



## Explanations in cognitive science: unification versus pluralism

Marcin Miłkowski<sup>1</sup>  · Mateusz Hohol<sup>2</sup> 

Received: 29 June 2020 / Accepted: 30 June 2020 / Published online: 15 July 2020  
© Springer Nature B.V. 2020

### Abstract

The debate between the defenders of explanatory unification and explanatory pluralism has been ongoing from the beginning of cognitive science and is one of the central themes of its philosophy. Does cognitive science need a grand unifying theory? Should explanatory pluralism be embraced instead? Or maybe local integrative efforts are needed? What are the advantages of explanatory unification as compared to the benefits of explanatory pluralism? These questions, among others, are addressed in this *Synthese*'s special issue. In the introductory paper, we discuss the background of the questions, distinguishing integrative theorizing from building unified theories. On the one hand, integrative efforts involve collaboration between various disciplines, fields, approaches, or theories. These efforts could even be quite temporary, without establishing any long-term institutionalized fields or disciplines, but could also contribute to developing new interfield theories. On the other hand, unification can rely on developing complete theories of mechanisms and representations underlying all cognition, as Newell's "unified theories of cognition", or may appeal to grand principles, as predictive coding. Here, we also show that unification in contemporary cognitive science goes beyond reductive unity, and may involve various forms of joint efforts and division of explanatory labor. This conclusion is one of the themes present in the content of contributions constituting the special issue.

**Keywords** Cognitive science · Explanation · Integration · Mechanisms · Philosophy of cognitive science · Pluralism · Unification

---

✉ Marcin Miłkowski  
marcin.milkowski@gmail.com

Mateusz Hohol  
mateusz.hohol@uj.edu.pl

<sup>1</sup> Institute of Philosophy and Sociology, Polish Academy of Sciences, Warsaw, Poland

<sup>2</sup> Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Kraków, Poland

## 1 Introduction: various kinds of explanatory unity

Does cognitive science need a grand unifying theory? Should explanatory pluralism be embraced instead? Or maybe local integrative efforts are needed? These are the questions that are addressed in this special issue. Its purpose is to discuss the nature of explanations in cognitive science in the context of the debate between defenders of explanatory unification and explanatory pluralism.

Since its very beginnings, cognitive science has been an interdisciplinary endeavor. The thought was that, to explain cognitive phenomena, multiple research perspectives must converge together. Hence, collaboration across scientific disciplines in explaining cognitive phenomena may result from the nature of cognitive science. This collaboration, however, need not imply the emergence of a highly theoretically unified, new discipline. As things stand, scientific practice in cognitive science does reflect divergent views on whether cognitive science should become really unified or merely loosely coordinated and integrated.

On the one hand, if collaboration among various disciplines, research fields, or traditions that make up cognitive science ceases to be essential for explanatory success, this could imply that cognitive science has lost its appeal and exhausted its potential as a research program (Núñez et al. 2019). Thus, at least some integrative effort is required for cognitive science to maintain its existence. This, however, need not imply that cognitive science should embrace a single, coherent view on cognition or on any other subject matter of interest to cognitive scientists. At least, this was the stance of the founders of cognitive science who were predominantly pluralistic (Gentner 2019). On the other hand, one of the goals of (at least some) the founders of cognitive science was to create a unified research discipline with a particularly pluralistic methodology, but a single object of research. Some even argued that unified theories of cognition should avoid fragmentation, isolationism, and ad hoc assumptions that seemed to plague cognitive psychology (Newell 1973).

These two approaches actually reflect the philosophical views on the unity of science that were already present in the Vienna Circle. On the one hand, Neurath (1937) argued that sciences should be coordinated and that there should be no isolated islands among them (Potochnik 2011). In the analysis proposed by Potochnik, this kind of coordination consists of three kinds of collaboration: gathering evidence from multiple sources, terminological clarification, and, most ambitiously, building integrated models. In the last case, coordination leads to *integration*.<sup>1</sup> On another hand, a minority position in the Vienna Circle was that all sciences should be reduced to one grand unifying theory (Carnap 1928). This view of *reductive unity* became much more prominent in the 1960s, not least thanks to the programmatic manifesto of Oppenheim and Putnam (1958). Note, however, that the distinction

<sup>1</sup> While we applaud Potochnik's (2011) analysis of coordination, her term "coordinate unification" may be misleading because there are alternative approaches to unity that go beyond reduction or coordination between approaches, methods, theories, or fields. Our choice of terminology is somewhat arbitrary, as there is no consensus in the use of these terms but it allows us to avoid possible fallacies of equivocation.

between these two kinds of unity is not a dichotomy because it is not exhaustive: More kinds may exist.

The fact that these two options do not exhaust the realm of possibilities is particularly important because reductive unity in the sense defended by Oppenheim and Putnam seems a lost case. In philosophy, the issues of theoretical unity have been traditionally discussed in the context of reduction, which was criticized by appeals to multiple realizability (Fodor 1974; Kincaid 1990; Putnam 1967). Multiple realizability undermines traditional reduction because the kinds in the theory to be reduced cannot be neatly identified with (collections of) the kinds in the reducing theory, in which there are usually multiple possible kinds that could realize the kinds of the theory to be reduced. A highly cited example was introduced by Fodor (1974): The kind money in economics cannot be identified with one physical bearer, such as shells, gold coins, or paper banknotes; the set of physical kinds can only be listed in a potentially infinite (and most likely, uncountably infinite) disjunctive predicate. Nonetheless, the crucial problem—even if multiple realizability did not exist or was inessential for reduction (for recent discussion, see Aizawa 2018; Polger and Shapiro 2016)—is that the received view on reduction is barely applicable to cognitive theories. This is because the received view requires theories to be stated in a propositional manner and include lawlike statements. Only then may the reducing theory imply a corrected image of an old theory, given that connections between vocabularies of both theories are also added in terms of so-called bridge principles (Bickle 1998; Hooker 1981a, b, c). But lawlike statements are hardly found in cognitive theorizing (Cummins 2000), even if some argue that they should be (Raja 2019). Moreover, cognitive theories are sometimes stated in terms of formal specifications in artificial languages, not all of which are declarative; some are stated in a mixture of diagrams and verbal descriptions, and some as software (Cooper and Guest 2014). Thus, they are difficult to bring to the traditional propositional format to perform logical derivations. For this reason, from the practical point of view, the issue of reductive unity, at least in its received version, is rather moot. Yet scientific practice need not appeal to reduction.<sup>2</sup> As we will see in what follows, alternative approaches to unification are dominant in cognitive science.

These issues mesh with the debate over what counts as a satisfactory explanation in cognitive science. While defenders of dynamical systems in cognitive science appeal to the received covering law account of explanation (Walmsley 2008; Chemero 2009; cf. Hempel and Oppenheim 1948), it remains a minority view: Most cognitive science does not offer anything like laws. The two main alternatives are the functional account of explanation and the new mechanistic proposal. The functional explanation may be understood in non-causal terms of functional analysis (Cummins 1975, 1983), or in non-spatiotemporal causal terms (Shapiro 2019; Weiskopf 2011). The cognitive phenomenon is understood as a capacity of a complex system, which results from constituent subcapacities of the functional components of that

<sup>2</sup> Arguably, derivational account of reduction is one among many other possible accounts. For example, one can view constitutive mechanistic explanations as reductive (Hensel 2013; Wright 2007).

system. Functional explanation is often linked with computational modeling of cognitive capacities.

One important issue for the functional account of computational modeling is what counts as relevant to cognitive function. The most influential answer usually appeals to the methodology of computational explanation that was proposed by Marr (1982), who required that the complete explanation of a computational system answer three kinds of questions, framed as constitutive for levels of explanation. The computational level is supposed to answer the question of why the computation occurs in a given environment, as well as specify its mathematical function. The level of algorithms and representations yields particular algorithms and representations over which they are defined to compute that function, and the implementation level provides the physical detail of how these algorithms and representations are physically realized. Because Marr's methodology is so influential in cognitive science, it also sparked a long-lasting debate on how to interpret it (see, for example, Bechtel and Shagrir 2015; McClamrock 1991; Peacocke 1986; Shagrir 2010), how to develop it further (Poggio 2012), and whether it is sufficient (Miłkowski 2013; Piccinini and Craver 2011).

The mechanistic approach goes further as it not only requires that these components be spatiotemporally localized but also that the overall explanation be causal (Machamer et al. 2000). The mechanistic approach in cognitive science is usually discussed in the context of computational mechanisms (Bechtel 2008; Miłkowski 2013; Piccinini 2015), even though the explanatory framework is not limited to computational factors and may include distributed processes as well (Miłkowski et al. 2018).

## 2 Integrative coordination

Coordination between various disciplines, fields, approaches, or theories is typical for interdisciplinary collaboration. Thus, integrative efforts are necessary for cognitive science to remain interdisciplinary. These efforts are usually required for explanatory purposes, and collaborations could even be quite ephemeral, without establishing any long-term institutionalized fields or disciplines (Nathan 2017).<sup>3</sup> But they could also coalesce to form new interfield theories (Darden and Maull 1977).

For fields, theories, or explanations to be coordinated or integrated, it is required that these are not fully autonomous. While the derivational reduction of psychology to neurobiology is, at best, a hopeful wish, it does not mean that psychological theorizing is unconstrained by biological facts. As Boone and Piccinini (2016) argue, the connections between cognitive science and neuroscience have become so strong that one can even speak of a scientific revolution, which is underpinned by mechanistic explanations in cognitive neuroscience. The resultant integrative discipline—cognitive neuroscience—was not created by deriving cognitive science from

<sup>3</sup> Nathan calls such integrated explanations *unificatory*, even if he strives to distinguish integration from unification, which only shows again that the terminology in this debate is still in flux.

neuroscience, however. The authors claim that the search for mechanisms of cognition simply forces collaboration and integration of cognitive science and neuroscience. Their view meshes well with the view of scientific unity of neuroscience, termed “mosaic unity” by Craver (2008; see also Miłkowski et al. 2019). Although he uses slightly different terminology to that used in this introduction, Craver defends the view that mechanistic explanations force piecemeal integration. This is not only because constitutive mechanistic explanations are a kind of reductive explanation (Hensel 2013; Wright 2007), but also because various kinds of evidence may be brought to bear on complex explanations of mechanisms.

Craver defends his account in terms of constraints on the space of possible mechanisms, which he identifies with findings (or evidence). Nonetheless, as Miłkowski (2016a) argues, the constraints cited by Craver include methodological principles, such as “causes explain effects and not vice versa;” thus, it is more plausible to think of constraints as shaping the space of representations of mechanisms. Simply put, representations of mechanisms can constrain each other, and can be constrained by methodological principles. The operation of such constraints from discipline B on representations of mechanisms in discipline A implies that A is not autonomous from B, while at the same time, there is no derivational reduction of A from B.

This implies that the mechanistic account of integration is compatible with non-extreme versions of explanatory pluralism: As long as divergent methodologies or theories lead to explanations that can be mutually constrained, cognitive science can still be considered integrated, at least to some extent. Some theories or models may also remain complementary if they have a somewhat different scope. Moreover, critics of the optimistic approach to integrative efforts in cognitive (neuro)science stress that integration is far from being smooth, given that different fields or traditions may understand their terms in a different manner and appeal to divergent methodological standards and experimental protocols, which makes experimental research in neuroscience difficult to bear on psychological research (Sullivan 2009).

The preceding reasons make it unsurprising that the contemporary default position in cognitive science seems to be explanatory pluralism (Colombo and Wright 2017; Dale 2008; Dale et al. 2009; Horst 2016), which can be taken to lead not only to integrative efforts (McCauley and Bechtel 2001; Richardson 2009) but also to isolationist pluralism (Bouwel 2014) and eliminativist pluralism without any integrative ambitions (Barker 2019; Taylor and Vickers 2017) as well. Apparently, proponents of pluralism do not agree that the complete unification of cognitive science is even desired. The reason may be that their cherished theoretical virtues diverge from the ones accepted by proponents of unification. Let us then turn to unification.

### 3 Unifying cognitive science

Newell’s (1973) plea for unification was a direct response to what he considered signs of degeneracy in cognitive science: Ad hoc assumptions in interpretation of experiments, grand questions answered without broader theoretical background, and no cumulation of knowledge.

But discoveries of new phenomena, new task designs, and ad hoc interpretations of observed effects do not suffice for a theory to appear. This assessment is now being repeated in the context of both the replication crisis (Oberauer and Lewandowsky 2019), and the theory crisis in psychology. Regarding the latter, McPhetres et al. (2020) recently found that only about 15% of research published in *Psychological Science* over the last decade explicitly state to test predictions inferred from theories. Newell's answer remains attractive: It is the theory that should allow psychologists to give up on fishing for effects.

Because unified theories are supposed to answer the crisis, which was spawned by fishing effects in a non-theoretical fashion and ad hoc theorizing, Newell does not offer his unification proposal to address the crisis in traditional terms of theory reduction. While philosophers tended to discuss reduction in the context of unifying cognitive science, psychologists and cognitive scientists kept on discussing the issue of how to make their theories unified without an appeal to reduction (Gaj 2016; Sternberg 2005). But this only shows that there is another sense of unification that goes beyond coordination and reduction. Furthermore, it is related to the philosophical proposals that unification is closely linked to understanding and explanation (Bangu 2017; Friedman 1974; Kitcher 1981).

While coordination and reduction senses of unity are related to how pieces of knowledge are brought together, the accounts of explanatory unification need not start from multiple pieces to be connected at all. Instead, defenders of explanatory unification press the case that theoretical unity has unique virtues. Not only should it address the issue of crisis, but it should also remove ad hoc assumptions and provide a systematic understanding of the whole domain of a given theory and all the phenomena that it explains (Bangu 2017). While one could argue that reduction could also serve this purpose, the usual solutions in cognitive science rely on other strategies. They can be brought down to two distinct proposals: Either build a theory of entities (or a single entity) that is responsible for all observed phenomena, or study general principles that govern them.

The first proposal can be understood in terms of functional explanation, as long as the entity in question is understood as having functional capacities to be explained, or, alternatively, in terms of mechanisms, in the sense of the new mechanistic approach to explanation. Thus, this proposal is mostly defended in mainstream cognitive science, which seems not to rely on covering-law explanations. Regardless of whether one endorses functionalism or mechanism, this strategy should produce relatively simple, systematic theories with a large scope, devoid of ad hoc assumptions (as the whole theory is just a theory of this entity, nothing else). They would be intuitively unified because they speak of a single entity. If there are more entities (or kinds of entities), the parsimony of the theory may decrease, but at the same time, it could still provide insight into phenomena. Notice that the entities in question may be complex (say, the whole cognitive system), or simply its building blocks. In a mechanistic version, this simply boils down to explaining all phenomena with a shared mechanism, or several shared (types of) mechanisms (Miłkowski 2016b). This strategy, *prima facie*, should at least alleviate the crisis.

One important proposal of this kind is developing complete theories of mechanisms and representations underlying all cognition, which are technically termed

“unified theories of cognition” in cognitive science (thanks to Newell, who was an influential proponent).<sup>4</sup> A lot of effort has been put into making unified theories of cognition prominent (Byrne 2012; Newell 1990). However, building such theories has remained relatively rare in cognitive science, which still proceeds with its divide-and-conquer strategies. Several factors could explain this fact.

First, building such unified architectures is a daunting task, and the learning curve is quite steep. The coverage of a unified theory, if it is supposed to be a grand theory, is extremely ambitious: It should explain all cognitive phenomena, full stop. But we do not even know how many such phenomena exist, not to mention how to explain them. One could make the goal more realistic by focusing on experimental tasks in cognitive science labs, and there are at least 3000 tasks to include (Carroll 1993)—still a steep number, but at least a tangible aim. Finally, one could aim for a unified but not a grand theory of cognition. Such a theory would not cover all of cognition but some selected aspects thereof. Realistically, even though there are now 84 actively developed cognitive architectures (Kotseruba and Tsotsos 2018), none of them is a grand theory that explains all phenomena known to cognitive science. But at least a fair number of their developers has the ambition to strive for completeness.

Second, while there are relatively clear evidential standards for individual models of cognitive phenomena, testing cognitive architectures remains somewhat elusive. Since these architectures are complex, it is extremely difficult, if not entirely impossible, to falsify them experimentally in their entirety without any remaining doubts. Instead, one could treat them as research programs that develop over time (Cooper 2007), but this still leaves a number of doubts. What should be a sign of real theoretical progress in cognitive architectures? Some propose to test architectures against sets of essential requirements (Anderson and Lebiere 2003), but such evaluations remain informal. Moreover, the requirements themselves are theory-laden, but it is exactly what unified theories of cognition are supposed to bring about: A new theory of cognition. Thus, it is unlikely that the requirements will remain stable; instead, they should undergo revision along with any progress in understanding cognition. But given the disarray of evaluation standards—whether it is the best possible coverage of multiple phenomena, predictive value for new phenomena, biological plausibility, variety of evidence, or empirical falsification—it is difficult to expect that one unified theory of cognition, posed in terms of a cognitive architecture, will become a dominant one with regards to all evaluation standards in cognitive science. As Laudan and Laudan (1989) argue, a theory becomes commonly accepted when it wins over its competition under any methodological standard of evaluation. But this is highly unlikely for any cognitive architecture right now.

<sup>4</sup> It was also argued that these theories are best understood as encompassing both mechanistic explanations and generalizations that should link artificial cognitive architectures with any kind of possible cognitive performance (Gervais 2020). While this might be true of some of them, the variety of approaches to building architectures makes such sweeping claims somewhat unsafe; there are both biologically plausible architectures focusing on neural ensembles (Eliasmith 2013) and ones that are stated in terms of symbolic production rules (Anderson 1983). It is unlikely that their developers subscribe to the same set of methodological principles.

Third, we lack a systematic analysis of foundations and conditions of the unification in cognitive science including, in particular, explanatory practices and the role of explanation in various fields of cognitive science, for example, in computational neuroscience or cognitive psychology, but also cognitive anthropology. The work on unified theories of cognition hitherto almost entirely ignores the cultural and social underpinnings of cognitive processing, even though these have become a major topic in prominent journals of cognitive science (Bender 2019; Bender et al. 2010). Moreover, it is doubtful that these unified theories avoid ad hoc assumptions: Because they are frequently presented as pieces of software, it remains unclear which parts of the software embody theoretical choices, and which are simply required for the software to work but have no representational function (Cooper and Guest 2014; Cooper and Shallice 1995).

The second kind of unification strategy may appeal to grand principles. Of course, one can embed a grand principle in a cognitive architecture (and their development is usually based on substantive theoretical principles), but some grand principles were proposed independently from this first strategy. For example, in the early work on cybernetics, it was believed that the deep principle of cognition was negative feedback (Sluckin 1954). Needless to say, this abstract principle, even if true, does not offer a local kind of understanding: It does not suffice to explain individual phenomena. This is probably the most critical problem for the strategy of global principles: Either one proposes one global principle, or a host of independent ones, which could then become unwieldy and disunified.

Global principles of cognition may yield unification to varying degrees. Although many proponents of predictive coding accept mechanistic assumptions and at least partially sketch some organization of the cognitive architecture (Beni 2018; Clark 2016; Gładziejewski 2019; Hohwy 2013), this theory offers global principles first of all (Friston 2010, 2019). These principles need not, however, cover all cognitive phenomena, and it is natural to understand dynamical models as offering some principled kinds of explanations in a unified manner (Kelso 1995). The unification offered by such principles should be understood to lie on a spectrum. Thus, even research heuristics could contribute to unifying cognitive science at least to some degree (Miłkowski and Nowakowski 2019).

There is one such principle that deserves special mention. What is called classical cognitive science, or simply cognitivism, relies on the belief that cognition requires computation over cognitive representations (Fodor 1975). This principle, just like the negative feedback principle, cannot yield an understanding of particular phenomena (Miłkowski 2018; Taatgen 2003). Yet, as imbued deeply in mainstream cognitive science, it offers a common language to talk about cognitive processing, which at least contributes to lessening possible theoretical diversity. But it is far too bland to address the issues noted by Newell (1973): The studies that incited his criticism mostly followed this principle to little avail. Obviously, one could respond that rejecting this principle is feasible, but implies a decrease of unity in cognitive science. This is probably what incited the study that aimed to show that cognitive science is disintegrated by showing that it has not transformed into a self-standing discipline (Núñez et al. 2019).

## 4 Unification and pluralism in practice

Cognitive science is clearly not unified right now; thus, unificatory accounts cannot be simply descriptively accurate. However, Newell's argument that without unification, cognitive science will cease to be cumulative, cataloging experimental effects instead, has not been rebutted by pluralists either. At the same time, developing systematic, unified, and parsimonious theories is simply difficult, and different researchers hold different evaluation standards. For defenders of grand principles, simplicity and the universal scope of a theory may be more important than its evidential support. Integration seems more related to evidential support and consistency instead. The defenders of the new mechanistic approach to explanation, for example, usually treat generality as an optional feature of scientific explanation (Craver 2008).

Thus, a defender of integration may proceed in a different fashion than a defender of unificatory strategies. This makes the debate over unification and integration (and explanatory pluralism) relevant to scientific practice.<sup>5</sup> We hope this much is clear in the papers that were selected for this special issue, as they all refer to choices of researchers that imply integrative or unificatory approaches.

As we have already mentioned, in the mainstream cognitive (neuro)science, both computational and mechanistic modes of explanation are massively adopted. Relationships between these modes are investigated in this special issue by Jens Harbecke (2020) as well as Elber-Dorozko and Shagrir (2019). Harbecke (2020) starts with the following observation: Despite the fact that computational models play an inestimable role in the understanding of the architecture of cognition (Churchland and Sejnowski 1994; Eliasmith and Anderson 2004; Thagard 2005), there is a question about the relationship among the levels that these models involve. So far, many credible frameworks aiming to sort it out have been proposed, and two of them became particularly prominent. The first is Marr's (1982) description of computational, algorithmic, and implementation levels of analysis. The second proposal is Craver's (2008) hierarchy of mechanistic levels. Harbecke argues that, taken separately, none of these approaches contributes to satisfactory explanations of cognitive phenomena. Instead, he is looking for a way of combining Marr's and Craver's proposals into a single mechanistic-computational account, which could perform an integrative function for cognitive science. Harbecke is grounding his standpoint in two ways, namely, in scientific practice and in metaphysical considerations. According to him, a satisfactory explanatory model of a cognitive architecture should integrate phenomena at all the levels (in the mechanistic sense), accounting simultaneously for computational processes involved in the relevant component parts.

The question of how computational explanations could be integrated with the mechanistic hierarchy is also pursued by Elber-Dorozko and Shagrir (2019). According to them, the integration problem arises from different relations that

<sup>5</sup> Perhaps this makes this debate more practical than arguments about the (im)possibility of reduction understood in traditional terms. It is striking, for example, that multiple realization is hardly even mentioned in the psychological discussion over unification in psychology (see, for example Gaj 2016; Sternberg 2005).

computational and mechanistic modes of explanation deploy. The former uses the homomorphism mapping relation, and as a consequence, is focused on the implementation of a computational system composed of abstract variables and functions in the physical. The latter, by contrast, employs the mereological “part-whole” relation and hence is focused on the causal organization of component parts and operations within a mechanism. Although using differing kinds of order relation, both computational and mechanistic explanations are hierarchical. This is clearly presented in examples of computational (Sutton and Barto 1998) and neurocognitive (Doya 2000) attempts to elucidate the phenomenon of reinforcement learning. Elber-Dorozko and Shagrir discuss two possible ways to deal with the integration problem. The first one is coherent with Piccinini and Craver’s (2011) proposal of considering computational explanations as mechanism sketches, and assumes that the mechanistic hierarchy builds at the corresponding levels of computational and implementational properties. The second way originates from the assumption of explanatory autonomy of computational models (e.g., Cummins 1983), and postulates the implementation relation of both hierarchies, namely computational and mechanistic. Finally, the authors claim that, although both ways could be successfully applied to the methodological investigation of reinforcement learning, in practice the understanding of computational explanations by modelers and other scientists reaches beyond mere sketches, and thus is more consistent with the second way.

Cognitive science studies are driven by grand approaches competing with each other, at least in certain respects. These approaches could be investigated through the lenses of various methodological frameworks, including Kuhn’s (1962) paradigms, Lakatos’ (1976) research programs, and Laudan’s (1977) research traditions. Miłkowski and Nowakowski (2019) investigate the contribution of embodied cognition to representational unification of cognitive phenomena, wherein unification is defined as involving four dimensions: Simplicity, generality and scope, non-monstrosity, and systematicity. To this end, Miłkowski and Nowakowski try to determine the methodological status of embodied cognition. While taking embodied cognition as a research program, in Lakatos’ sense, has gained popularity (Shapiro 2007), the authors claim that it is more fruitful to analyze it in terms of a research tradition in Laudan’s sense, since embodied cognition is composed of multiple, sometimes mutually contradictory, subtraditions, starting from representation-compatible ne empiricism (Barsalou 1999; Prinz 2004) and ending with non-representational radical embodiment (Chemero 2009). According to Miłkowski and Nowakowski, even if embodied cognition fails as a proposal of the grand unification of cognitive science, it shows that unification constitutes a notable virtue of research traditions.

While mechanistic explanations and computational methodology are at the core of cognitive science, they do not remain uncontested. Erdin (2020) attempts to evaluate one of the most recognizable applications of the mathematical theory of dynamical systems to cognition, i.e., radical embodiment as defended by Chemero (2009), which is contrasted with computationalism. To this end, Erdin thoroughly analyzes the requirements of continuity and autonomy of scientific progress adopted from Lakatos’ methodology, assuming that each research program is constituted by two following elements: A hard core, that cannot be falsified experimentally, and the positive heuristic providing the guidelines for constructing protective belts for the

hard core. Erdin's investigation reveals shortcomings in Chemero's non-representationalism. Moreover, he rejects a claim that the explanatory power of dynamicism and computationalism is equal. Erdin's paper not only assesses the radical embodiment or non-representational approach to cognitive science but also shows convincingly that Lakatos' methodology of scientific research programs constitutes a useful philosophical tool, or even a necessary constraint for various methodological discussions within cognitive science.

Within this special issue, the explanatory power of dynamicism is also discussed by Raja (2020). He claims that the dynamical and non-representational approach built on the concept of resonance could serve as the framework unifying neuroscience and psychology, namely disciplines interested, respectively, in neural and behavioral spatiotemporal dimensions of cognition. In a nutshell, the author's strategy is to show that resonance constitutes the process by which the neural dynamics are constrained by the same informational variable constraining the dynamics at the behavior scale. The concept of resonance between an internal cognitive state and an invariant structure of external information has been proposed for the first time within ecological psychology. Although Gibson (1966) used the concept of resonance in a metaphorical way, Raja defends the view that nowadays it constitutes the crucial component of a conceptual framework of radically embodied cognition (Chemero 2009). Raja substantiates the unification power of the resonance-based framework by bridging theories of behavioral dynamics (Warren et al. 2001) and neural reuse (Anderson 2014). Toward the end, the author discusses the applicability of the mechanistic description of the resonance. Although Raja indicates possible components of the alleged resonance mechanism merging brain, cognitive phenomena, and the organism-environment interplays, at the same time he argues that the constitutive nature of mechanisms (Craver 2008) is difficult to reconcile with the resonance-based framework.

Carls-Diamante (2019) also pursues the issue of dynamical explanations of cognition. She does this, however, in the context of octopus cognition, especially goal-directed motor control. While dynamicism and representationalism are frequently seen as competing theoretical frameworks (Van Gelder 1995), the author presents a case study of the phenomenon that calls for a pluralistic (hybrid) explanation. This phenomenon is known as fetching; it consists of grasping an object with a single arm, and approaching it point-to-point towards the mouth. Although representation-based models of fetching have been proposed (Flash and Sejnowski 2001), they turn out to be irrelevant—as Carls-Diamante claims—to the octopus case, due to the neuroanatomical and morphological constraints of this sea creature. Moreover, a credible model of octopus' fetching (Sumbre et al. 2006) involves some operation accounted for dynamically. In consequence, Carls-Diamante is defending explanatory pluralism in cognitive science. This is not, however, a general statement to the effect that we can freely or interchangeably appeal to dynamical (non-representational) and representational accounts of explanation. Instead, the author claims that at least some convergent phenomena (cognitive processes, tasks, etc.) require the non-redundant use of distinct explanatory accounts, wherein some component parts and operations within a mechanism are described dynamically, while others are described in a representational way. In more general terms, Carls-Diamante's

contribution clearly shows that comparative studies on cognitive processes in animals inhabiting radically differing niches could deliver meaningful insights into long-standing discussions within cognitive science and its philosophy. (It is also worth noting that this paper received the Werner Callebaut Prize from the International Society for the History, Philosophy and Social Studies of Biology in 2019.)

Many theoretical proposals which aspire to unify cognitive science and life sciences, to list only evolutionary psychology (Barkow et al. 1992) or a core knowledge approach (Spelke and Kinzler 2007), have a notion of innateness in their background. It is also arguably one of the unifying general principles in a number of cognitive explanations. Ritchie's (2020) paper contributes to a debate on the legitimacy of nativism in psychological explanations. The author undertakes a defense of the frequently challenged notion of nativism, which identifies inborn cognitive skills with "not learned" ones. He calls this view "the minimal conception of innateness." First of all, he claims that this minimal conception is indispensable in recognizing correctly a time-honored dispute between nativism and empiricism, which has gone beyond philosophy and has pervaded cognitive science. Next, he rejects pleas toward nativism recurring again and again. They involve a risk of overgeneralizations, i.e., recognizing some cognitive capacities as innate despite the existence of convincing evidence to the contrary, defining what is inborn through what is not, suppressing research that could lead to new explanations of developmental issues, and underestimating the role of learning processes in psychology. Next, Ritchie separates the minimal conception of innateness from primitivism, namely, the view that a given capacity is innate when it cannot be accounted for developmentally. The author also accounts for the problem of innateness as a natural kind. Last but not least—and essential for the aims of this Special Issue—Ritchie defends the claim that the minimal conception of innateness is well suited to consideration as a "bridge" between psychology and life sciences.

This Special Issue ends with Dupre's (2019) paper on the explanatory economy in one of the founding theories of cognitive science: Generative linguistics (Chomsky 1965). As the author notes, generative linguistics can be accused of excluding post hoc inconvenient data. If generativists could immunize their theory from any kind of counterevidence, the theory is unfalsifiable, and therefore unscientific. Dupre proposes a methodological account called "explanatory economy," which aims to reject the above objection through indicating (before inquiry) the data that can be excluded from the scope of a theory, and likewise the data which cannot. Without going into technical details, the author claims that an approach to corroborating or testing the explanatory power of a single theory in the light of data without any appeal to other theories is a flawed strategy, at least in the case of linguistics. This form of methodological individualism goes back to Popper (1959/2002), and as Dupre argues, it should be replaced by a kind of collectivism. According to the author, whenever the theorist decides to exclude some data from a scope of the theory, he or she should consider whether the data are accountable for by another non-contending theory. An open question that Dupre's contribution raises concerns the opportunity and fruitfulness of applying the methodology of the explanatory economy to other branches of cognitive science. The lesson from this paper is that in cognitive science, some theories could be considered complementary, and that they should not

be overgeneralized. This naturally leads to a certain form of explanatory pluralism. Generative linguistics is not a grand theory of all of cognition, but a theory of some aspects of linguistic capacities that human beings display.

We hope that the collection of papers constituting this Special Issue will contribute to a deeper understanding of the importance of unification and pluralism in cognitive science. As we can see, the issue of whether unification is to be preferred over integration is far from settled. Nonetheless, as papers in this issue also attest, unification does not boil down to reductive explanation. At the same time, the way unification or coordination is approached depends, obviously, on the general framework that one adopts in understanding satisfactory explanations.

**Acknowledgements** We wish to thank all reviewers for their work on all papers of the Special Issue and Catarina Dutilh Novaes, one of Editors-in-Chief of *Synthese* for her valuable and timely guidance in managing the issue. We also thank Przemysław Nowakowski, Piotr Litwin, and Witold Wachowski for their comments on the previous version of the manuscript.

**Funding** Our work on this Special Issue of *Synthese* was supported by Marcin Miłkowski's Grant "Cognitive Science in Search of Unity: Unification and Integration of Interdisciplinary Research" funded by the National Science Centre, Poland (2014/14/E/HS1/00803). See <http://cognitive.ifispan.pl> for the official project webpage.

## Compliance with ethical standards

**Conflict of interest** The authors declare that there is no conflict of interest.

## References

Aizawa, K. (2018). Multiple realization and multiple "ways" of realization: A progress report. *Studies in History and Philosophy of Science Part A*, 68, 3–9. <https://doi.org/10.1016/j.shpsa.2017.11.005>.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA: The MIT Press.

Anderson, J. R., & Lebiere, C. (2003). The Newell test for a theory of cognition. *The Behavioral and Brain Sciences*, 26(5), 587–601. <https://doi.org/10.1017/S0140525X0300013X>. (discussion 601–648).

Bangu, S. (2017). Scientific explanation and understanding: unificationism reconsidered. *European Journal for Philosophy of Science*, 7(1), 103–126. <https://doi.org/10.1007/s13194-016-0148-y>.

Barker, M. J. (2019). Eliminative pluralism and integrative alternatives: the case of SPECIES. *The British Journal for the Philosophy of Science*, 70(3), 657–681. <https://doi.org/10.1093/bjps/axz057>.

Barkow, J., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Biological approaches to mind and culture*. Oxford: Oxford University Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, 22(4), 577–609. (discussion 610–660).

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.

Bechtel, W., & Shagrir, O. (2015). The non-redundant contributions of Marr's three levels of analysis for explaining information-processing mechanisms. *Topics in Cognitive Science*, 7(2), 312–322. <https://doi.org/10.1111/tops.12141>.

Bender, A. (2019). The value of diversity in cognitive science. *Topics in Cognitive Science*, 11(4), 853–863. <https://doi.org/10.1111/tops.12464>.

Bender, A., Hutchins, E., & Medin, D. (2010). Anthropology in cognitive science. *Topics in Cognitive Science*, 2(3), 374–385. <https://doi.org/10.1111/j.1756-8765.2010.01082.x>.

Beni, M. D. (2018). The reward of unification: A realist reading of the predictive processing theory. *New Ideas in Psychology*, 48, 21–26. <https://doi.org/10.1016/j.newideapsych.2017.10.001>.

Bickle, J. (1998). *Psychoneural reduction: The new wave*. Cambridge, MA: The MIT Press.

Boone, W., & Piccinini, G. (2016). The cognitive neuroscience revolution. *Synthese*, 193(5), 1509–1534. <https://doi.org/10.1007/s11229-015-0783-4>.

Bouwel, J. V. (2014). Pluralists about pluralism? Different versions of explanatory pluralism in psychiatry. In M. C. Galavotti, D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, & M. Weber (Eds.), *New directions in the philosophy of science* (pp. 105–119). Dordrecht: Springer.

Byrne, M. D. (2012). Unified theories of cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(4), 431–438. <https://doi.org/10.1002/wcs.1180>.

Carls-Diamante, S. (2019). Make up your mind: octopus cognition and hybrid explanations. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-019-02102-2>.

Carnap, R. (1928). *The logical structure of the world; Pseudoproblems in philosophy*. (R. A. George, Trans.). Berkeley: University of California Press, 1967.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Churchland, P. S., & Sejnowski, T. J. (1994). *The computational brain*. Cambridge, MA: The MIT Press.

Clark, A. (2016). *Surfing uncertainty: Prediction, action and the embodied mind*. Oxford: Oxford University Press.

Colombo, M., & Wright, C. (2017). Explanatory pluralism: An unrewarding prediction error for free energy theorists. *Brain and Cognition*, 112, 3–12. <https://doi.org/10.1016/j.bandc.2016.02.003>.

Cooper, R. P. (2007). The role of falsification in the development of cognitive architectures: Insights from a Lakatosian analysis. *Cognitive Science*, 31(3), 509–533. <https://doi.org/10.1080/15326900701326592>.

Cooper, R. P., & Guest, O. (2014). Implementations are not specifications: Specification, replication and experimentation in computational cognitive modeling. *Cognitive Systems Research*, 27, 42–49. <https://doi.org/10.1016/j.cogsys.2013.05.001>.

Cooper, R. P., & Shallice, T. (1995). Soar and the case for unified theories of cognition. *Cognition*, 55(2), 115–149. [https://doi.org/10.1016/0010-0277\(94\)00644-Z](https://doi.org/10.1016/0010-0277(94)00644-Z).

Craver, C. F. (2008). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, 72(20), 741–765.

Cummins, R. (1983). *The nature of psychological explanation*. Cambridge, MA: The MIT Press.

Cummins, R. (2000). “How does it work” versus “what are the laws?”: Two conceptions of psychological explanation. In F. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 117–145). Cambridge, MA: MIT Press.

Dale, R. (2008). The possibility of a pluralist cognitive science. *Journal of Experimental & Theoretical Artificial Intelligence*, 20(3), 155–179. <https://doi.org/10.1080/09528130802319078>.

Dale, R., Dietrich, E., & Chemero, A. (2009). Explanatory pluralism in cognitive science. *Cognitive Science*, 33(5), 739–742. <https://doi.org/10.1111/j.1551-6709.2009.01042.x>.

Darden, L., & Maull, N. (1977). Interfield theories. *Philosophy of Science*, 33, 43–64. <https://doi.org/10.2307/187099>.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10(6), 732–739. [https://doi.org/10.1016/S0959-4388\(00\)00153-7](https://doi.org/10.1016/S0959-4388(00)00153-7).

Dupre, G. (2019). Linguistics and the explanatory economy. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-019-02290-x>.

Elber-Dorozko, L., & Shagrir, O. (2019). Integrating computation into the mechanistic hierarchy in the cognitive and neural sciences. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-019-02230-9>.

Eliasmith, C. (2013). *How to build the brain: A neural architecture for biological cognition*. New York: Oxford University Press.

Eliasmith, C., & Anderson, C. H. (2004). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. Cambridge, MA: The MIT Press.

Erdin, H. O. (2020). Appraisal of certain methodologies in cognitive science based on Lakatos's methodology of scientific research programmes. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism, ed. by M. Milkowski & M. Hohol). <https://doi.org/10.1007/s11229-020-02612-4>.

Flash, T., & Sejnowski, T. J. (2001). Computational approaches to motor control. *Current Opinion in Neurobiology*, 11(6), 655–662. [https://doi.org/10.1016/S0959-4388\(01\)00265-3](https://doi.org/10.1016/S0959-4388(01)00265-3).

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28(2), 97–115. <https://doi.org/10.1007/BF00485230>.

Fodor, J. A. (1975). *The language of thought* (1st ed.). New York: Thomas Y. Crowell Company.

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, 71(1), 5–19. <https://doi.org/10.2307/2024924>.

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>.

Friston, K. J. (2019). A free energy principle for a particular physics. Retrieved December 29, 2019 from arXiv:1906.10184 [q-bio]. <http://arxiv.org/abs/1906.10184>.

Gaj, N. (2016). *Unity and fragmentation in psychology: The philosophical and methodological roots of the discipline*. Basingstoke: Taylor & Francis.

Gentner, D. (2019). Cognitive science is and should be pluralistic. *Topics in Cognitive Science*, 11(4), 884–891. <https://doi.org/10.1111/tops.12459>.

Gervais, R. (2020). The multiplicity of explanation in cognitive science. *Foundations of Science*. <https://doi.org/10.1007/s10699-020-09653-5>.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gladziejewski, P. (2019). Mechanistic unity of the predictive mind. *Theory & Psychology*, 29(5), 657–675. <https://doi.org/10.1177/0959354319866258>.

Harbecke, J. (2020). The methodological role of mechanistic-computational models in cognitive science. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism, ed. by M. Milkowski & M. Hohol). <https://doi.org/10.1007/s11229-020-02568-5>.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135–175.

Hensel, W. M. (2013). On reduction and interfield integration in neuroscience. In M. Milkowski & K. Talmont-Kamiński (Eds.), *Regarding the mind, naturally: Naturalist approaches to the sciences of the mental* (pp. 167–181). Newcastle upon Tyne: Cambridge Scholars Publishing.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hooker, C. A. (1981a). Towards a general theory of reduction. Part I: Historical and scientific setting. *Dialogue*, 20(01), 38–59. <https://doi.org/10.1017/S0012217300023088>.

Hooker, C. A. (1981b). Towards a general theory of reduction. Part II: Identity in reduction. *Dialogue*, 20(02), 201–236. <https://doi.org/10.1017/S0012217300023301>.

Hooker, C. A. (1981c). Towards a general theory of reduction. Part III: Cross-categorical reduction. *Dialogue*, 20(03), 496–529. <https://doi.org/10.1017/S0012217300023593>.

Horst, S. (2016). *Cognitive pluralism*. Cambridge, MA: The MIT Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: The MIT Press.

Kincaid, H. (1990). Molecular biology and the unity of science. *Philosophy of Science*, 57(4), 575–593. <https://doi.org/10.1086/289580>.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, 48(4), 507–531. <https://doi.org/10.1086/289019>.

Kotseruba, I., & Tsotsos, J. K. (2018). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-018-9646-y>.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1976). Falsification and the methodology of scientific research programmes. In S. Harding (Ed.), *Can theories be refuted* (pp. 205–259). Dordrecht: Springer.

Laudan, L. (1977). *Progress and its problems*. Berkeley: University of California Press.

Laudan, R., & Laudan, L. (1989). Dominance and the disunity of method: Solving the problems of innovation and consensus. *Philosophy of Science*, 56(2), 221–237. <https://doi.org/10.1086/289484>.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1–25. <https://doi.org/10.1086/392759>.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

McCauley, R. N., & Bechtel, W. (2001). Explanatory pluralism and heuristic identity theory. *Theory & Psychology*, 11(6), 736–760. <https://doi.org/10.1177/0959354301116002>.

McClamrock, R. (1991). Marr's three levels: A re-evaluation. *Minds and Machines*, 1(2), 185–196. <https://doi.org/10.1007/BF00361036>.

McPhetres, J., Albayrak-Aydemir, N., Barbosa Mendes, A., Chow, E. C., Gonzalez-Marquez, P., Loukras, E., et al. (2020). *A decade of theory as reflected in psychological science (2009–2019)* (preprint). PsyArXiv. <https://doi.org/10.31234/osf.io/hs5nx>.

Miłkowski, M. (2013). *Explaining the computational mind*. Cambridge, MA: The MIT Press.

Miłkowski, M. (2016a). Integrating cognitive (neuro)science using mechanisms. *Avant: Journal of Philosophical-Interdisciplinary Vanguard*, VI(2), 45–67. <https://doi.org/10.26913/70202016.0112.0003>.

Miłkowski, M. (2016b). Unification strategies in cognitive science. *Studies in Logic, Grammar and Rhetoric*, 48(1), 13–33. <https://doi.org/10.1515/slgr-2016-0053>.

Miłkowski, M. (2018). From computer metaphor to computational modeling: The evolution of computationalism. *Minds and Machines*, 28(3), 515–541. <https://doi.org/10.1007/s11023-018-9468-3>.

Miłkowski, M., Clowes, R. W., Rucińska, Z., Przegalińska, A., Zawidzki, T., Gies, A., et al. (2018). From wide cognition to mechanisms: A silent revolution. *Frontiers in Psychology*, 9, 2393. <https://doi.org/10.3389/fpsyg.2018.02393>.

Miłkowski, M., Hohol, M., & Nowakowski, P. (2019). Mechanisms in psychology: The road towards unity? *Theory & Psychology*, 29(5), 567–578. <https://doi.org/10.1177/0959354319875218>.

Miłkowski, M., & Nowakowski, P. (2019). Representational unification in cognitive science: Is embodied cognition a unifying perspective? *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism, ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-019-02445-w>.

Nathan, M. J. (2017). Unificatory explanation. *The British Journal for the Philosophy of Science*, 68(1), 163–186. <https://doi.org/10.1093/bjps/axv022>.

Neurath, O. (1937). Unified science and its encyclopaedia. *Philosophy of Science*, 4(2), 265–277. <https://doi.org/10.1086/286457>.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Núñez, R., Allen, M., Gao, R., Rigoli, C. M., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behaviour*, 3(8), 782–791. <https://doi.org/10.1038/s41562-019-0626-2>.

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26(5), 1596–1618. <https://doi.org/10.3758/s13423-019-01645-2>.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Concepts, theories, and the mind-body problem. Minnesota studies in the philosophy of science* (Vol. 1–47, pp. 3–36). Minneapolis: Minnesota University Press.

Peacocke, C. (1986). Explanation in computational psychology: Language, perception and level 1.5. *Mind and Language*, 1(2), 101–123.

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford: Oxford University Press.

Piccinini, G., & Craver, C. F. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183(3), 283–311. <https://doi.org/10.1007/s11229-011-9898-4>.

Poggio, T. (2012). The levels of understanding framework, revised. *Perception*, 41(9), 1017–1023. <https://doi.org/10.1068/p7299>.

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford, NY: Oxford University Press.

Popper, K. R. (2002). *The logic of scientific discovery*. London: Routledge.

Potochnik, A. (2011). A Neurathian conception of the unity of science. *Erkenntnis*, 74(3), 305–319. <https://doi.org/10.1007/s10670-010-9228-0>.

Prinz, J. J. (2004). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: The MIT Press.

Putnam, H. (1967). Psychological predicates. In W. H. Capitan & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: University of Pittsburgh Press.

Raja, V. (2019). J. J. Gibson's most radical idea: The development of a new law-based psychology. *Theory & Psychology*, 29(6), 789–806. <https://doi.org/10.1177/0959354319855929>.

Raja, V. (2020). Resonance and radical embodiment. *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism, ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-020-02610-6>.

Richardson, R. C. (2009). Multiple realization and methodological pluralism. *Synthese*, 167(3), 473–492. <https://doi.org/10.1007/s11229-008-9387-6>.

Ritchie, J. B. (2020). What's wrong with the minimal conception of innateness in cognitive science? *Synthese* (SI: Explanations in cognitive science: Unification vs pluralism, ed. by M. Miłkowski & M. Hohol). <https://doi.org/10.1007/s11229-020-02543-0>.

Shagrir, O. (2010). Marr on computational-level theories. *Philosophy of Science*, 77(4), 477–500. <https://doi.org/10.1086/656005>.

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy Compass*, 2(2), 338–346. <https://doi.org/10.1111/j.1747-9991.2007.00064.x>.

Shapiro, L. (2019). A tale of two explanatory styles in cognitive psychology. *Theory & Psychology*, 29(5), 719–735. <https://doi.org/10.1177/0959354319866921>.

Sluckin, W. (1954). *Minds and machines*. Harmondsworth Middlesex: Penguin Books.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, 10(1), 89–96. <https://doi.org/10.1111/j.1467-7687.2007.00569.x>.

Sternberg, R. J. (2005). *Unity in psychology: Possibility or pipedream?*. Washington, DC: American Psychological Association.

Sullivan, J. A. (2009). The multiplicity of experimental protocols: A challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese*, 167(3), 511–539. <https://doi.org/10.1007/s11229-008-9389-4>.

Sumbre, G., Fiorito, G., Flash, T., & Hochner, B. (2006). Octopuses use a human-like strategy to control precise point-to-point arm movements. *Current Biology*, 16(8), 767–772. <https://doi.org/10.1016/j.cub.2006.02.069>.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: The MIT Press.

Taatgen, N. A. (2003). Popperizing the Newell test. *Behavioral and Brain Sciences*, 26(5), 621–622. <https://doi.org/10.1017/S0140525X03390132>.

Taylor, H., & Vickers, P. (2017). Conceptual fragmentation and the rise of eliminativism. *European Journal for Philosophy of Science*, 7(1), 17–40. <https://doi.org/10.1007/s13194-016-0136-2>.

Thagard, P. (2005). *Mind: Introduction to cognitive science*. Cambridge, MA: The MIT 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>.

Walmsley, J. (2008). Explanation in dynamical cognitive science. *Minds and Machines*, 18(3), 331–348. <https://doi.org/10.1007/s11023-008-9103-9>.

Warren, J., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, 4(2), 213–216. <https://doi.org/10.1038/84054>.

Weiskopf, D. A. (2011). The functional unity of special science kinds. *The British Journal for the Philosophy of Science*, 62, 233–258. <https://doi.org/10.1093/bjps/axq026>.

Wright, C. D. (2007). Is psychological explanation becoming extinct? In M. Schouten & H. L. de Jong (Eds.), *The matter of the mind: Philosophical essays on psychology, neuroscience, and reduction* (pp. 249–274). Hoboken: Wiley-Blackwell.

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