

Commentary: Investigating the concept of representation in the neural and psychological sciences

A commentary on

Investigating the concept of representation in the neural and psychological sciences

by Favela, L. H., & Machery, E. (2023) *Front. Psychol.* 14:1165622. doi: 10.3389/fpsyg.2023.1165622

Introduction

Favela & Machery (2023) describe four experiments probing the role of the concept *representation* in the brain sciences. They show that, given short descriptions of brain activity, neuroscientists and psychologists are generally not confident whether it should be described as a representation or not. Favela & Machery interpret this to mean that the scientists are unsure what it takes for brain activity to be, or *count as*, a representation. And they conclude that the concept *representation* should either be eliminated from the brain sciences, or reformed.

The experiments are revealing, and constitute an important methodological advance on existing approaches to the concept *representation*, which mostly use a priori reflection and case studies (Baker et al., 2022; Poldrack, 2020; Ramsey, 2007; Shea, 2018). But this commentary will argue that the study's design is not well-suited to its ultimate goal, and that Favela & Machery's conclusion relies on an implausible assumption about scientific concepts. Discarding that assumption will make room for important work building on Favela & Machery's contribution.

Scientific concepts and how to understand them

Each experiment probed "scientists' willingness to use different kinds of descriptions" (3)¹ of brain activity, focusing on ones that describe brain activity as "representing" its environment. After seeing a "cover story about a neuroscientific study recording brain response to various stimuli" (3), participants were asked whether they agreed with a statement asserting that the brain's response *represented* the stimuli, responding on a Likert scale from "strongly agree" to "strongly disagree." They were also asked about other descriptions, some involving representational notions (like "being about") and some causal ones (like "responding to").

Three of the four experiments modulated a particular feature of the brain activity (its scale, relation to the stimulus, and function in the brain) to probe its effect on the acceptability of representational descriptions. The fourth investigated participants' willingness to describe brain activity as *misrepresenting* stimuli. In each case, participants were told of the brain's response to certain stimuli, and they were asked how willing they were to describe that response as (among other things) *representational*. In other words, they were asked to *categorize* the brain's responses, or *taxonomize* them, into *representations* and *non-representations*.

For causal descriptions, like "the brain area responds to the stimulus," responses clustered around the ends of the Likert scale. But for representational descriptions, like "the brain area represents the stimulus," the answers cluster around the middle of the scale. The natural interpretation is that although scientists are confident in some (especially causal) categorizations

¹ All page numbers refer to (Favela & Machery, 2023).

of brain activity, they are not confident in their categorizations of brain activity as *representing* a stimulus or not.²

This is an interesting finding, and supports an interesting conclusion: whatever the concept *representation* contributes to the brain sciences, it doesn't contribute a clear taxonomy of brain activity into the categories *representational* and *non-representational*. But Favela & Machery conclude that the concept of representation needs to either be reformed, or simply eliminated from the brain sciences. How do you get that conclusion from those findings? Only by assuming that what the concept *representation* contributes could only be a taxonomy of neural activity into the categories *representational* and *non-representational*, or that whatever it contributes must depend on that taxonomy.

This picture of a concept's scientific role will be dubious to anyone familiar with the psychology of concepts or the nuances of scientific practice. I'll summarize two reasons, before returning to the positive lessons of Favela & Machery's study. First, scientific practice shows us that taxonomy is not all scientific concepts do. When scientists conceive of misinformation as a virus, they do not assume that the concept *virus* sorts the world into two kinds of things, viruses and non-viruses, and that misinformation falls into the former category. Rather, they are using the concept to introduce modeling tools, assumptions, and conceptual frameworks to study disinformation (Kucharski, 2016). And when fluid mechanics is applied to model traffic, there is no assumption that traffic *is a fluid*, or that the correct description of traffic is *as a fluid* (Sun et al., 2011).³ The idea is to introduce modeling resources that are applicable to traffic for reasons that, while interesting, do not involve traffic's *being a fluid*. A study that presented scientists with different traffic scenarios, asking them whether they agreed with statements like "the traffic is a fluid," would not capture the work that the concept *fluid* is doing for this area of science.

Second, there is already work that applies psychological methods to study how concepts figure into explanation; this is closely related, for obvious reasons, to questions of how concepts figure into science. Consider Lombrozo & colleagues' paradigmatic work on the explanatory role of the concept *function*. Some of this work asks which things tend to be attributed functions by which populations (Lombrozo et al., 2007). But often, and more informatively, it asks what participants can *do* once they've characterized a target in terms of the concept *function*, e.g., what predictions or generalizations they can make given functional as opposed to mechanistic descriptions of a system (Lombrozo, 2009). Because the concept *function* might contribute something to explanations other than a taxonomy (into things that have functions and things that don't), this research has found ways to probe what functional descriptions are (or can be) *used to do*, rather than just what conditions elicit them.

Discussion

Favela & Machery provide evidence that representational concepts in cognitive science do not provide a clear taxonomy of neural activity into the categories *representational* and *non-representational*. This is important, but it does not support the conclusion that the concept of representation does no useful work for cognitive science. I haven't aimed to *defend* the concept of representation here. Whether it serves an important scientific role or not, whether it should be retained, reformed, or eliminated, depends on what it does for science. And that can be

² At least when they're asked by philosophers: a potential problem, if they are aware that philosophers tend to be very particular in how they define the concept *representation*.

³ Though, to be fair, that kind of assumption is precisely what characterizes some fascinating Ig Nobel Prize-winning work (Fardin, 2014).

investigated partly with a priori and case study methods (Richmond, n.d.), but it can also be investigated experimentally. And if we set aside their more ambitious conclusions, Favela & Machery provide a great starting-point for that investigation.

References

- Baker, B., Lansdell, B., & Kording, K. P. (2022). Three aspects of representation in neuroscience. *Trends in Cognitive Sciences*, 26(11), 942–958.
<https://doi.org/10.1016/j.tics.2022.08.014>
- Fardin, M. A. (2014). On the rheology of cats. *Rheology Bulletin*, 83(2), 16–30.
<https://doi.org/10.1002/actp.1984.010350801>
- Favela, L. H., & Machery, E. (2023). Investigating the concept of representation in the neural and psychological sciences. *Frontiers in Psychology*, 14(1165622), 1–13.
<https://doi.org/10.3389/fpsyg.2023.1165622>
- Kucharski, A. (2016). Post-truth: Study epidemiology of fake news. *Nature*, 540(7634), 525.
<https://doi.org/10.1038/540525a>
- Lombrozo, T. (2009). Explanation and categorization: How “why?” informs “what?” *Cognition*, 110(2), 248–253. <https://doi.org/10.1016/j.cognition.2008.10.007>
- Lombrozo, T., Kelemen, D., & Zaitchik, D. (2007). *Inferring Design: Evidence of a Preference for Teleological Explanations in Patients With Alzheimer’s Disease*. 18(11), 999–1006.
- Poldrack, R. A. (2020). The physics of representation. *Synthese*. <https://doi.org/10.1007/s11229-020-02793-y>
- Ramsey, W. M. (2007). *Representation Reconsidered*. Cambridge University Press.
- Richmond, A. (n.d.). What is a Theory of Neural Representation For? *Forthcoming*.
- Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press.
- Sun, D., Lv, J., & Waller, S. T. (2011). In-depth analysis of traffic congestion using computational fluid dynamics (CFD) modeling method. *Journal of Modern Transportation*, 19(1), 58–67. <https://doi.org/10.1007/bf03325741>