

Leonard Dung

## Against the explanatory argument for enactivism

Abstract:

Sensorimotor enactivism is the view that the content and the sensory modality of perceptual experience are determined by implicit knowledge of lawful regularities between bodily movements and patterns of sensory stimulation. A proponent of the explanatory argument for sensorimotor enactivism holds that this view is able to provide an intelligible explanation for why certain material realizers give rise to certain perceptual experiences, while rival accounts cannot close this “explanatory gap”.

However, I argue that the notion of the “material realizer” of perceptual experience is ambiguous. On a narrow construal, the explanatory gap cannot be bridged, not even by enactivism. On a wide construal, enactivism gets a grip on the explanatory gap, but to the same extent do established theories of consciousness. Thus, on both horns of the dilemma, the explanatory power of enactivism is not superior to more traditional theories of conscious experience.

### 1. The explanatory argument

Roughly, sensorimotor enactivism<sup>1</sup> is the view that the content of perceptual experience is determined not (mainly) by brain processes, but through dynamic interaction with the external world. Many enactivists claim that an enactivist account of perceptual experience partially bridges the explanatory gap obtaining between conscious experience and its material realizer (Hurley and Noë, 2003; Noë, 2007, pp. 462-467; O’Regan, 2011).<sup>2</sup> This is the *explanatory argument* for enactivism.

Helpfully, Hurley and Noë (2003) distinguish the *absolute*, the *comparative intermodal* and the *comparative intramodal* explanatory gap. The absolute explanatory gap refers to the lack of explanation for why a material realizer should give rise to consciousness at all. By contrast, the comparative explanatory gaps concern why a particular material realizer gives rise to an experience in a particular sensory modality rather than another one, e.g. visual rather than auditory, (intermodal) or why it gives rise to a particular experience rather than another one in the same modality, e.g. seeing yellow rather than seeing purple (intramodal). Enactivism is supposed to help with both comparative explanatory gaps, rather than the absolute one. Thus, the

---

<sup>1</sup> Henceforth, I use ‘enactivism’ as shorthand for ‘sensorimotor enactivism’.

<sup>2</sup> The *material realizer* of a conscious experience is the physical, probably neuroscientific, property in virtue of which the subject has the experience. This terminology aims to be neutral between different varieties of physicalism and dualism.

former will be our topic in what follows.<sup>3</sup>

I will first outline the basic structure of the explanatory argument. Thereafter, I will scrutinize its assumption that its material realizer has to be explanatory of conscious experience itself. Finally, I will argue that there is a sense in which the material realizer indeed has to be explanatory and that – in this sense – other non-enactivist conceptions can allow that the realizer of conscious experience fulfills this explanatory demand. Thus, enactivism possesses no advantage over competing views in tackling the explanatory gap.

I reconstruct the enactivist argument in the following way which is especially faithful to Alva Noë's presentation of the argument:

p1: We should prefer accounts of perceptual experience which allow an explanation for why the material realizer of consciousness gives rise to a certain experience or an experience of a certain sensory modality.

p2: There is no explanation for why a purely neural realizer of consciousness gives rise to a certain experience or an experience of a certain sensory modality.

p3: An enactivist account of perceptual experience explains why the material realizer of consciousness gives rise to a certain experience or an experience of a certain sensory modality.

C: We should prefer an enactivist account of perceptual experience to every account according to which the material realizer of conscious experience is purely neural.

I will now elaborate on the specific propositions composing this argument. p1 encapsulates the demand that the realizer of a conscious experience has to explain the properties of the realized experience. This demand is less strong as it might seem, because it is limited in two ways: First, Noë does not require that the properties of the realizer (maybe in conjunction with background conditions) a priori entail the properties of the experience. Hence, Noë's demand does not exclude the conceivability of zombies. It is sufficient when the realizer makes *intelligible* why the experience has the character it has. This leads to an "aha!" reaction (Hurley and Noë, 2003, p. 160) or to an insight about how the material facts "shap[e] the contours of the perceptual facts" (Noë, 2007, p. 465). Second, the properties of the realizer do not have to explain why they are associated with some phenomenal experience at all (that is, the absolute explanatory gap). They just illuminate why – given that they are associated with *some* experience – the experience is of a certain sensory modality or has the specific character it has (Hurley and Noë, 2003, pp. 147-148;

---

<sup>3</sup> Bridging the explanatory gap requires an account which explains in general terms how the connection between an experience and its material realizer can be intelligible, i.e., not arbitrary. The capacity of enactivism to solve the explanatory gap problem can be discussed independently of the fact that there may be *specific* pieces of empirical evidence which enactivism predicts or explains better or worse than its rivals. A general review of the empirical evidence for or against enactivism does not concern us here. For this reason, I set aside other arguments for enactivism, for instance its accounts of neural plasticity and sensory substitution (Hurley and Noë, 2003).

O'Regan, 2012, pp. 124-125).<sup>4</sup>

According to Noë, it is not strictly necessary that realizers carry out some explanatory work, however it is required for all practical purposes. The reason is that the belief in a certain realization-relation can never be justified if the realizer is not intelligibly connected to the realized property. The evidence for the claim that  $x$  realizes  $y$  has to partly consist in the fact that  $x$  explains properties of  $y$ . One might deny Noë's claim on the grounds that there are other ways to collect evidence for the existence of a realization-relation. For instance, exception-less correlation between  $x$  and  $y$  and the possibility of interventions on  $x$  which change  $y$  might be taken to provide sufficient evidence that  $x$  realizes  $y$ . Nevertheless, even under this assumption it is still plausible that the possibility of an explanation of  $y$  via  $x$  provides solid evidence that  $x$  realizes  $y$ . Hence, it is at least the case that something should be treated as a realizer if it explains better than the alternatives *and is otherwise on a par*.

p2 is based on the claim that neural properties are “qualitatively inscrutable” (Hurley and Noë, 2003, p. 132), i.e., the neural properties themselves give no indication of the experience which they might realize. O'Regan (2012) supports this claim with the observation that one can always intelligibly ask why a certain neural property should give rise to a particular experience. Since it might be considered the received view on the topic (see Levine, 1983), I will accept p2 for the sake of the argument.

According to sensorimotor enactivism, the content of our perceptual experience (as well as its modality) is determined by our implicit knowledge of sensorimotor dependencies. Sensorimotor dependencies are lawlike relations between bodily movements and sensory stimulation. Perception, respectively perceptual experience, consists in the active bodily exercise of this know-how. Crucially, different sensory modalities are governed by different sensorimotor regularities. For instance, eye movements affect visual sensations in systematic ways, while they do not influence olfactory experience. A further example is that only vision is governed by laws of occlusion, i.e., you can see an object only if its sight is not blocked by another object. This has sensorimotor consequences for it entails patterns in how movements can bring an object into or remove it from view.

Enactivism predicts that the modality, in which sensory stimulations are experienced, changes when the sensorimotor dependencies associated with the stimulation are transformed to the dependencies typical for a different modality. Noë and O'Regan interpret studies on tactile visual sensory substitution systems (TVSS) (devised by Bach-y-Rita, 1972) as lending support for this enactivist claim. In TVSS, a camera is attached to a blind subject. This camera transmits visual input to a device on the body of the subject. In response to and depending on the camera

---

<sup>4</sup> Many enactivists frequently emphasize the power to explain the sensory modality of an experience, as opposed to explanations of intramodal differences. Nevertheless, enactivism is claimed to achieve both.

input, this device stimulates the skin of the subject in a corresponding pattern. More specifically, it looks like the sensorimotor dependencies characteristic of vision are (partially) transferred to the tactile stimulation via the device, i.e., movements affect tactile stimulation in a way characteristic of vision. Enactivism predicts that a partial emulation of the sensorimotor patterns of vision leads to partially visual experience, although neither the eyes nor the visual cortex are experimentally manipulated.

The experimental results may be regarded as vindication of this prediction. After a short period of adaptation in which subjects perform various kinds of movement-involving goal-directed activity, they start to report quasi-visual experiences of looming objects (Clark, 2014, p. 224). In addition, they report experiencing typically visual effects and illusions (Hurley and Noë, 2003, p. 144). Noë and O'Regan conclude that these experiences are indeed visual in character and that, hence, TVSS provides evidence for enactivism. Crucially, the adaptation only occurs in case the subject is allowed to freely interact with and explore the environment and – in doing so – to intentionally control the camera.

At least in principle, the same ideas apply to variation within one sensory modality. Since perceptual experience is determined by implicit knowledge of sensorimotor dependencies, fine-grained changes in sensorimotor dependencies should lead to changes in perceptual experience (Hurley and Noë, 2003, pp. 147-148). Different objects afford different possibilities for interaction and consequently different sensorimotor dependencies, hence they are experienced differently.

I will not attempt to evaluate the overall plausibility of and evidence for the enactivist theory here. I will grant that it can explain the character of our perceptual experience to the extent that the explanatory argument requires. The fact that certain sensory stimulation is governed by sensorimotor dependencies characteristic of vision is an illuminating explanation for why the resulting experience is visual. The correlation between patterns of skill-mediated sensorimotor interactions with the world and certain experiences does not seem as arbitrary as the correlation between experiences and neural activity. Noë regards these patterns of interaction with external objects, which are exercises of sensorimotor skills, as the material realizers of conscious experiences (Noë, 2007, p. 465). If this assumption is conceded, it follows p3, i.e. that enactivism explains how the material realizer of perceptual experience can be explanatory. In conjunction with the other premises, this establishes that the enactivist account of perceptual experience possesses additional explanatory power relative to every account which depicts the material realizer of perceptual experience as purely neural.

## **2. Why purely neural explanations are always uninformative**

I claim – contra the explanatory argument – that enactivist and “neuro-centric” views are

explanatorily entirely on a par. There is nothing that the material realizer of conscious experience explains on an enactivist account which could not also be accounted for by a purely neural realizer. I will postpone the question which of the premises of the argument it is that I reject. The reason is that these premises can be interpreted in different ways. Any of them can be correctly judged as true or false, depending on how the key term “material realizer” is understood.

I agree with proponents of the explanatory argument that neural activity considered in isolation – without being embedded in a wider theoretical framework – is not explanatory. But neither do I think that this fact is surprising nor that it indicates a serious explanatory gap. On the contrary, it reflects a well-known insight that is shaping theorizing in cognitive science since many decades. This insight consists in the recognition that low-level facts about the physical organization of a cognitive system do not suffice to attain real understanding of what the system is doing and how it manages to be successful. This claim is encapsulated in Marr’s (1982) distinction between the computational, algorithmic and implementational level of description. Despite the criticism this methodological approach has attracted, the idea that cognitive systems have to be understood on more than the implementational level is still up to date.<sup>5</sup> The reasons for that seem to be twofold.

First, an understanding of a system as distinctively *cognitive* requires abstraction from the physical details. Only when you move to a higher level of description, you are able to grasp the properties that different cognitive systems, while differing widely in their physical micro-organization, share. Second – and more importantly – cognitive systems are too complex to predict, let alone understand, their behavior, complex cognitive states and experiences from their implementational properties alone (Marr, 1982; Dennett, 1987, pp. 24-27). Employing higher levels of description simplifies tasks of prediction and especially explanation, since they abstract from the bewildering amount of physical detail found on the implementational level(s). Hence, we cannot expect to understand a cognitive phenomenon, like perceptual experience, by considering its neural realizer alone.

Clark (2009, p. 972) points out that it is *in principle* possible to derive from information about neural activity explanations of why people tend to behave the way they do, including their verbal reports. However, the explanatory argument demands not (only) the *in-principle* availability of an explanation, but its *actual* availability. Enactivism would still be explanatorily superior to neuro-centric accounts if it could provide an explanation of experiential properties in terms of sensorimotor dependencies while standard-accounts could only insist on the in-principle possibility of a neural explanation. Furthermore, one must show that neural properties actually

---

<sup>5</sup> The main misgivings are that the three levels are significantly more interdependent than Marr thought and that the levels, especially the implementational level, are extremely heterogenous. They do not undermine this key tenet.

explain properties of perceptual experience in order to do justice to the enactivist contention that  $x$  must explain  $y$  for us to have sufficient evidence that  $x$  realizes  $y$ . The explanation does not have to be complete, it might be sketchy and programmatic, but it has to be able to illuminate why variations in brain activity give rise to variations in experience. Otherwise, we would still lack evidence that purely neural properties are the realizers of perceptual experience.

To summarize, I concede that facts about neural realizers in isolation do not explain properties of the experience they realize, i.e. one possible interpretation of p2. However, if one conceives of realizers in this way, I reject the assumption that such an explanatory link is necessary, i.e. p1. Neural properties do not explain cognitive phenomena in isolation but by being embedded in a wider theoretical framework. The connection between neural activity and perceptual experience is not arbitrary, if certain kinds of neural activity are especially suited to play a functional role in a theory of consciousness which is – as a whole – capable of explaining properties of experience. Since the wider framework possesses the requested explanatory power, we are able to receive evidence for the framework which constitutes also evidence for a certain realizer-relation because the relation is postulated by the framework. Before I spell out this claim in more depth by means of a concrete example, I want to note that even p3 begins to look suspect on this narrow construal of a material realizer.

If the material realizer is construed as being on the level of low-level implementational detail, it seems that even the enactivist explanation of properties of perceptual experience is not available. Counting body and world as part of the realizer is not by itself sufficient to attain some explanatory benefit. The enactivist explanation succeeds only by virtue of abstraction from details of the realizers and the introduction of *implicit knowledge of sensorimotor dependencies* as theoretical construct. The intelligible connection between realizer and experience obtains in virtue of the former's role in this more abstract description, i.e., because body and world are essential *for* the exercise of sensorimotor know-how.

If one limits oneself to the low-level material details alone, the extended brain-, body- and world-encompassing realizers, which the enactivist postulates, cannot explain perceptual experience either – both in the comparative intramodal and the comparative intermodal sense. If one takes the connection between realizers and higher-level theories into account, then neural realizers alone suffice to attain the same level of understanding of why a certain experience occurs as the enactivist supplies. In that case, p2 is false. I will show that now.

### **3. Making neural realization explanatory**

I am going to demonstrate how a cognitive theory of consciousness equips neural properties with an explanatory role. The global (neuronal) workspace theory of consciousness (GWT) will serve as example, but the lessons are generalizable to other neuro-centric theories. GWT holds that

mental representations are conscious iff they are made globally available for use in a wide variety of cognitive processes (Dehaene & Naccache, 2001; Mashour et al., 2020). The property of being conscious might even be *identical* to the property of being globally available (in the right kind of information-processing system). Global availability of information achieves coordination and control between different domain-specific subsystems and enables higher cognitive functions which require the manipulation of information processed in many different subsystems.

This function is subserved by a neuronal workspace consisting of a distributed set of cortical neurons which are able to receive and to send information to homologous neurons in distant cortical areas through long-range excitatory axons (Dehaene, Changeux and Naccache, 2011, p. 56). It is hypothesized that pyramidal neurons of layer 2 and 3, which are especially elevated in prefrontal, parietal-temporal and cingulate associative cortices, serve as these “workspace neurons”. Due to their long axons and dense connectivity, it is easy to see why they would be ideally suited to perform long-range range transmission of information through the cortex.

GWT explains why a certain neural representation becomes conscious. It becomes conscious, because it manages – via a combination of stimulus-driven and top-down amplification – to reach a certain threshold of activity, to excite the workspace neurons and to enter the global workspace. According to the traditional story which suffices here, neurons in the sensory cortices function as “feature detectors”, i.e., their activation correlates with the presence of certain perceptible features in the environment. Certain networks of neurons code for certain features. This establishes a non-arbitrary link between the activation of certain neurons in sensory cortex and the contents of a subject’s perception. A deeper explanation would need to resort to one of the competing theories about how the content of neural representations is determined. With these two ingredients, we have already arrived at an intelligible explanation for why certain neural activity gives rise to certain conscious contents. GWT explains why neural activity  $n$  is conscious and  $n$ ’s role as feature detector – or even a theory of representational content – illuminates why  $n$  is associated with a certain content. Even if the contents of perceptual experience and the contents of the underlying neural representations are not identical, they are clearly intimately intertwined. This amounts to an explanation of why  $n$  gives rise to a certain conscious content. By the same token, variations in neural activity are intelligibly connected to variations in conscious content. Since the content of a neural representation entails its sensory modality,<sup>6</sup> there is besides an explanatory link forged between  $n$  and its experienced modality.

---

<sup>6</sup> To see this, it might help to bring to mind that neurons in different sensory cortices are sensitive to different kinds of low-level features. If there exist indeed neurons in different sensory cortices which respond to the same features, then their role as, e.g., visual neurons is determined by their embedding into visual cortex.

This approach makes intelligible why certain neuronal activity in sensory cortices and workspace neurons gives rise to certain conscious contents and to contents in a certain sensory modality. The “aha!” reaction typical of explanatory understanding might be elicited by the realization that a subject’s particular experience is explained by the fact that (i) an assembly of neural representations with corresponding content is active and this content is globally broadcast and (ii) that – according to a theory which purports to explain the functional features of consciousness – this process of global broadcasting is sufficient to bring about consciousness. Thus, the explanatory power of this approach equals enactivism’s.<sup>7</sup> It follows that p2 is false if material realizers are conceived of in the context of a wider cognitive theory. I conclude that there exists no relevant conceptualization of material realizers on which all three premises are true. The argument is therefore unsound.

Conversely, it has been shown that enactivist explanations are not “deeper”, i.e., it is not in general the case that enactivism explains correlations between material realizer and perceptual experience more satisfactorily than neuro-centric views. The enactivist might, however, still argue that specific perceptual experiences, respectively their variations, cannot be explained equally well by neuro-centric views, like e.g. in TVSS. But that would be a different sort of argument: It amounts not to an argument to the effect that enactivism provides in general more intelligible explanations of the content and modality of perceptual experiences, but that enactivism predicts specific empirical phenomena, like the adaptation of subjects in TVSS studies, more successfully than its neuro-centric rivals.

Before concluding, I will address a possible counterargument. An enactivist might complain that GWT, as proposed by me, can be interpreted in an externalist fashion or even needs to be interpreted this way to plausibly overcome the explanatory gap. For it is arguable that experiences acquire their content via their relation to entities in the environment. If experiences are individuated by their contents, entities in the environment would play an indispensable role in the explanation of why a certain brain state gives rise to a certain experience. In this way, my argument might even support the contention that the material realizer of consciousness extends beyond the brain.

However, this objection is ineffective since it conflates two distinct and orthogonal forms of externalism. *Content externalism* posits that the *content* of some mental states – in this case perceptual content – is partly determined by the subject’s relations to entities in its external environment, rather than being entirely determined by factors internal to the subject (Putnam, 1975). This view is distinct from the externalist view considered here, according to which the *material realizer* of a mental state – in this case perceptual experience – involves objects in body

---

<sup>7</sup> It might even be considered more powerful, if one regards it as an explanation for why neuronal activity becomes conscious *at all*.



and environment. Content externalism pertains to the content of mental states, enactivism makes a claim about the realizer or constituents of mental states. Even if the story furnished here required us to ally GWT with content externalism, this would not imply that the neuro-centric view regarding the realizer of experience is false (Rupert, 2004). Externalists about the content of perceptual experience do not need to believe that the material processes underlying experience extend into the environment (Siegel, 2021).

#### **4. Conclusion**

To summarize, sensorimotor enactivists claim that the correlation between the material realizer of perceptual experience and the experience itself can be explained by positing that this realizer extends into body and world. According to this line of thought, sensory stimulation gives rise to particular perceptual experiences because our interaction with the sensed objects exerts implicit knowledge of a range of sensorimotor dependencies which is characteristic for this type of experience.

Against this thesis, I argued that the notion of the “material realizer” of perceptual experience is ambiguous. On a narrow construal, the explanatory gap cannot be bridged, not even by enactivism, since details about low-level physical organization, considered in isolation, are frequently incapable of explaining higher-level cognitive phenomena. On a wide construal, when the notion of a material realizer is conceived of as embedded in an encompassing high-level theory, enactivism gets a grip on the explanatory gap, but to the same extent do established theories of consciousness like, for instance, the global workspace theory. Thus, on both horns of the dilemma, the explanatory power of enactivism is not superior to more traditional theories of conscious experience. Consequently, the explanatory argument fails. That being said, the preceding argument is directly targeted only at the sensorimotor (knowledge) variety of enactivism. Other forms of enactivism, when they refrain from positing a theoretical mediator between experience and its material realizer (like implicit knowledge of sensorimotor dependencies), may evade the objection developed here.

## References

- Bach-y-Rita, P. (1972) *Brain mechanisms in sensory substitution*, New York: Academic Press.
- Clark, A. (2009) Spreading the joy? Why the machinery of consciousness is (probably) still in the head, *Mind*, 118(472), 963-993. <https://doi.org/10.1093/mind/fzp110>
- Clark, A. (2014) *Mindware: an introduction to the philosophy of cognitive science*, 2<sup>nd</sup> edn. Oxford: Oxford University Press.
- Dehaene, S., & Naccache, L. (2001) Toward a cognitive neuroscience of consciousness: Basic evidence and a workspace framework, *Cognition*, 79(1-2), 1-37. [https://doi.org/10.1016/S0010-0277\(00\)00123-2](https://doi.org/10.1016/S0010-0277(00)00123-2)
- Dehaene, S., Changeux, J. and Naccache, L. (2011) The global neuronal workspace model of conscious access: from neuronal architectures to clinical applications, in Dehaene, S. and Christen, Y. (eds.) *Characterizing consciousness: from cognition to the clinic?*, Berlin: Springer Verlag Berlin Heidelberg, 55-84.
- Dennett, D. (1987) True believers: the intentional stance and why it works, in Dennett, D. (ed.) *The intentional stance*, Cambridge Massachusetts: The MIT Press, 13-35.
- Hurley, S. and Noë, A. (2003) Neural plasticity and consciousness, *Biology and Philosophy*, 18(1), 131-168. <https://doi.org/10.1023/A:1023308401356>
- Levine, J. (1983) Materialism and qualia: The explanatory gap, *Pacific Philosophical Quarterly*, 64, 354-361. <https://doi.org/10.1111/j.1468-0114.1983.tb00207.x>
- Marr, D. (1982) *Vision: a computational investigation into the human representation and processing of visual information*, San Francisco: W.H. Freeman and Company.
- Mashour, G. A., Roelfsema, P., Changeux, J., & Dehaene, S. (2020) Conscious processing and the global neuronal workspace hypothesis, *Neuron*, 105(5). <https://doi.org/10.1016/j.neuron.2020.01.026>
- Noë, A. (2007) Magical realism and the limits of intelligibility: What makes us conscious?, *Philosophical Perspectives*, 21(1), 457-474. <https://doi.org/10.1111/j.1520-8583.2007.00132.x>
- O'Regan, J. K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness*, New York, USA: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199775224.001.0001>
- O'Regan, J. K. (2012) How to build a robot that is conscious and feels, *Minds & Machines*, 22(2), 117-136. <https://doi.org/10.1007/s11023-012-9279-x>
- Putnam, H. (1975) The meaning of 'meaning', *Minnesota Studies in the Philosophy of Science* 7, 131-193.
- Rupert, R. D. (2004) Challenges to the hypothesis of extended cognition, *Journal of Philosophy* 101(8), 389-428. <https://doi.org/10.5840/jphil2004101826>
- Siegel, S. (2021) The contents of perception, in Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2021.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/perception-contents/>. Accessed 9 February 2022.



