Expectation, Representation, and Enactivism

Jean-Charles Pelland (jean.pelland@uib.no)
Department of Psychosocial Science, Christies gate 12 (4. etasje)
Bergen 5015, Norway

Abstract

This paper presents a challenge to enactivist approaches to cognition (e.g. Ward, D., Silverman, D. & Villalobos, M. 2017) that is based on the theoretical commitments behind forms of looking time studies that have been extensively used to probe into the cognitive abilities of infants and nonhuman animals. I briefly summarize the Violation of Expectation (VoE) paradigm (Ginnobili & Olmos 2021) to illustrate why such methods might pose a problem for enactivists and their conception of cognition as a largely representation-free dynamic coupling between organism and environment. I argue that despite the lack of clarity regarding how the notion of expectation should be applied to the minds of neonates and nonhuman animals, there is an inherently representational aspect to expectation, given that it embodies satisfaction conditions. The challenge is, then: given that many forms of enactivism seem to reject the notion of representation as it is used in the VoE literature, how can enactivists make sense of data and results obtained using such research methods?

Keywords: Violation of expectation; looking time paradigm; habituation-dishabituation; developmental psychology; ethology; philosophy of mind; enactivism; representation; cognitivism; situated cognition; 4E cognition; Embodied cognition

Introduction

Thanks to methodological and technological advances across its disciplines, cognitive science has grown by leaps and bounds since its birth in the mid twentieth century. Nowadays, it is safe to say that the sciences of the mind have made tremendous progress in identifying some of the systems and processes responsible for our mental lives, including their neural correlates, developmental milestones and possible evolutionary precursors. Surprisingly, despite it being a comparatively young branch of scientific investigation, cognitive science has already gone through a number of important changes throughout its short history, including the rise in popularity of connectionist models that were initially rejected by influential figures in the field (e.g. Fodor & Pylyshyn 1988).

Throughout much of its short-yet-productive history, cognitive science essentially relied on a computational theory of mind where the mind is often characterized as a computer that processes representations stored in the brain.

While this computational approach to the mind and its reliance on the notion of representation directed research in the many disciplines that make up cognitive science for decades, issues related to the importance of our body, our environment, and the material support that populate our mental lives (among others) have highlighted potential limits in seeing the mind as a brainbound computer that operates on representations stored in neuronal format.

In the past few decades, revisionist approaches to cognitive science that reject parts or all of the classical, computationalist view of the mind have blossomed, often embracing a version or other of what is often referred to as *4E Cognition* (Newen, A., de Bruin, L., & Gallagher, S. 2018). While the revisionist schools embrace a wide variety of theoretical commitments, a shared thesis is that of framing the mind as embodied, enactive, embedded, and/or extended and distributed. Nowadays, such alternative conceptions of the mind have mushroomed in many fields, often yielding productive frameworks with which to direct and interpret novel experimental practices (see e.g. part III of Robbins & Aydede 2009). For some, this heralds a second major change in cognitive science, its abandonment of the computational approach to the mind and the accompanying heavy reliance on the notion of 'representation'.

In this short paper, I wish to explore to which extent one particular brand of revisionist cognitive science – namely, enactivism (Varela, F., Thompson, E., & Rosch, E. 1991) – is well-equipped to make sense of a particularly productive and representation-heavy experimental paradigm that has been extensively used in studying the mental lives of preverbal infants and nonhuman animals – namely, Violation of Expectation (VoE). More precisely, I wish to argue that VoE poses a challenge to varieties of enactivism: on the one hand, given that VoE is such a widespread method of investigating what are sometimes called 'basic' minds, any form of revisionism about the mind needs to be able to accommodate its widespread use and make sense of its results under the revisionist approach's theoretical toolkit. On the other hand, despite there being considerable conceptual ambiguity regarding the principles and empirical commitments driving the VoE paradigm, its constitutive reliance on a strong notion of representation seems to make it incompatible with various versions of enactivism, especially the self-proclaimed radical brand (e.g. Hutto & Myin 2013), and their accompanying widespread rejection of representations in so-called 'basic' minds.

The paper unfolds as follows: in section one I sketch the main lines of the looking time method and its incarnation in VoE. Then, in section two I summarize some of the main

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

motivations behind the enactivist approach to the mind, focusing on reasons behind enactivist rejection of representations. Section three sets up the challenge that VoE represents for enactivism, while section four explores an answer to this challenge in the form of re-interpreting the VoE literature in enactivist-friendly terms, and finds it wanting, due to the fact that expectations come with satisfaction conditions.

Violation of Expectation

A difficult hurdle when probing infant and animal minds is that they cannot verbally report on the causes of their behavior. Thankfully, methodological breakthroughs that exploit objective, observable features of animal and infant subjects appear to have opened the door into their cognitive repertoire. Techniques such as looking time paradigms, which detect changes in cognitive states by measuring the amount of time a subject's gaze is directed to a particular scene, as well as manual search paradigms, where subjects are allowed to search for objects they have seen researchers interact with, have allowed scientists to peer into the mental lives of subjects that were out of reach when Piaget first started looking into the minds of infants.[1]

In recent years, the VoE paradigm has become the most widespread methodological tool used to investigate the representational repertoire of infants and animals (Carey 2009; Ginnobili & Olmos 2021). It has been used in thousands of studies targeting the cognitive abilities of infants as well as a wide variety of animal species, testing for the presence of systems and abilities ranging from tracking object permanence (e.g. Baillargeon, R., Spelke, E. S., & Wasserman, S. 1985) to what is often regarded as the rudiments of moral cognition (e.g. Onishi & Baillargeon 2005).[2] VoE is a specific application of looking time studies, which were initially geared towards detecting perceptual skills in infants (e.g. whether or not they could see colors). In VoE, researchers exploit subjects' tendency to look longer[3] at stimuli that presents novel or unexpected scenes. The underlying rationale is that we can detect the presence of cognitive systems geared towards certain aspects of the world by observing whether or not manipulations of this aspect of the world incur longer looking times in the exposed subjects. In other words, subjects would not stare longer at scenes that violate the way they expect the world to behave unless they have an ability to detect this aspect of the world and expectations for its behaviour.

Many VoE experiments involve little – if any – training, simply exposing subjects to unexpected scenes and reading their reaction. Other VoE paradigms involve repeated exposure to perceptually similar scenes. In such habituation paradigms, researchers exploit the fact that subjects eventually lose interest (i.e. habituate) when repeatedly exposed to stimuli displaying the same properties. Typically, after a series of trials that fail to present novel features, subjects' gaze starts to wander away from the stimulus, which appears to be presenting the same content over and over again.

As an indicator of habituation, researchers typically set a value that is proportional to the looking time for the first stimulus. For example, if subjects first looked at the stimuli for four seconds, when they start to look at it for 2 seconds before looking away, it is taken as a sign that the stimuli are no longer presenting anything new or interesting to the subject.[4] Once subjects have habituated, researchers expose them to stimuli that are meant to regain their interest, either by showing them possible modifications of previously displayed stimuli, or modifications that, at least to an adult, would be impossible or unexpected. The important behavioral datum here is whether looking time varies between habituation stimuli and novel stimuli.

Here the idea is that if repeated exposure to the same stimulus eventually bores subjects, when a change in certain features of the stimulus is able to get looking time back up, it might be a sign that the subjects' reaction is based on a cognitive system that is sensitive to those changed features. In particular, stimuli that display impossible outcomes would elicit longer looking times given that they violate the expected outcome as determined by the constraints on the cognitive systems responsible for the subject's reaction to the stimuli.

As an example, consider the pioneering work of Renée Baillargeon and her colleagues (Baillargeon et al. 1985), who were the first to test the presence of representations of object permanence using a VoE paradigm that involved playing a trick on the infant subjects. In this study, four-month old infants were shown a platform that rotated back and forth 180 degrees, like a drawbridge that falls in two directions. Once habituated to the motion of this platform, infants were shown an object introduced directly in the path of its downward motion. Normally, this would mean that the platform could not complete a full 180-degree rotation, since the object had been placed in a way such as to prevent this motion.

In some trials, infants were shown the expected outcome, and the platform halted when it touched the occluding object before going back in the other direction. In other trials, an impossible outcome was shown to the infants: without the infants noticing, researchers surreptitiously removed the occluding object from where infants had seen it placed. Thus, instead of stopping on the object, the platform continued its downward motion as it had done when there was no occluding object. Infants looked longer at the impossible outcome, thus suggesting that they expect objects to be solid

[1] Carey 2009 offers an extensive discussion of how these paradigms extended past Piaget's work.

[2] See Ginnobili & Olmos 2021 for an exhaustive list.

[3] While the majority of VoE studies involve measuring gaze duration, other measures including heartrate and pupil dilation are also sometimes used.

[4] In some experiments, variations in looking time are relatively small, for example, around two seconds (e.g. Xu & Carey 1996), but in all cases, they are reliable and replicable differences in behavior.

to the extent that one cannot pass through the other. Importantly, the fact that infants looked longer at the impossible outcome suggests that the cognitive system underlying their behavior in this setup can keep track of objects – in this case, the object that was supposed to prevent the platform from moving down – even when they are not being directly perceived. The standard interpretation of this study has been that infants are equipped with an innate representation of object permanence and that their longer looking time is a sign that the expectation triggered by the tokening of an object file associated with the occluding object has not been satisfied.

As this example shows, a great advantage of the VoE paradigm over previous methods is that it allows researchers to probe deep into the cognitive inventory of subjects that cannot produce introspective reports – especially, preverbal infants and animals – using a non-invasive behavioral reading. With VoE, we can piece together the domains and patterns of activity of cognitive systems at play in infant and animal minds during various tasks by measuring various aspects of their outwardly observable behaviour as they navigate their way through possible and impossible scenarios. If subjects react in ways that display the presence of underlying cognitive abilities with the same features as those of human adults, VoE may even point to the presence of ontogenetic or phylogenetic continuity regarding cognitive systems underlying certain rational or perceptual tasks.

While I have been careful here to remain as neutral as I could about whether or not explaining subject behaviour in such studies requires the notion of representation, speaking for the most part in terms of cognitive abilities and systems, it is important to mention that virtually all the literature couches the theoretical commitments behind VoE in representational terms (see e.g. Carey 2009; Ginnobilly & Olmos 2021). While space constraints prevent me from offering a long discussion of what representations might be, the reasoning here is intuitive enough: if a subject is able to respond to a variation in the environment, that means that they have something like a mental picture, something inside their head that is *about* that part of their environment – in other words, a representation. Since VoE shows that subjects respond to variations of specific aspects of the environment, the intuitive conception of representation at play here suggests that infants have representations concerning those aspects of the environment.

The fact that VoE is universally framed in representational terms is not surprising, since representations are at the heart of mainstream cognitive science, which for the most part is concerned more with gathering data than clarifying foundational notions. And yet, despite the importance of this notion for empirical practices in many branches of cognitive science, some revisionist schools of thought that do focus on such foundational notions argue that the notion of 'representation' should be abandoned, at least in part. In the next section, I sketch some of the thinking behind this revisionism towards representations as it is framed in some enactivist approaches to the mind

Broad church enactivism vs. representations

Despite there being significant diversity in both the motivating lines behind branches of enactivism and how enactivist theses are fleshed out,[5] most enactivist schools still display strong commitment to its roots, which can broadly be described as an attempt to grasp how organisms and their environments mutually determine and constrain each other. One consequence of this conceptual foundation is to see human thinking as simply one way in which an organism adapts to its environment and in so doing gives meaning to it. Another is to see many cognitive processes – especially those involved in basic perception and action – as constitutively distributed across the brain, the body and the environment.

With this rejection of the brainboundedness of cognition comes many questions, one of which is the status of representations: if representations are typically thought to reside in our heads, and if enactivism doesn't see cognition as taking place entirely in our heads, where does that leave representations? For many enactivists, getting rid of representations for at least some forms of cognition is part of the answer to this question.

Indeed, on top of the many positive and original enactivist ideas that seem to offer novel insight on difficult problems facing cognitive science,[6] of the main unifying threads that runs across enactivist schools of thought is their rejection, to varying degrees, of classical cognitive science's reliance on the notion of representation.

In a nutshell, one of the main problems most enactivists see with loose attribution of representations to basic building blocks of cognition like neurons and firing rates of neuronal assemblies is that these conflate representation with co-variation. A common illustration of this issue is that of a thermostat's behaviour: while we would agree that a thermostat's state varies in response to features of the environment, most of us would be uncomfortable in saying that the thermostat's states are representational. We would probably be reluctant to use such language because of an important aspect of representations, which is that they are

[5] It is customary to split enactivism up into three main schools, based mainly on which aspect of enactivist thought they dedicate their attention to: autopoietic, sensorimotor, and radical enactive cognition, or REC. While these all share a commitment to ideas present in Dynamic Systems Theory, including that of dynamical coupling, according to which cognition *emerges* from sensorimotor interaction of an organism with its environment, they disagree on the extent to which this implies ridding cognitive science of representations. Sensorimotor enactivism can be seen as the least hostile towards representations, while radical enactivism, as its name suggests, offers the most radically eliminativist attitude towards representations. See chapter two of Hutto & Myin 2013, as well as Ward et al. 2017 for more on the differences between the main schools of enactivism. Where the battle lines are drawn does not affect the overall argument presented here, so we will not dive further into this issue.

[6] See e.g. Gallagher 2017 for such a positive enactivist spin on many problems.

supposed to be *about* something. And it doesn't look like a thermostat's behaviour should be *about* anything.

Building on these intuitions, the enactivist reflection goes something like this: if we are uncomfortable in attributing representational status to a thermostat's co-variation with its environment, why would it be acceptable to do so for the building blocks of brainbound cognition, like firing patterns of neurons, for example? Why would it be acceptable to say that the type of co-variation displayed by neuronal assemblies is representational, while that of the thermometer is not? For enactivists, neither case is one of representation, since neither case involves the possibility of *mis*representation.

An assumption behind this reasoning seems to be that one of the main criteria something should meet in order to be considered to be a representation is the possibility that it is misused, or fails to obtain. For example, if we consider the word 'dog' to be about dogs, it is because we have come to associate its use to point to specific thing in the world – i.e. dogs. We say the content of the word 'dog' is the things in the world the word applies to – and only those things. Implicit here is the assumption that such a word could be applied to things that aren't dogs, and such an application would be a mistake – a misrepresentation.[7]

Given that representations come with satisfaction conditions, however, how can we use this word to apply to things that appear to be mere cases of co-variation? For example, how can we consider a neuronal firing pattern to represent anything, when all the neurons do is fire when they are presented with certain inputs? How can such purely mechanical, cause-consequence entities ever misrepresent? For some, this problem means representations shouldn't even be considered as a possibility for talking about cognition that is not couched in complex social agreements, such as the symbols of language and mathematics (e.g. Di Paolo, E. A., Buhrmann, T., & Barandiaran, X. E. 2017). This, in turn, means that cognition would not be grounded in representations.

So, if cognition isn't grounded in representations, what is it based in?

Instead of basing the study of cognition on the notion representation, enactivists view cognition as emerging from an organism's embodied dynamic sensorimotor interaction with its environment. To get an idea of how this could rival – and even replace – classical approaches to cognition, one need only look at one of the main sources of inspiration for enactivist thought, namely, connectionist architecture (e.g. Clark 1993).

As the advent of artificial intelligence and deep learning architectures has taught us, connectionist networks can display behaviour that is adapted to its environment, to the point of mimicking intelligent actions. And yet, the nodes within such network are never said to represent anything. Rather, the intelligent-looking behavior is said to *emerge* from the dynamic interaction of the network itself with its environment. This notion of intelligent behaviour emerging from the dynamic interaction of its constituents is one of the main doctrines shared across varieties of enactivism. The idea is to avoid talk of discrete representations to describe cognition and other aspects of our mental lives as much as possible and to replace it with the global, organism-centric vocabulary of dynamic systems theory. Here, cognition emerges from an organism's attempt to make sense of its interactions with its environment, to which it is coupled. This sensemaking applies as much to human cognition as it does to the behaviour of organisms like bacteria: both cases involve an autopoietic system that is the product of reciprocal processes involving the organism and its environment.

In sum, while classical approaches to cognition locate meaning in the organism's internal representations because these are supposed to be about something external to them, enactivists see meaning as emerging from a dynamic interaction between organism and environment.

On top of framing cognitive structures in terms of an organism's patterns of sensorimotor interaction with its environment, a related thread running through enactivist thought is to see perception as something organisms *do*, actively, rather than something that happens to them as part of a passive process of receiving information. Perception is seen as an active part of an organism's attempt to maintain itself and make sense of its relation to its environment. This is one reason why enactivists share many of the doctrines of Gibson's (1979) ecological approach to perception, where we perceive not features of things, but rather affordances, which can be summed up as actions we could do with things.

This helps explain why details about physical implementation of organisms – their bodies – matter to enactivists: the way in which an organism will interact with its environment is dictated by its physical implementation, and these interactions and their specific dynamics will in turn give rise to or 'bring forth' specific structures that characterize the organism's cognitive regime.

It is easy to see how enactivism's main conceptual motivations set it apart from traditional, representation-friendly approaches to cognition. This alternative perspective on cognition has predictably attracted considerable philosophical disagreement, but also, some concrete innovations. Indeed, like other revisionist approaches to cognitive science, enactivism has been applied in a wide variety of fields, from robotics (Clark 2008; Egbert & Barandiaran 2022) to archeology (Malafouris 2013) to physiology (Colditz 2020) and virtual reality (Rolla, G., Vasconcelos, G., & Figueiredo, N. M., 2022), to name a few.

VoE: A challenge for enactivism

While this presentation of the main lines of enactivism had to remain criminally short due to space constraints, it nonetheless should suffice to highlight how VoE may represent a challenge for enactivism's aversion to representations: can enactivism make sense of the results

[7] More elaborate presentation of this issue can be found in Hutto & Myin 2013 and Ramsey 2007.

obtained using VoE without appealing to the notion of representation?

Note that the challenge here is not the 'usual' one of finding an enactivist-friendly explanation for what is sometimes called representation-hungry cognition (see e.g. Clark & Toribio 1994; Zahnoun 2021). The hunger-based challenge is usually associated with 'high-level' cognition like imagination, reasoning, and abstract thought, which can proceed offline when compared to the online, embodied cognition involved in actions like catching a ball. The idea is that since enactivism focuses on the active interaction of an organism with its environment, it would be ill-equipped to deal with abstract thoughts that seem detached from our environment, such as numerical cognition.[8] Here, the challenge applies to what appear to be the building blocks of our cognitive regimes, given that much of VoE targets infants that are a few months old and animals that may not be capable of abstract thought. In this sense, this challenge threatens to undermine more of the enactivist repertoire than those based on representation-hungry cognition.

Is enactivist re-interpretation an option?

At first glance, it looks like enactivism is simply not equipped to integrate VoE. After all, the VoE literature includes thousands of studies using representation-speak. How can enactivists deny the value of such an important body of empirical research – especially if this denial is based on purely conceptual grounds?

Of course, the vocabulary commonly used in the existing literature should not be a determinant factor in deciding whether a scientific practice is based on solid conceptual grounds. After all, if this were to be adopted as a general principle it would effectively stifle innovation and conceptual change. But more importantly, enactivists have proposed to replace talk of representation with vocabulary that is more in line with their perspective in many domains,[9] so there is no reason in principle to claim that this reinterpretation strategy could not work for VoE. Here, the strategy would be to claim that while it is common to talk of representations in VoE, the term is used there in a unreflective manner, and there is a way to re-interpret its widespread use in the empirical literature that is consistent with enactivist principles. Perhaps the R-word here is just, as Gallagher put it, "a place-holder for an explanation that needs to be cast in dynamical terms of an embodied, environmentally embedded, and enactive model" (Gallagher 2008, p. 365).

At second glance, there does indeed appear to be wiggle room for enactivism to take grip here after all: taking a closer look at the assumptions underlying VoE reveals considerable ambiguity regarding the central notion of expectation, how these are supposed to be violated, and the type of cognitive machinery required for this to manifest itself in observable, measurable ways of the sort exploited by the VoE paradigm. That is, while the myriad incarnations of VoE do share a general perspective where subjects are seen as possessing expectations that can be violated, the exact nature of these expectations is not fleshed out in any significant way across the literature, which suggests that the theoretical commitments of this paradigm have not received extensive conceptual analysis of the sort that would make enactivist re-interpretation less likely.

Indeed, as Ginnobilly & Olmos (2021) point out in one of the rare articles that focus on the conceptual foundations of the VoE paradigm, despite there being a trend to interpret this method as indicating the presence of high-level cognitive abilities, VoE is actually 'neutral' with respect to the nature of the expectation itself. As they point out, this interpretative neutrality is common in scientific practice across disciplines and need not be considered a flaw in this specific paradigm.

So, not only does it look like there is room for enactivist re-interpretation of expectations here, but such an enactivist re-reading of VoE could actually help anchor this popular paradigm on more solid conceptual grounds.

However, this optimistic outlook seems to clash with some of the most obvious ways of making the notion of expectation more precise. For example, while Ginnobily & Olmos (2021) do highlight how the empirical literature rests on an ambiguous use of the notion of expectation, any attempt to flesh out this notion is done in terms of varieties of (representational) content, which, at first glance, is not compatible with an enactivist re-reading.[10]

The problem here is that, unlike more enactivist-friendly notions like perception and action, the expectations of the VoE paradigm appear to have a feature that makes them inherently representational: satisfaction conditions. That is, while there is considerable disagreement on whether perception requires talk of normativity or regarding how action must be discussed via 'contentspeak' or 'representationese', there seems to be less disagreement concerning the fact that expectations are things that can be met – and fail to be met – *by definition*. So even if it turns out that the mental entity referred to by the word 'expectation' in VoE isn't satisfactory, it still looks like it has to refer to

[8] Against this, both Gallagher 2017 and Hutto 2019 offer attempts to frame numerical cognition in enactivist terms.

[9] For example, here is an extract from Hutto's (2019) formula for capturing numerical cognition under radical enactivist principles, "My recommended formula for creating a satisfactory enactivist account of mathematical cognition is to enact the following procedure: Subtract any residual commitment to mental representation, information-processing stories, and neuro-fetishism. Add, in place of these items, a more Andersonian account of neural reuse – one that focuses on the pluripotent, protean brains and which places the greater weight on the contributions of socio-cultural practices in establishing mathematical content and competencies" (Hutto 2019, p.16).

[10] For example, they write that "we do not mean that VoE is neutral with respect to every empirical question–if it were, it would have no empirical content at all—but only with respect to certain specific controversies about the nature of the postulated mental *contents* (are they beliefs, capacities, abilities, are they innate, are they learned, etc.)." (Ginnobily and Olmos 2021, p.106, emphasis added)

something that can obtain and fail to obtain. In this sense, it is difficult to see how enactivists could re-read expectations in terms that do not involve something like representational content.

The claim here is that, insofar as we associate representational content with correctness conditions– which, as mentioned earlier, seems to be what many representation-avoiding enactivists see as being a defining trait of representations – it looks like expectations have to be thought of as things that involve the presence of representations. This possibility of our expectations being met or not, I argue, indicates the presence of representations about how the world is *supposed* to behave, on a standard reading of what representations (and expectations) are.

Importantly, this does not mean that there is no way for enactivists to explain subjects' behaviour in these experiments. After all, what representations are is far from obvious, so it may be possible to flesh out what expectations are without appealing to representational content. It's just not obvious how to do so, given the appearance of a strong association of expectations and satisfaction conditions. So the argument here is not that there are no reasons to banish the R-word from the theoretical foundations of VoE. It is simply that until such reasons are given, it looks like an important tool for studying 'basic' minds is off limits to enactivists, given that this paradigm relies heavily on expectations and representations.

The challenge for enactivists seems to be to either come up with such an explanation of VoE results that does not rely on the notion of expectations, or to appeal to a representation-lite version of expectations. While neither has been explicitly addressed in the enactivist literature (as far as I can tell), enactivism may have resources to attempt both. Space once again is our enemy here, so I limit myself to brief speculative sketches of how this might go.

Regarding the first option, an enactivist could potentially offer a reinterpretation of behaviour displayed in VoE studies in terms of a subject's learning to interact with its environment, rather than in terms of expectations being violated.[11] On this re-interpretation, there is no violation of expectation, there is merely increased attention to a novel scene as part of a learning procedure. It looks like such a re-interpretation could fit well with the enactivist conception of perception as action, to the extent that the subject's exploring of its world forces it to pay more attention to new things. Subjects would look longer at impossible outcomes because these display novelty that warrants further attention and exploration, as part of the organism's sensemaking activities.

As for the second option, one way this could be explored is by reframing the intentionality of the expectations in terms of an enactivist-friendly portrayal of this problematic notion. Here the fact that expectations are supposed to be about the world is cashed out in terms of an organism's coupled relation with its environment and its skillful interaction with it, as opposed to a part of the brain's being somehow directed towards the world. Though this is an almost criminally gross oversimplification, one question that comes up here is how making intentionality into a organism-wide notion rather than something parts of the organism have helps explain what (representation-lite) expectations are.

Perhaps the main issue facing an enactivist in both options is that of explaining why there is longer looking time for certain forms of novelty but not others. In VoE studies, experimenters are measuring looking time in relation to novelty presented in a target domain (e.g. color, numerosity, surface area, etc.) in comparison to control domains. This means they often will present stimuli that display novelty in many aspects of the environment, to see if varieties of new stimuli elicit similar responses in control domains as they do in the target domain. In the standard view, the peculiar form of novelty of impossible outcomes can often be explained by appealing to expectations being violated in cases of impossible outcomes vs expectations not being violated in the same way for cases of novel-but-possible outcomes (e.g., sudden colour change vs moving through an object). The enactivist will have to offer an expectation-free explanation of why some types of novelty elicit longer looking times and others do not, since both seem to offer learning opportunities.

Conclusion

I have argued that enactivist approaches to cognition face a challenge when trying to assimilate an important paradigm used to probe the cognitive lives of preverbal infants and animals. In itself, this challenge has nothing particularly special about it, if not perhaps for the fact that it targets forms of cognition that are usually considered within the explanatory wheelhouse of enactivism. Given the myriad experimental paradigms used across the cognitive sciences, chances are that similar empirically-motivated challenges lurk nearby. In a sense, the fact that infant and animal cognition is well-worn terrain for enactivists means that they likely have more resources at their disposal to answer this particular challenge than those concerning more abstract forms of cognition. In any case, the purpose of this paper was certainly not to argue that enactivists do not have the resources to answer such challenges. On the contrary, it is to urge further reflection on actual scientific practice so that enactivism and other alternative approaches to cognition get closer to mainstream scientific practice in the cognitive sciences. While science does sometimes undergo wholesale conceptual change (e.g. Kuhn 1962), until empirical challenges of the sort laid out here are met, it is far from obvious that the time has come for full-scale revolution in cognitive science, despite what some may say (Meyer & Brancazio 2022).

Acknowledgments

This work has been made possible by financial support from the "QUANTA: Evolution of Cognitive Tools for

[11] I am thankful to an anonymous reviewer for suggesting this possibility.

Quantification" project, which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 951388). I am thankful to Olga Dudojc as well as two anonymous referees for helpful comments on initial drafts of this paper.

References

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5- month-old infants. *Cognition*, 20, 191–208.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Clark, A. (1993). *Associative engines: Connectionism, concepts, and representational change.* The MIT Press.

Clark, A. (2008). *Supersizing the Mind.* Oxford: OUP.

Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese, 101*(3), 401–431.

Colditz, I. G. (2020). A consideration of physiological regulation from the perspective of Bayesian enactivism. *Physiology & behavior*, *214*, 112758.

Di Paolo, E. A., Buhrmann, T., & Barandiaran, X. E. (2017). *Sensorimotor life: An enactive proposal*. Oxford: Oxford University Press.

Egbert, M.D., Barandiaran, X.E. (2022). Using enactive robotics to think outside of the problem-solving box: How sensorimotor contingencies constrain the forms of emergent autononomous habits. *Frontiers in Neurorobotics*16.https://doi.org/10.3389/fnbot.2022.847054

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition* 28, 3-71.

Gallagher, S. (2008). Are minimal representations still representations? *International Journal of Philosophical Studies,* 16(3), 351–69.

Gallagher, S. (2017). *Enactivist Interventions: Re-Thinking the Mind.* Oxford: OUP.

Gibson, J. J. (1979) *The ecological approach to visual perception.* Erlbaum.

Ginnobili, S., & Olmos, A.S. (2021). Empirical assumptions behind the violation of expectation experiments in human and non-human animals. *HPLS* **43**, 106. https://doi.org/10.1007/s40656-021-00459-7

Hutto, D. D. (2019). Re-doing the math: Making enactivism add up. *Philosophical Studies*, 176(3), 827–837.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Meyer, R., & Brancazio, N.(2022). Putting down the revolt: Enactivism as a philosophy of nature. *Front Psychol*. Oct 21;13:948733. doi: 10.3389/fpsyg.2022.948733.

Newen, A., De Bruin, L., & Gallagher, S. (Eds.) (2018). *The Oxford Handbook of 4E Cognition*. Oxford: OUP https://doi.org/10.1093/oxfordhb/9780198735410.001.0001

Onishi, K.H. and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–8.

Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.

Robbins, P., & Aydede, M. (Eds.) (2008). *The Cambridge Handbook of Situated Cognition* (Cambridge Handbooks in Psychology). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511816826

Rolla, G., Vasconcelos, G., & Figueiredo, N. M. (2022). Virtual Reality, Embodiment, and Allusion: an Ecological-Enactive Approach. *Philosophy & Technology*, *35* (4), 1-23.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge, MA: MIT Press.

Ward, D., Silverman, D. & Villalobos, M. (2017). Introduction: The Varieties of Enactivism. *Topoi* **36**, 365–375 https://doi.org/10.1007/s11245-017-9484-6

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology, 30*, 111–153.

Zahnoun, F. (2021). On representation hungry cognition (and why we should stop feeding it). *Synthese* 198 (Suppl 1), 267–284. https://doi.org/10.1007/s11229-019-02277-8