section 5), that its key concepts are conceptually problematic (Section 6), and that the internal factors it identifies as explanatory may not exist (Section 7). I will argue that none of these arguments succeed, especially in light of the success of cognitivism discussed in Section 3. As a result, I conclude that eAC proponents should adopt a pluralist rather than an eliminativist stance, at least as things stand thus far (Section 8).

Before proceeding, I want to make one final clarification. Although I do not evaluate the success of eAC in this paper due to space constraints, I do take eAC to be a successful and fruitful family of frameworks. Common criticisms of eAC’s explanatory power—such as those by Zednik (2011) and Vernazzani (2019), which dismiss it as merely descriptive—often beg the question by assuming a mechanistic view of explanation. As others have argued, including some not aligned with eAC, biological sciences can accommodate interactivist and mathematical explanatory models that are distinct from mechanistic ones (e.g., Chirimuuta, 2018; Ross, 2024). For this reason, and in line with arguments from eAC proponents themselves, I take it for granted that eAC can be both explanatory (e.g., Silberstein & Chemero, 2013) and fruitful (see Chemero, 2009, ch. 5). Accordingly, just as I argue that eAC should not eliminate cognitivism, I also maintain that cognitivists should not be eliminativists about eAC.<sup>13</sup>

## 5 Assessing explanatory reasons to reject cognitivism

As shown in Section 3, cognitivism succeeds in providing compelling explanations for cognitive phenomena such as memory. Nonetheless, eAC eliminativists contend that cognitivism suffers from explanatory shortcomings that warrant its elimination as a scientific framework, along with its associated concept of cognition. That is, the concept *cognition*<sub>1</sub> cannot track explanatory factors and, therefore, should be eliminated. The main criticisms fall into three broad categories: that cognitivist explanations are incomplete (e.g., Gallagher, 2018b), dispensable (e.g., Chemero, 2000; Rolla, 2021), or shallower/worse than eAC explanations (e.g., Favela, 2024). These are versions of the epistemic argument, according to which the concept of *cognition*<sub>1</sub>—or the regularities it is meant to track—suffers from explanatory shortcomings and should therefore be replaced by a concept that avoids those shortcomings, *cognition*<sub>2</sub>. These three broad categories correspond to distinct types of alleged shortcomings identified by proponents of eAC. In this section, I examine these arguments and show that they rest on a misunderstanding of what constitutes scientific explanation or result from a superficial grasp of cognitivism.

The first criticism is that cognitivist explanations are incomplete. Gallagher (2018b, p. 9, emphasis added), for example, critiques cognitivist accounts of social cognition for being limited to “... the idea that we can find a *complete explanation* of how we understand others in

---

<sup>13</sup> In general, cognitivists who adopt an eliminativist stance adhere to the notion of *cognition*<sub>1</sub>. However, it is unlikely that any cognitivist would refute the notion that environmental and bodily factors serve as explanatory factors for behavior, and therefore, merit investigation. In contrast, eAC eliminativists tend to consider that cognitivist factors, such as representation and computation, are not explanatory for behavior. This topic will be discussed in the next section.

a set of mechanisms that are entirely contained in the individual brain...” He argues that these accounts fail because they ignore environmental and bodily factors. Similarly, Rolla (2021, p. 30, emphasis added, my translation), after describing a sequence of mundane behaviors, asks: “Would it be possible to capture the nature of the mental as a whole within a unique theoretical perspective?” He concludes that cognitivism fails to offer such a perspective. Before doing so, he introduces cognitivism in the following way: “In the last three decades of the past century, the cognitive sciences *promised* an elegant explanation of the mind *in its integrity*.”

As argued in Section 2, however, such criticisms reveal a fundamental misunderstanding of the aims of cognitivism. By being a Galilean science, cognitivism aims not at exhaustive explanation of behavior (or the mind) but at isolating and theorizing about a specific set of internalist explanatory factors. Cognitivism, in other words, never promised a complete or integrative theory of the mind. A well-known illustration of this point comes from the exchange between Pinker and Fodor: after Pinker’s *How the Mind Works* was published, Fodor responded with *The Mind Doesn’t Work That Way*. In the very first page of that book, Fodor (2001, p. 1) observes that although he supports the computational theory of mind, he never believed it could constitute “more than a fragment of a full and satisfactory cognitive psychology; and that the most interesting—certainly the hardest—problems about thinking are unlikely to be much illuminated by any kind of computational theory we are now able to imagine. I guess I sort of took it for granted that even us ardent admirers of computational psychology were more or less agreed on that.” Hence, criticizing cognitivism for failing to deliver a complete explanation is a case of holding it accountable for aims it never professed. In other words, it is the result of a misunderstanding of cognitivism.

Although cognitivist explanations are not complete, Section 3 demonstrated that they are deep—i.e., they support robust, stable and multiple forms of intervention (Hitchcock & Woodward, 2003, Woodward, 2010). eAC proponents may insist that completeness is a necessary criterion for scientific adequacy (e.g., Rolla, 2021; Favela, 2024). But I remain skeptical of the very idea of a “complete” explanation of any complex phenomenon, as I will discuss after introducing the second criticism.

The second line of attack is that cognitivist explanations are dispensable or unnecessary. This claim is typically supported by pointing to models that successfully explain behavior without invoking internal representations. One example comes from dynamical systems theory, which allegedly explains behavior without recourse to representations, thereby rendering them superfluous (e.g., van Gelder, 1998; Chemero, 2000; Sanches de Oliveira et al., 2021). Another

example is the development of robotics systems that mimic human behavior without relying on internal representations (e.g., Chemero, 2009; Hutto & Myin, 2013; Rolla, 2021).

These criticisms—like the first—rest on a misunderstanding of the nature of scientific explanation. Scientific explanations aim to identify stable dependence relations that account for phenomena of interest (Woodward & Hitchcock, 2003; Potochnik, 2017; Ross, 2024). But it is widely recognized that no scientific explanation can capture all such relations: explanations are thus partial, selective, and aspectual (Ylikoski & Kuorikoski, 2010; Ross, 2024). Even dynamical systems models rely on selecting a subset of relevant factors from incomplete datasets (van Gelder, 1998; Leonelli, 2016). As such, no model can claim to be complete, and, hence, no model—dynamical or otherwise—can render another wholly dispensable because it omits certain factors that are crucial to that other model.

Moreover, the claim that representations are unnecessary simply because some models omit them is misguided. As Section 3 showed, representations are explanatory in other models, particularly in those targeting internalist factors. As a matter of analogy, consider the case of early radical embodied cognitive science, a particular version of eAC, and its proposal of dynamical systems models. In its early stages, this framework notably omitted brain dynamics from their explanations (e.g., van Gelder, 1998; Chemero, 2000, 2009). Despite this omission, radical embodied cognition theorists today acknowledge the explanatory importance of neural dynamics and actively work to integrate them into their models (e.g., Raja, 2021, 2024; Favela, 2024). None of these theorists would conclude that brain dynamics are explanatorily dispensable merely because they were absent from initial dynamical models.

Yet, when it comes to representations, the same theorists argue that their absence from dynamical models does show that they are unnecessary or dispensable (e.g., Chemero, 2000). This asymmetry is hard to justify. The same point applies to the robotics literature: the fact that some robots function without representations does not imply that all forms of behavior can be explained without them. One must first clarify the notion of representation at stake, since some have argued that these robots do count as representational in the relevant sense (e.g., Bechtel, 2008). Moreover, as Vernazzani (2019) has shown, dynamical systems models and representational models can be complementary rather than mutually exclusive when explaining perception. Hence, the explanatory success of one eAC model does not entail the outright dispensability of cognitivist models.

In sum, the fact that a given phenomenon can be explained without invoking representations does not mean that representations are dispensable in all explanatory contexts.

Explanations are selective by nature. The success of dynamical models is no reason to abandon other explanatory models that appeal to internal representations—particularly when, as demonstrated in Section 3, those models have yielded fruitful and deep scientific insights.

A related objection arises from another line of criticism offered again by radical embodied cognition proponents. It is sometimes argued that, since dynamical models do not support decomposition and localization, cognitive systems—presumed to be dynamical—are likewise resistant to decomposition and localization (van Gelder, 1998; Chemero & Silberstein, 2008). This argument, however, conflates models with the phenomena they aim to describe—it mistakes the map for the territory, as eAC theorists themselves are fond of saying (cf. Ciaunica, 2022, Kirchhoff, 2025b). Like scientific explanations, explanatory models represent only *aspects* of a phenomenon (Bokulich, 2011; Hochstein, 2016; Massimi, 2022). Therefore, concluding that a phenomenon cannot be decomposed or localized simply because a dynamical model of it resists such treatment is a *non sequitur*.

As Mitchell (2002) argues, scientific models rely on idealizations, and each model idealizes its subject differently. Consequently, even models that yield seemingly conflicting conclusions can be integrated. This is because their conclusions conflict about their specific *idealized objects*, not the actual phenomenon they aim to explain. True competition arises only when models track the same set of causal factors (cf. Potochnik & Sanches de Oliveira, 2020). As established in Section 4, cognitivism and eAC analyze distinct causal factors. They can therefore be integrated.

In this sense, dynamical systems models emphasize interactive features of behavior, which may resist decomposition and localization. By contrast, more stable internal features, such as engrams, are better suited to a decompositional approach. In consequence, they would end up not being competing hypotheses. For example, one may construct dynamical models of memory performance (Grieben et al., 2020), but this does not preclude the investigation of memory through decomposition and localization methods, as demonstrated in the engram research. Indeed, Silberstein and Chemero (2013) eventually acknowledge this very point. A model's explanatory limits should not be mistaken for ontological boundaries on the phenomena it represents.<sup>14</sup>

---

<sup>14</sup> Another related proposal that resists the search for stable decomposition patterns is the Protean Brain Hypothesis, which holds that brain activity is so dynamic and ever-changing that it can only be understood within specific contexts and, therefore, cannot support the kind of generalizations endorsed by the Galilean principle central to cognitivism (Hutto et al., 2017). Although this proposal does not rely on a misunderstanding of scientific explanation, it fails to acknowledge that the dynamic nature of brain activity does not preclude the identification of stable and generalizable patterns across contexts (cf. Chirimuuta, 2024, ch. 7). Research on engrams—

The final criticism offered by eAC proponents is that eAC allegedly provides *better* explanations than cognitivism (e.g., Hutto & Myin, 2018; Gallagher, 2018). In this case, both radical embodied cognition and radical enactivism researchers appeal to this argument. There are, of course, widely accepted objective criteria for evaluating explanatory adequacy—such as the stability and invariance of the dependence relations described, as well as the specificity of the explanation (Hitchcock & Woodward, 2003; Woodward, 2010, 2016; Ross, 2015, 2024, forthcoming). eAC theorists may therefore claim that, by these measures, their explanations outperform those offered by cognitivism.

However, such evaluative criteria are always constrained by the pragmatic context of inquiry—specifically, by the aspects of the phenomenon that researchers aim to control or understand (Ross, 2015, 2024, forthcoming; Woodward, 2016). For this reason, there is no single “complete” or “best” explanation across all domains—only those that are best suited to specific explanatory or practical purposes.

Consider, for instance, the retrieval of a specific memory during the early stages of Alzheimer’s disease. As shown in Section 3, environmental and bodily factors perform poorly compared to artificial reactivation of silent engrams using optogenetic methods (Josselyn & Tonegawa, 2020). For this specific purpose, engram-based explanations are more stable and specific than eAC explanations of memory retrieval (compare with Ianì, 2019). This does not imply, however, that we should discard embodied or environmental models outright (Righetti, 2021, 2024). For other purposes—such as understanding the political implications of memory—engram research may play no role, and eAC-based accounts are far more relevant (cf. Campbell, 2014).

This purpose-sensitive view of explanation aligns better with the epistemological commitments of at least some of eAC frameworks—particularly those that subscribe to ecological psychology principles—than does the search for a context-independent “best” explanation (cf. Sanches de Oliveira, 2022, 2023; Sanches de Oliveira et al., 2023, Malavolta et al., 2025). If human cognition—including scientific cognition (Rolla & Novaes, 2022)—is situated within ecological-enactive contexts, then the models and explanations it yields should also be understood as context-bound. Accordingly, explanatory evaluations—such as judgments

---

particularly studies on memory trace implantation—demonstrates that the brain possesses stable properties that allow for reliable manipulation of memory-related behavior, independently of specific environmental or embodied contexts. These findings challenge the idea that cognitive neuroscience cannot generalize and suggest that at least some brain mechanisms conform to the explanatory ideals associated with cognitivism.

of which explanation is “best” or “most appropriate”—must be situated within specific research practices and purposes, rather than framed in absolute or decontextualized terms.

In summary, eAC’s objections to the explanatory merits of cognitivism rest on an overly narrow or unexamined conception of scientific explanation. Indeed, if one accepts eAC as a broader philosophy of nature (Gallagher, 2017, Meyer & Brancazio, 2022), then one is committed to a situated epistemology—favoring eAC pluralism over eliminativism—unless further arguments are provided to justify eliminating cognitivism. In the following sections, I address additional reasons offered by eAC proponents in support of such elimination.

## **6 Assessing conceptual reasons to reject cognitivism**

As demonstrated in Section 4, the dispute between cognitivism and eAC is arguably headed for a conceptual disagreement. To resolve this, some eAC proponents have presented conceptual justifications for the elimination of cognitivism. This is another type of epistemic argument. In this case, the target is not cognition<sub>1</sub> itself, but the cognitivist framework’s ability to investigate such regularities (or any regularities at all). The claim is that cognitivism is internally flawed because of conceptual inconsistencies. Thus, even though this argument does not reject cognition<sub>1</sub>, it does reject cognitivism as a viable framework for studying it.

Perhaps the clearest case is Radical Enactivism, as advanced by Hutto and Myin (2013, 2017, 2018) and further developed by Hutto (2018, 2023). This framework aims to revolutionize the cognitive sciences by refining their conceptual foundations, eliminating what it sees as problematic notions, most notably, representation and content. According to its proponents, these notions are not aligned with scientific practice and should therefore be excluded from the conceptual toolkit of cognitive science. Hutto and Myin even appeal to Thagard’s (1992) analysis of scientific revolutions to support their view. However, Thagard’s analysis shows that conceptual revolutions occur when a prior conceptual framework exhibits serious explanatory failures. As argued in the previous section, this is not the case for cognitivism, at least not thus far.

Another example comes from Favela and Machery’s (2023) empirical investigation of the concept of representation in cognitive science. They found significant disagreement among scientists on basic questions—for example, when a neural pattern should be said to *represent* a stimulus, or at what level representational content is instantiated (see also Favela & Machery, 2025). Similar concerns have been raised by Raja (2018, 2024) about the concept of *stimulus*,

whose use ranges from proximal retinal inputs to entire events, such as those described in Section 3.

In their initial article, Favela and Machery (2023) concluded that the lack of conceptual clarity surrounding representation left us with two options: reform or eliminate the concept. Radical Enactivism, as noted, opts for the latter (Hutto, 2018). However, in a subsequent paper, Favela and Machery (2025) revised their position, recognizing that imprecise concepts often play productive roles in science. Accordingly, they propose investigating how imprecise concepts contribute to scientific practice rather than calling for their immediate reformation or elimination. This is not a new idea in the philosophy of science.

One key way that imprecise concepts such as *representation* can be scientifically fruitful is through enabling analogical reasoning (Chirimuuta, 2024). As noted throughout this paper, I follow Carnap's (1950) notion of fruitfulness, according to which scientific concepts should facilitate further research by encouraging new hypotheses and experiments (see also Dutilh Novaes & Reck, 2017; Dutilh Novaes, 2020a, 2020b). As seen in Section 3, the concept of representation was crucially fruitful in the engram research program, despite its conceptual imprecision. Chirimuuta (2024) explains that this fruitfulness stems from the analogy between representation and intentional communication. This analogy guided researchers to search for a vehicle of a sign, which motivated the hunt for neural correlates of memory, such as engrams (Andonovski, 2024). Furthermore, the analogy between cognition and computation fostered the exploration of the manipulation of these vehicles, i.e., the engrams to increase the possible ways to control memory behavior.

Additionally, imprecise concepts enable what Feest (2017) and Colaço (2022) call *kinding processes*: the temporary grouping of phenomena under a common label, allowing for fruitful comparison and differentiation. This is exemplified by the concept of *memory*, which is itself highly contested (Feest, 2025), yet continues to generate productive research by lumping together diverse phenomena for empirical comparison. Thus, through both analogical use and kinding, imprecise and contested concepts can be scientifically valuable—and for that reason, should neither be eliminated nor immediately reformed. As shown in Section 3, at least the analogical function of the concept of representation has proven fruitful in the context of cognitivism.<sup>15</sup>

---

<sup>15</sup> Another example of a concept that is imprecise and disputed but that is far from being eliminated from the sciences is the very concept of cognition (Allen, 2017).

In response to these points, Radical Enactivism offers a more forceful conceptual critique: the Hard Problem of Content (HPC). The core of the argument is as follows:

P1. The only respectable notion of information in the sciences is information-as-covariance.

P2. Content (and therefore representation) is a notion of information that cannot be reduced to information-as-covariance.

C1. Therefore, content is not a respectable notion of information in the sciences.

P3. Any non-respectable notion of information should be eliminated from scientific practice.

C2. Therefore, content should be eliminated from scientific practice.

As discussed in Section 2, this argument targets *strongly semantic* notions of representation and content, which are not representative of all uses in the cognitive sciences. In fact, other, less semantic notions of representation—such as the one employed in this paper—have been defended in the literature (cf. Milkowski, 2015; Shea, 2018; Ramsey, 2023). Due to space constraints, I will not delve into those responses here. However, it is important to note that HPC is an instance of what Chemero (2009, p. 7) calls a *Hegelian argument*. Hegelian arguments consist of arguing “... that a class of scientific approaches is doomed to fail based on theoretical posits and little or no empirical evidence.” HPC fits this pattern, as it predicts the failure of cognitivism purely on the grounds of its reliance on a supposedly flawed concept, without any empirical evidence of that failure. As previously discussed in this section and in Section 3, cognitivism possesses a remarkably fruitful and successful conceptual framework, despite its imprecision.

Moreover, as Hutto and Myin (2013, 2017, 2018) recognize frequently, other notions of information may turn out to be respectable, but they are skeptical about it. Their expectations of which concepts will turn out to be respectable, however, is not an empirical evidence. Furthermore, there is not much evidence for P1, as the use of information in the sciences is very plural and Shannon’s notion is very respectable in the sciences (Floridi, 2010). Even some eAC advocates adopt such notion of information (De Carvalho & Rolla, 2020), which is one of the basis for the notion of content and representation in the cognitive sciences (Gallistel & King, 2009; Martínez, 2019). Hence, HPC is based on a premise (P1) that lacks empirical evidence, thus constituting a Hegelian argument. As Chemero (2009) demonstrates well, Hegelian arguments have no space in scientific practice, which is evident by the fact that HPC has played little role in any conceptual framework adopted in the cognitive sciences.<sup>16</sup> Even non-

---

<sup>16</sup> For instance, a search for "Hard Problem of Content" in *PubMed* returns no scientific article referring to that argument, as of July 2025.

cognitivist frameworks do not base their account in HPC, but in other, much more solid grounds, such as the dynamical systems theory.

Therefore, a significant part of eAC criticism toward the conceptual framework used by cognitivists fail because either it does not recognize the importance and fruitfulness of imprecise/contended concepts in science or because it appeals to Hegelian arguments that have no place in scientific practice. For this reason, as of now cognitivist conceptual frameworks should not be eliminated from the cognitive sciences.

### **7 Assessing metaphysical reasons to reject cognitivism**

This section addresses a metaphysical concern that has been used against cognitivism. The reasoning proceed as follows: even if one accepts that cognitivism is explanatorily powerful and its concepts scientifically fruitful, one could nonetheless reject it as a metaphysically accurate depiction of the world. That is, by adopting a form of antirealism about scientific explanation (cf. Chirimuuta, 2024), one might concede that cognitivism generates useful models and experimental strategies without committing to the reality of the entities or processes it posits. This reasoning has been advocated by Kirchhoff (2025b). He argues that cognitivist explanations, particularly their appeal to representations, are useful fictions or metaphors given their high degree of idealization, and that such models carry no genuine metaphysical import.

However, the mere fact that a framework posits non-real or fictional entities is not in itself a reason to reject it—at least not from the standpoint of scientific practice, as he recognizes. Many scientific models are explicitly fictional or idealized (e.g., toy models) and yet play indispensable roles in theory development and experimentation. Nevertheless, one might respond that eAC is not only a philosophy of science but also a metaphysical thesis about the nature of mind and cognition (cf. Gallagher, 2017, Meyer & Brancazio, 2022, 2023). In that case, even if cognitivism is valuable within the epistemological practices of science, it might still be rejected on metaphysical grounds for failing to reflect the true nature of cognition. According to this metaphysical argument (Section 4), cognition<sub>1</sub> does not correspond to anything in the real world and should therefore be eliminated.

In response, a major point should be made as this argument faces serious difficulties. As seen in Sections 2 and 3, the explanatory concepts central to cognitivism—such as representation—are individuated by their causal profiles and employed in causal explanations. Many eAC theorists also appeal to causal explanations (cf. Potochnik & Sanches de Oliveira, 2020; Raja & Sanches de Oliveira, 2024). To reject cognitivist explanations on metaphysical

grounds despite their causal-explanatory utility risks slipping into a form of global antirealism that would undermine all causal explanations in cognitive science, including those preferred by eAC proponents (e.g. Chirimuuta, 2024). This would lead to an undesirable position for eAC theorists, as they tend to be realists (cf. Hutto, 2018, Raja & Sanches de Oliveira, 2024).

In response, Kirchhoff (2025a, 2025b) argues that cognitivism, particularly its representationalist commitment, fails to capture actual causal explanations and therefore precludes intervention. He claims that “no evidence is provided for the claim that [neural representation] is apt to serve cognitive science’s aim for intervention” (2025b, p. 17). As shown in Section 3, however, this is arguably false. On an interventionist account of causation, the evidence discussed there supports the claim that neural representations are causally efficacious (cf. footnote 11). On a manipulativist account of explanation, they also qualify as causally explanatory. Kirchhoff’s objection therefore does not hold.

A second argument he raises concerns idealizations. Kirchhoff contends that idealizations in the cognitive sciences differ significantly from those in the physical sciences (2025a, 2025b). In his understanding, a consequence of this disanalogy is that physical sciences employ idealizations that track real causal factors, whereas cognitive sciences cannot make the same claim. To defend this disanalogy, he argues that idealizations in cognitive science fail to distinguish irrelevant causal factors or to highlight relevant ones (2025a). He further maintains that there is greater disagreement over the ontological status of entities in cognitive science than in physical science, since physical scientists, he claims, possess a more solid understanding of the parts involved in their explanatory models (2025b).

These arguments are problematic for several reasons. Firstly, when Kirchhoff claims that physical scientists have a more secure grasp of the parts in their models than cognitive scientists, he overlooks the central role of idealizations in physics. Physical sciences have long relied on idealized models, even when their parts were poorly understood. For instance, Maxwell’s models of electromagnetism and of gases as billiard balls were developed when almost no empirical knowledge about these phenomena was available (Suárez, 2024). Galileo’s geometric model of motion was produced under similarly limited epistemic conditions (Hughes, 2010). Moreover, the supposed “solid understanding” of the parts often depends on other, highly idealized models that are themselves far less well understood (Potochnik, 2017; Massimi, 2022, Suárez, 2024, ch. 2). For instance, a well-developed model of the atom depends on much less understood models of the electron and, even more so, of other subatomic particles

(Massimi, 2022). This undercuts the claimed disanalogy between physical and cognitive sciences idealizations.

Secondly, Kirchhoff's claim that disputes over ontology are greater in cognitive science rests on selective evidence. Nonetheless, his examples are drawn solely from philosophical debates about cognitive science. If one considered the equally extensive philosophical debates about the ontology of physics, the level of disagreement would not appear any less. The comparison is therefore made on uneven grounds.

Thirdly, and most importantly, Kirchhoff's assertion that cognitive science idealizations fail to highlight causal factors overlooks clear counterexamples, including research on the engram (Section 3). As discussed there, representational assumptions guided scientists to search for causal processes and entities capable of fulfilling representational roles, such as vehicles and their manipulation to control memory behavior. Moreover, as Chirimuuta (2024) argues, the representationalist framework also helps exclude less stable causal factors that would hinder explanatory aims. In this respect, cognitivist idealizations perform functions analogous to those of the physical sciences: they highlight causal factors and render them more tractable. Thus, there is no significant disanalogy between idealizations in physical and cognitive sciences.<sup>17</sup>

Another metaphysical argument might claim that eAC itself constitutes a metaphysical hypothesis about the nature of cognition, and that accepting it entails the falsity of cognitivism. In other words, one could treat the debate between cognitivism and eAC as a zero-sum metaphysical dispute: if one is true, the other must be false. But here again, two clarifications are needed.

Firstly, even when eAC is understood as a metaphysical hypothesis, this does not entail that it must deny the reality of all entities or processes posited by cognitivism. As seen in Section 5, one can accept that eAC captures certain real aspects of cognition without denying that other aspects are better captured by internalist, representational models. Moreover, as already noted, different models can be compatible and integrated even when they advance seemingly incompatible theses, since they often target different causal patterns (Mitchell, 2002; Potochnik & Sanches de Oliveira, 2020; Kirchhoff, 2025a). In such cases, each thesis may be true of its own target while false of the other's. Notably, one needs not to be an anti-realist to accept it (cf. Giere, 2006, Massimi, 2022, contra Kirchhoff, 2025b). This situation also characterizes the dispute between cognitivism and eAC. Hence, the eliminativist stance is an unmotivated choice,

---

<sup>17</sup> An additional problem is that Kirchhoff's (2025a, 2025b) criticism is directed exclusively at strongly semantic notions of representation.

not a necessary consequence of eAC’s metaphysical commitments. Indeed, as shown in Section 4, some eAC proponents—including those who continue to identify with the label “radical”—explicitly adopt a pluralist perspective.

Secondly, given the empirical and philosophical considerations discussed throughout this paper, a metaphysically strong version of eAC—one that entails the falsity of cognitivism—appears unjustified. Instead, what follows from our discussion is that eAC should adopt a more relaxed metaphysical stance, compatible with pluralism. That is, instead of claiming that the eAC model is the *only* true description of cognition, one should acknowledge that it may be true of *certain* aspects of cognition, while other models (including cognitivist ones) may be true of *other* aspects.

This more moderate metaphysical stance aligns with what Raja and Sanches de Oliveira (2024) call pragmatic realism: the idea that different scientific frameworks explore different causal patterns and thereby illuminate different aspects of a complex phenomenon. As discussed in Section 4, pragmatic realism (or, causal patterns realism) aligns better with a pluralist interpretation of eAC rather than an eliminativist one as cognitivism and eAC do explore different causal patterns. On this view, eAC and cognitivism are not mutually exclusive but rather complementary, depending on the aspect of cognition under investigation and the explanatory goals of the inquiry.

In conclusion, metaphysical arguments for the elimination of cognitivism either rest on questionable assumptions or lead to implausible consequences. A more promising route for eAC is to embrace a pragmatic, pluralist metaphysics—one that accommodates multiple explanatory models and recognizes the multifaceted nature of cognition.

## 8 Conclusions

In this paper, I analyzed eAC as offering alternatives to cognitivism. After introducing the cognitivist framework and presenting a concrete example of its success, I distinguished between two radical positions within eAC: pluralism and eliminativism. On the one hand, eAC pluralists argue that there are valuable frameworks—such as eAC—that do not adopt the tenets of cognitivism but are nonetheless successful and fruitful, all while acknowledging the explanatory power and fruitfulness of cognitivism itself. On the other hand, eAC eliminativists aim to terminate cognitivism as a valuable framework altogether.

Throughout the paper, I critically assessed the main arguments offered by eAC proponents in support of eliminativism. These included claims about cognitivism’s alleged

explanatory inadequacy, its conceptual problems, and its supposed metaphysical failure. I argued that these criticisms ultimately fail to justify the rejection of cognitivism. Many stem from misunderstandings about the nature of scientific explanation, the role of concepts in scientific practice or of cognitivism itself.

As discussed in Section 4, eAC and cognitivism explore different causal patterns that are explanatory for behavior, thereby illuminating distinct aspects of cognitive phenomena. This pluralism fits well with the core epistemological commitments of some versions of eAC, namely those associated with ecological psychology—particularly the idea that scientific knowledge and explanatory frameworks are situated within ecological-enactive contexts. Given this situatedness, we should expect a diversity of theories and frameworks, each appropriate to specific explanatory purposes. In this context-sensitive landscape, there is no principled reason to exclude cognitivism, which continues to offer valuable tools for understanding and intervening in aspects of cognition.

## References

- Abdou, K., Shehata, M., Choko, K., Nishizono, H., Matsuo, M., Muramatsu, S., & Inokuchi, K. (2018). Synapse-specific representation of the identity of overlapping memory engrams. *Science*, *360*(6394), 1227–1231. <https://doi.org/10.1126/science.aat3810>
- Adams, F., & Aizawa, K. (2010). Defending the Bounds of Cognition. In *The Extended Mind* (pp. 67–80). The MIT Press. <https://doi.org/10.7551/mitpress/9780262014038.003.0004>
- Aizawa, K. (2015). What is this cognition that is supposed to be embodied? *Philosophical Psychology*, *28*(6), 755–775. <https://doi.org/10.1080/09515089.2013.875280>
- Aizawa, K. (2017). Cognition and behavior. *Synthese*, *194*(11), 4269–4288. <https://doi.org/10.1007/s11229-014-0645-5>
- Allen, C. (2017). On (not) defining cognition. *Synthese*, *194*(11), 4233–4249. <https://doi.org/10.1007/s11229-017-1454-4>
- Allott, N., Lohndal, T., & Rey, G. (2021). Chomsky’s “Galilean” Explanatory Style. In N. Allott, T. Lohndal, & G. Rey (Eds.), *A Companion to Chomsky* (1st ed., pp. 515–528). Wiley. <https://doi.org/10.1002/9781119598732.ch33>
- Andonovski, N. (2021). Causation and mnemonic roles: On Fernández’s Functionalism. *Estudios de Filosofía*, *64*, 139–153. <https://doi.org/10.17533/udea.ef.n64a07>
- Andonovski, N. (2023). Auto-noesis and the Galilean science of memory: Explanation, idealization, and the role of crucial data. *European Journal for Philosophy of Science*, *13*(3), 42. <https://doi.org/10.1007/s13194-023-00548-3>
- Andonovski, N. (2024). Engrams as mental files. *Synthese*, *204*(6), 154. <https://doi.org/10.1007/s11229-024-04801-x>
- Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2020). *Memory* (Third edition). Routledge, Taylor & Francis Group.
- Baggs, E., & Sanches De Oliveira, G. (2024). Rewilding psychology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1910). <https://doi.org/10.1098/rstb.2023.0287>

- Baker, B., Lansdell, B., & Kording, K. P. (2022). Three aspects of representation in neuroscience. *Trends in Cognitive Sciences*, 26(11), 942–958. <https://doi.org/10.1016/j.tics.2022.08.014>
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. Psychology - L. Erlbaum.
- Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180(1), 33–45. <https://doi.org/10.1007/s11229-009-9565-1>
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61(1–2), 127–148. <https://doi.org/10.1007/BF00385837>
- Brancazio, N. (2018). Irreducible Aspects of Embodiment: Situating Scientist and Subject. *Australasian Philosophical Review*, 2(2), 219–223. <https://doi.org/10.1080/24740500.2018.1552101>
- Campbell, S. (2014). *Our faithfulness to the past: The ethics and politics of memory*. Oxford University Press.
- Cao, R. (2022). Putting representations to use. *Synthese*, 200(2), 151. <https://doi.org/10.1007/s11229-022-03522-3>
- Carnap, R. (1950). *Logical foundations of probability*. Univ. of Chicago Pr.
- Carvalho, E. M. D., & Rolla, G. (2020). An Enactive-Ecological Approach to Information and Uncertainty. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00588>
- Chang, H. (2014). *Is water H<sub>2</sub>O? Evidence, realism and pluralism*. Springer.
- Chemero, A. (2000). Anti-Representationalism and the Dynamical Stance. *Philosophy of Science*, 67(4), 625–647. <https://doi.org/10.1086/392858>
- Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.
- Chemero, A., & Silberstein, M. (2008). After the Philosophy of Mind: Replacing Scholasticism with Science\*. *Philosophy of Science*, 75(1), 1–27. <https://doi.org/10.1086/587820>
- Chirimuuta, M. (2018). Explanation in Computational Neuroscience: Causal and Non-causal. *The British Journal for the Philosophy of Science*, 69(3), 849–880. <https://doi.org/10.1093/bjps/axw034>
- Chirimuuta, M. (2024). *The Brain Abstracted: Simplification in the History and Philosophy of Neuroscience*. The MIT Press. <https://doi.org/10.7551/mitpress/13804.001.0001>
- Chomsky, N. (1978). A theory of core grammar. *Glott*, 1:1-26.
- Chomsky, N. (1980). *Rules and representations*. Columbia Univ. Press.
- Choucry, A., Nomoto, M., & Inokuchi, K. (2024). Engram mechanisms of memory linking and identity. *Nature Reviews Neuroscience*, 25(6), 375–392. <https://doi.org/10.1038/s41583-024-00814-0>
- Ciaunica, A. (2022). The map, the territory, and the cartographer: Linking the “pure” formal models to the “murky” material world. *Behavioral and Brain Sciences*, 45. <https://doi.org/10.1017/s0140525x22000590>
- Colaço, D. (2022). What Counts as a Memory? Definitions, Hypotheses, and “Kinding in Progress.” *Philosophy of Science*, 89(1), 89–106. <https://doi.org/10.1017/psa.2021.14>
- Colombo, M., & Piccinini, G. (2024). *The Computational Theory of Mind*. Cambridge University Press. <https://doi.org/10.1017/9781009183734>
- Craver, C. F. (2013). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience* (First published in paperback, reprinted 2013). Clarendon Press.
- Crippen, M. (2025). Enactivism: A newish name for mostly old ideas? *Phenomenology and the Cognitive Sciences*, 24(1), 103–127. <https://doi.org/10.1007/s11097-024-10019-6>
- De Brigard, F. (2020). The Explanatory Indispensability of Memory Traces. *The Harvard Review of Philosophy*, 27, 23–47. <https://doi.org/10.5840/harvardreview202072328>

- De Brigard, F. (2024). Episodic memory without auto-noetic consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 379(1913), 20230410. <https://doi.org/10.1098/rstb.2023.0410>
- Deisseroth, K. (2011). Optogenetics. *Nature Methods*, 8(1), 26–29. <https://doi.org/10.1038/nmeth.f.324>
- Di Paolo, E. A., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic Bodies: The Continuity between Life and Language*. The MIT Press. <https://doi.org/10.7551/mitpress/11244.001.0001>
- Dutilh Novaes, C. (2020a). Carnap meets Foucault: Conceptual engineering and genealogical investigations. *Inquiry*, 1–27. <https://doi.org/10.1080/0020174x.2020.1860122>
- Dutilh Novaes, C. (2020b). Carnapian explication and ameliorative analysis: A systematic comparison. *Synthese*, 197(3), 1011–1034. <https://doi.org/10.1007/s11229-018-1732-9>
- Dutilh Novaes, C., & Reck, E. (2017). Carnapian explication, formalisms as cognitive tools, and the paradox of adequate formalization. *Synthese*, 194(1), 195–215. <https://doi.org/10.1007/s11229-015-0816-z>
- Einstein, G. (2012). Situated Neuroscience: Exploring Biologies of Diversity. In R. Bluhm, A. J. Jacobson, and H. L. Maibom (Eds.), *Neurofeminism: Issues at the Intersection of Feminist Theory and Cognitive Science*. New York: Palgrave-Macmillan.
- Favela, L. H. (2024). *The Ecological Brain: Unifying the sciences of brain, body, and environment*. Routledge.
- Favela, L. H., & Chemero, A. (2015). The animal-environment system. In Y. Coello & M. H. Fischer (Eds.), *Foundations of embodied cognition. Volume 1: Perceptual and emotional embodiment* (pp. 59–74). Routledge, Taylor & Francis Group.
- Favela, L. H., & Machery, E. (2023). Investigating the concept of representation in the neural and psychological sciences. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1165622>
- Favela, L. H., & Machery, E. (2025). The concept of representation in the brain sciences: The current status and ways forward. *Mind & Language*, 40(2), 215–225. <https://doi.org/10.1111/mila.12531>
- Feest, U. (2017). Phenomena and Objects of Research in the Cognitive and Behavioral Sciences. *Philosophy of Science*, 84(5), 1165–1176. <https://doi.org/10.1086/694155>
- Feest, U. (2025). *Operationism in Psychology: An Epistemology of Exploration*. University of Chicago Press.
- Floridi, L. (2010). *Information: A very short introduction*. Oxford University press.
- Fodor, J. A. (2001). *The mind doesn't work that way: The scope and limits of computational psychology* (1. paperback ed). MIT Press.
- Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press.
- Gallagher, S. (2018a). Building a stronger concept of embodiment. In *The Oxford Handbook of 4E Cognition* (1st ed., pp. 353–368). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198735410.013.18>
- Gallagher S. (2018b) Decentering the brain: Embodied cognition and the critique of neurocentrism and narrow-minded philosophy of mind. *Constructivist Foundations* 14(1): 8–21.
- Gallagher, S. (2018c). The Extended Mind: State of the Question. *The Southern Journal of Philosophy*, 56(4), 421–447. <https://doi.org/10.1111/sjp.12308>
- Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: Why cognitive science will transform neuroscience* (Nachdr.). Wiley-Blackwell.
- Giere, R. N. (2006). *Scientific perspectivism*. University of Chicago Press.

- Gładziejewski, P., & Miłkowski, M. (2017). Structural representations: Causally relevant and different from detectors. *Biology & Philosophy*, 32(3), 337–355. <https://doi.org/10.1007/s10539-017-9562-6>
- Goldman, A. I. (2012). A Moderate Approach to Embodied Cognitive Science. *Review of Philosophy and Psychology*, 3(1), 71–88. <https://doi.org/10.1007/s13164-012-0089-0>
- Griegen, R., Tekülve, J., Zibner, S. K. U., Lins, J., Schneegans, S., & Schöner, G. (2020). Scene memory and spatial inhibition in visual search: A neural dynamic process model and new experimental evidence. *Attention, Perception, & Psychophysics*, 82(2), 775–798. <https://doi.org/10.3758/s13414-019-01898-y>
- Hitchcock, C., & Woodward, J. (2003). Explanatory Generalizations, Part II: Plumbing Explanatory Depth. *Noûs*, 37(2), 181–199. <https://doi.org/10.1111/1468-0068.00435>
- Hochstein, E. (2016). One mechanism, many models: A distributed theory of mechanistic explanation. *Synthese*, 193(5), 1387–1407. <https://doi.org/10.1007/s11229-015-0844-8>
- Hughes, R. I. G. (2010). *The Theoretical Practices of Physics: Philosophical Essays*. Oxford University Press.
- Hutto, D. D. (2018). REC: Revolution Effected by Clarification. *Topoi*, 36(3), 377–391. <https://doi.org/10.1007/s11245-015-9358-8>
- Hutto, D. D. (2023). Remembering without a trace? Moving beyond trace minimalism. In A. Sant’Anna, C. J. McCarroll, & K. Michaelian (Eds.), *Current controversies in philosophy of memory* (pp. 61–81). Routledge.
- Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. MIT Press.
- Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. MIT Press.
- Hutto, D. D., & Myin, E. (2018). Going Radical. In *The Oxford Handbook of 4E Cognition* (1st ed., pp. 95–116). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198735410.013.5>
- Hutto, D. D., Peeters, A., & Segundo-Ortin, M. (2017). Cognitive ontology in flux: The possibility of protean brains. *Philosophical Explorations*, 20(2), 209–223. <https://doi.org/10.1080/13869795.2017.1312502>
- Hutto, D. D., & Satne, G. (2015). The Natural Origins of Content. *Philosophia*, 43(3), 521–536. <https://doi.org/10.1007/s11406-015-9644-0>
- Iani, F. (2019). Embodied memories: Reviewing the role of the body in memory processes. *Psychonomic Bulletin & Review*, 26(6), 1747–1766. <https://doi.org/10.3758/s13423-019-01674-x>
- Josselyn, S. A., & Frankland, P. W. (2018). Memory Allocation: Mechanisms and Function. *Annual Review of Neuroscience*, 41(1), 389–413. <https://doi.org/10.1146/annurev-neuro-080317-061956>
- Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science*, 367(6473), eaaw4325. <https://doi.org/10.1126/science.aaw4325>
- Kandel, E. R. (2007). *In search of memory: The emergence of a new science of mind*. Norton.
- Kirchhoff, M. D. (2025a). Idealization and mental fictionalism. *Philosophical Psychology*, 1–22. <https://doi.org/10.1080/09515089.2025.2552422>
- Kirchhoff, M. D. (2025b). *The idealized mind: From model-based science to cognitive science*. The MIT Press.
- Kriegeskorte, N., & Diedrichsen, J. (2019). Peeling the Onion of Brain Representations. *Annual Review of Neuroscience*, 42(1), 407–432. <https://doi.org/10.1146/annurev-neuro-080317-061906>
- Kunz, L., Deuker, L., Zhang, H., & Axmacher, N. (2018). Tracking Human Engrams Using Multivariate Analysis Techniques. In *Handbook of Behavioral Neuroscience* (Vol. 28, pp. 481–508). Elsevier. <https://doi.org/10.1016/B978-0-12-812028-6.00026-4>

- Lazzeri, F. (2016). On some common objections to a behavioral approach to psychological categories. *Philosophical Psychology*, 29(3), 405–418. <https://doi.org/10.1080/09515089.2015.1121540>
- Leonelli, S. (2016). *Data-centric biology: A philosophical study*. The University of Chicago Press.
- Luis-Islas, J., Luna, M., Floran, B., & Gutierrez, R. (2022). Optoception: Perception of Optogenetic Brain Perturbations. *Eneuro*, 9(3), ENEURO.0216-22.2022. <https://doi.org/10.1523/ENEURO.0216-22.2022>
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.
- Martínez, M. (2019). Representations Are Rate-Distortion Sweet Spots. *Philosophy of Science*, 86(5), 1214–1226. <https://doi.org/10.1086/705493>
- Massimi, M. (2022). *Perspectival realism*. Oxford University press.
- Meyer, R., & Brancazio, N. (2022). Putting down the revolt: Enactivism as a philosophy of nature. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.948733>
- Meyer R. & Brancazio N. (2023) Enactivism: Utopian & scientific. *Constructivist Foundations* 19(1): 1–11.
- Milkowski, M. (2015). The Hard Problem Of Content: Solved (Long Ago). *Studies in Logic, Grammar and Rhetoric*, 41(1), 73–88. <https://doi.org/10.1515/slgr-2015-0021>
- Mitchell, S. D. (2002). Integrative Pluralism. *Biology and Philosophy*, 17(1), 55–70. <https://doi.org/10.1023/A:1012990030867>
- Myin, E., & Van Dijk, L. (2022). The Is and Oughts of Remembering. *Topoi*, 41(2), 275–285. <https://doi.org/10.1007/s11245-021-09784-9>
- Najenson, J. (2021). What have we learned about the engram? *Synthese*, 199(3–4), 9581–9601. <https://doi.org/10.1007/s11229-021-03216-2>
- Najenson, J. (forthcoming). Encoding without perceiving: Can memories be implanted? *Philosophical Psychology*, 1–28. <https://doi.org/10.1080/09515089.2023.2295927>
- Nambu, M. F., Lin, Y.-J., Reuschenbach, J., & Tanaka, K. Z. (2022). What does engram encode?: Heterogeneous memory engrams for different aspects of experience. *Current Opinion in Neurobiology*, 75, 102568. <https://doi.org/10.1016/j.conb.2022.102568>
- Neander, K. (2017). *A mark of the mental: In defense of informational teleosemantics*. The MIT Press.
- Noë, A. (2004). *Action in perception*. MIT Press.
- Piccinini, G. (2015). *Physical Computation: A Mechanistic Account*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199658855.001.0001>
- Piccinini, G. (2020). *Neurocognitive mechanisms: Explaining biological cognition* (First edition). Oxford University Press.
- Piccinini, G. (2022). Situated Neural Representations: Solving the Problems of Content. *Frontiers in Neurobotics*, 16, 846979. <https://doi.org/10.3389/fnbot.2022.846979>
- Pietroski, P., & Rey, G. (1995). When Other Things Aren't Equal: Saving *Ceteris Paribus* Laws from Vacuity. *The British Journal for the Philosophy of Science*, 46(1), 81–110. <https://doi.org/10.1093/bjps/46.1.81>
- Potochnik, A. (2017). *Idealization and the aims of science*. The University of Chicago press.
- Potochnik, A., & Sanches De Oliveira, G. (2020). Patterns in Cognitive Phenomena and Pluralism of Explanatory Styles. *Topics in Cognitive Science*, 12(4), 1306–1320. <https://doi.org/10.1111/tops.12481>
- Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2022). The Best Game in Town: The Re-Emergence of the Language of Thought Hypothesis Across the Cognitive Sciences. *Behavioral and Brain Sciences*, 1–55. <https://doi.org/10.1017/S0140525X22002849>

- Raja, V. (2018). A Theory of Resonance: Towards an Ecological Cognitive Architecture. *Minds and Machines*, 28(1), 29–51. <https://doi.org/10.1007/s11023-017-9431-8>
- Raja, V. (2021). Resonance and radical embodiment. *Synthese*, 199(S1), 113–141. <https://doi.org/10.1007/s11229-020-02610-6>
- Raja, V. (2024). The motifs of radical embodied neuroscience. *European Journal of Neuroscience*, 60(5), 4738–4755. <https://doi.org/10.1111/ejn.16434>
- Raja, V., & Anderson, M. L. (2019). Radical Embodied Cognitive Neuroscience. *Ecological Psychology*, 31(3), 166–181. <https://doi.org/10.1080/10407413.2019.1615213>
- Raja, V., & Sanches De Oliveira, G. (2024). Two species of realism. *European Journal for Philosophy of Science*, 14(3). <https://doi.org/10.1007/s13194-024-00596-3>
- Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R. L., Ryan, T. J., & Tonegawa, S. (2013). Creating a False Memory in the Hippocampus. *Science*, 341(6144), 387–391. <https://doi.org/10.1126/science.1239073>
- Ramsey, W. M. (2025). The Hard Problem of Content is Neither. *Review of Philosophy and Psychology*, 16(1), 153–174. <https://doi.org/10.1007/s13164-023-00714-9>
- Righetti, F. (2021). The pre-reflective roots of the madeleine-memory: A phenomenological perspective. *Phenomenology and the Cognitive Sciences*, 22(2), 479–499. <https://doi.org/10.1007/s11097-021-09774-7>
- Righetti, F. (2024). Re-enactment and embodied resonance in episodic memory: Reconciling phenomenological approaches and constructive theories. *Phenomenology and the Cognitive Sciences*. <https://doi.org/10.1007/s11097-024-10026-7>
- Robins, S. K. (2016). Optogenetics and the mechanism of false memory. *Synthese*, 193(5), 1561–1583. <https://doi.org/10.1007/s11229-016-1045-9>
- Robins, S. K. (2023). The 21st century engram. *WIREs Cognitive Science*, 14(5), e1653. <https://doi.org/10.1002/wcs.1653>
- Robins, S. K. (forthcoming). A place for the memory trace. In Aronowitz, S. & Nadel, L. (eds.). *Memory, Space, and Time*. Oxford University Press.
- Rolla, G. (2018). Radically Enactive High Cognition. *Dissertatio*, 47, 26–41.
- Rolla, G. (2021). *A mente enativa*. Editora Fi. <https://doi.org/10.22350/9786559173341>
- Rolla, G. (2022). Do babies represent? On a failed argument for representationalism. *Synthese*, 200(4). <https://doi.org/10.1007/s11229-022-03728-5>
- Rolla, G., & Huffermann, J. (2022). Converging enactivisms: Radical enactivism meets linguistic bodies. *Adaptive Behavior*, 30(4), 345–359. <https://doi.org/10.1177/10597123211020782>
- Rolla, G., & Novaes, F. (2022). Ecological-enactive scientific cognition: Modeling and material engagement. *Phenomenology and the Cognitive Sciences*, 21(3), 625–643. <https://doi.org/10.1007/s11097-020-09713-y>
- Ross, L. N. (2018). Causal Selection and the Pathway Concept. *Philosophy of Science*, 85(4), 551–572. <https://doi.org/10.1086/699022>
- Ross, L. N. (2025). *Explanation in Biology* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009300940>
- Ross, L. N. (Forthcoming). Causal Control: A Rationale for Causal Selection. In: *Minnesota Studies in Philosophy of Science on "Causation in Biology*.
- Rupert, R. (2005). Minding One's Cognitive Systems: When Does a Group of Minds Constitute a Single Cognitive Unit? *Episteme*, 1(3), 177–188. <https://doi.org/10.3366/epi.2004.1.3.177>
- Russell, J. L. (2023). Problems for enactive psychiatry as a practical framework. *Philosophical Psychology*, 36(8), 1458–1481. <https://doi.org/10.1080/09515089.2023.2174423>

- Sanches De Oliveira, G. (2023a). Explanatory Diversity and Embodied Cognitive Science: Reflexivity Motivates Pluralism. In *Studies in Brain and Mind* (pp. 51–76). Springer International Publishing. [https://doi.org/10.1007/978-3-031-39744-8\\_4](https://doi.org/10.1007/978-3-031-39744-8_4)
- Sanches De Oliveira, G. (2023b). The strong program in embodied cognitive science. *Phenomenology and the Cognitive Sciences*, 22(4), 841–865. <https://doi.org/10.1007/s11097-022-09806-w>
- Sanches De Oliveira, G., Raja, V., & Chemero, A. (2021). Radical embodied cognitive science and “Real Cognition.” *Synthese*, 198(S1), 115–136. <https://doi.org/10.1007/s11229-019-02475-4>
- Sanches De Oliveira, G. S., & Chemero, A. (2015). Against Smallism And Localism. *Studies in Logic, Grammar and Rhetoric*, 41(1), 9–23. <https://doi.org/10.1515/slgr-2015-0017>
- Sanches De Oliveira, G., Van Es, T., & Hipólito, I. (2023). Scientific practice as ecological-enactive co-construction. *Synthese*, 202(1). <https://doi.org/10.1007/s11229-023-04215-1>
- Schacter, D. L. (2021). *The seven sins of memory: How the mind forgets and remembers* (Updated edition). Mariner Books ; Houghton Mifflin Harcourt.
- Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 773–786. <https://doi.org/10.1098/rstb.2007.2087>
- Schöner, G. (2016). *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press.
- Segundo-Ortin, M., & Hutto, D. D. (2021). Similarity-based cognition: Radical enactivism meets cognitive neuroscience. *Synthese*, 198(S1), 5–23. <https://doi.org/10.1007/s11229-019-02505-1>
- Shagrir, O. (2012). Structural Representations and the Brain. *The British Journal for the Philosophy of Science*, 63(3), 519–545. <https://doi.org/10.1093/bjps/axr038>
- Shea, N. (2018). *Representation in cognitive science* (First edition). Oxford University Press.
- Silberstein, M., & Chemero, A. (2013). Constraints on Localization and Decomposition as Explanatory Strategies in the Biological Sciences. *Philosophy of Science*, 80(5), 958–970. <https://doi.org/10.1086/674533>
- Suárez, M. (2024). *Inference and representation: A study in modeling science*. The University of Chicago Press.
- Sun, C., Yang, W., Martin, J., & Tonegawa, S. (2020). Hippocampal neurons represent events as transferable units of experience. *Nature Neuroscience*, 23(5), 651–663. <https://doi.org/10.1038/s41593-020-0614-x>
- Telakivi, P. (2023). *Extending the Extended Mind: From Cognition to Consciousness*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-35624-7>
- Thagard, P. (1992). *Conceptual revolutions*. Princeton University Press.
- Tonegawa, S., Morrissey, M. D., & Kitamura, T. (2018). The role of engram cells in the systems consolidation of memory. *Nature Reviews Neuroscience*, 19(8), 485–498. <https://doi.org/10.1038/s41583-018-0031-2>
- Tulving, E. (2005). Episodic memory and autonoesis: uniquely human? In: Terrace, H. & Metcalfe, J. (eds.). *The missing link in cognition: origins of self-reflective consciousness*. Oxford: Oxford University Press.
- Van Gelder, T. (1995). What Might Cognition Be, If Not Computation?: *Journal of Philosophy*, 92(7), 345–381. <https://doi.org/10.2307/2941061>
- Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21(5), 615–628. <https://doi.org/10.1017/s0140525x98001733>
- Vernazzani, A. (2019). The structure of sensorimotor explanation. *Synthese*, 196(11), 4527–4553. <https://doi.org/10.1007/s11229-017-1664-9>

- Vetere, G., Tran, L. M., Moberg, S., Steadman, P. E., Restivo, L., Morrison, F. G., Ressler, K. J., Josselyn, S. A., & Frankland, P. W. (2019). Memory formation in the absence of experience. *Nature Neuroscience*, 22(6), 933–940. <https://doi.org/10.1038/s41593-019-0389-0>
- Vilarroya, O. (2017). Neural Representation. A Survey-Based Analysis of the Notion. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01458>
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation* (1st ed.). Oxford University Press New York. <https://doi.org/10.1093/0195155270.001.0001>
- Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, 25(3), 287–318. <https://doi.org/10.1007/s10539-010-9200-z>
- Woodward, J. (2016). The problem of variable choice. *Synthese*, 193(4), 1047–1072. <https://doi.org/10.1007/s11229-015-0810-5>
- Woodward, J., & Hitchcock, C. (2003). Explanatory Generalizations, Part I: A Counterfactual Account. *Noûs*, 37(1), 1–24. <https://doi.org/10.1111/1468-0068.00426>
- Xie, H., & Miyata, K. (2014). Real-time simulation of lightweight rigid bodies. *The Visual Computer*, 30(1), 81–92. <https://doi.org/10.1007/s00371-013-0783-7>
- Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, 148(2), 201–219. <https://doi.org/10.1007/s11098-008-9324-z>
- Zednik, C. (2011). The Nature of Dynamical Explanation. *Philosophy of Science*, 78(2), 238–263. <https://doi.org/10.1086/659221>