This article is part of the "Progress & Puzzles of Cognitive Science" letter series.

Do Retinal Neurons Also Represent Somatosensory Inputs? On Why Neuronal Responses Are Not Sufficient to Determine What Neurons Do

Lotem Elber-Dorozko,[a,b] Yonatan Loewenstein[c,d,e,f]

[a]*Department of Humanities and Arts, Technion*
[b]*The Center for Philosophy of Science, University of Pittsburgh*
[c]*The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University*
[d]*Department of Neurobiology, The Alexander Silberman Institute of Life Sciences, The Hebrew University*
[e]*Department of Cognitive Sciences, The Hebrew University*
[f]*The Federmann Center for the Study of Rationality, The Hebrew University*

Abstract

How does neuronal activity give rise to cognitive capacities? To address this question, neuroscientists hypothesize about what neurons "represent," "encode," or "compute," and test these hypotheses empirically. This process is similar to the assessment of hypotheses in other fields of science and as such is subject to the same limitations and difficulties that have been discussed at length by philosophers of science. In this paper, we highlight an additional difficulty in the process of empirical assessment of hypotheses that is unique to the cognitive sciences. We argue that, unlike in other scientific fields, comparing hypotheses according to the extent to which they explain or predict empirical data can lead to absurd results. Other considerations, which are perhaps more subjective, must be taken into account. We focus on one such consideration, which is the purposeful function of the neurons as part of a biological system. We believe that progress in neuroscience critically depends on properly addressing this difficulty.

Keywords: Neuronal representation; Neuronal computation; Biological function; Competing hypotheses

Correspondence should be sent to Lotem Elber-Dorozko, Department of Humanities and Arts, Technion, 3200003, Haifa, Israel. E-mail: Lotem.elber@gmail.com

Much of neuroscientific research is focused on identifying what neurons represent, even if the definition of representation is debated. In practice, if neuronal activity correlates or causally relates to the relevant variable, the hypothesis about "representation"[1] of that variable is taken to be supported. For example, we believe that the somatosensory cortex encodes information about somatosensory stimulation. Some observations that support this are that somatosensory stimulation elicits a response in somatosensory cortical regions (e.g., Saadon-Grosman, Loewenstein, & Arzy, 2020) and that electrical stimulation of these cortical regions elicits a report of feeling somatosensory stimulation (e.g., Penfield & Boldrey, 1937). Similar experiments underlie the beliefs that faces are represented in the fusiform gyrus (Kanwisher, McDermott, & Chun, 1997), that "reward prediction error" is encoded in the ventral tegmental area (Schultz, Dayan, & Montague, 1997), and many similar beliefs.

Neuroscientists may disagree about neuronal representations.[2] Consider an experiment in which a neuron has been shown to respond when images of faces are presented to a participant. One neuroscientist may argue that a face-representing neuron was identified. Nevertheless, there are alternative interpretations. A second neuroscientist may argue that the neuron represents oval shapes. A third neuroscientist may argue that the neuron responded only because a nose has been presented. A common way of testing these competing hypotheses is to conduct additional experiments, in which the competing hypotheses predict different patterns of neuronal activity. For example, present nonface oval objects, nonoval faces, and noses in isolation to the participant.

However, this approach can lead to absurd conclusions. To see why, consider the question of what retinal ganglion cells represent. It is generally accepted that these cells represent a spot of light at a particular spatial angle (for simplicity, we ignore center-surround properties). It is well-known that pressing the eyeball elicits light sensations also known as "phosphenes" (Grüsser & Hagner, 1990). Taking these phosphenes as evidence for the activation of retinal ganglion cells in response to somatosensory stimulation, one may hypothesize that retinal ganglion cells represent both visual and somatosensory aspects of the world. If we were to test the competing hypotheses by recording the activity of retinal ganglion cells in response to somatosensory stimuli, we would be compelled to judge in favor of the visual+somatosensory representation hypothesis. Nonetheless, neuroscientists take the visual representation hypothesis to be correct and the visual+somatosensory hypothesis to be absurd. Similarly, while the activity of the vestibular system is affected by alcohol ingestion (Tianwu et al., 1995), it is not believed that the vestibular system also *represents* the recent history of alcohol consumption.

It has been argued before that a neuron's response to a stimulus is not sufficient to claim that it *represents* it (Dretske, 1988, chap. 3; Cao, 2022; Baker, Lansdell, & Kording, 2022). Moreover, even an analysis of a multiarea pattern of activation is not sufficient to determine between two different hypotheses about representation (Gessell, Geib, & De Brigard, 2021). Here, we make an even stronger claim: no amount of neuronal data, by itself, can be used to conclusively decide between two competing hypotheses about representation. A hypothesis that is more strongly supported by the data, such as the visual+somatosensory representation hypothesis, can be considered unlikely for other reasons.

We claim that, in practice, different considerations guide us when comparing hypotheses about neuronal representations. One reason why the neuroscience community takes retinal

ganglion cells to represent visual input but not somatosensory input is because it is believed that the eye has been adapted for visual processing. More generally, much neuroscientific work is guided by the intuition that representations are necessarily linked with purposeful, teleological *function*. According to popular philosophical views, representations (Dretske, 1988; Millikan, 1989; Neander, 2017) and computations (Fresco, 2021; Piccinini, 2015) are determined according to the functions they perform,[3] and some functions make more sense than others. A well-known example in philosophy is that the heart's function is to pump blood and not to make thumping sounds (Wouters, 2005). Similarly, we can say that retinal ganglion cells do not represent somatosensory input because they do not have the function of relaying somatosensory information.

The difficulty with biological function is that there is no clear-cut way to know the function of a system or subsystem. It is often posited that biological functions are relevant to survival and reproduction (Millikan, 1989; Piccinini, 2015, Chapter 6; but see Cummins, 1975). While these guidelines allow us to exclude certain hypothesized functions, such as making thumping sounds, they still leave much room for various possible functions assigned to the same system.

In the examples of retinal ganglion cells and the vestibular system, neuroscientific intuition is relatively clear. However, there are many cases where neuroscientists may disagree regarding function. Consider a neuron in the visual cortex that responds exclusively to elongated, rope-like shapes. One neuroscientist may hypothesize that the role of some neurons in the visual cortex is to facilitate avoidance of dangerous stimuli. On this view, the recorded neuron may represent snakes. The fact that a similar object, say a garden hose, elicits a similar response is interpreted as a misrepresentation, just like the vestibular system's response to alcohol is a misrepresentation. Another scientist may assume that neurons in the visual cortex aim to accurately represent the visual scene, of which elongated objects are occasionally a part. On this view, there is no misrepresentation. Crucially, if one is convinced by the previous examples that response properties may differ from representation, it becomes evident that the question of whether this neuron represents "snakes" or "elongated objects" cannot be resolved solely by conducting more experiments. While response properties play an important part in identifying what neurons represent, alone they are not decisive, as each neuroscientist may legitimately argue that some response properties are "misrepresentations." What will settle this debate is agreement on the function of the neuron.

One invaluable way to understand neuronal activity is to examine it in the context of behavior (Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017; Niv, 2021). Nonetheless, a behavioral response cannot replace nor determine "function." To see the usefulness, but also the limitations of behavioral experiments, consider a sea-creature with receptors that, like retinal ganglion cells, respond both to light and touch. Studying the creature's behavior, we discover that it moves in the direction of the activated receptors, regardless of the source of activation. As both visual and somatosensory inputs lead to a behavioral response, the behavioral experiment is consistent with various hypotheses about representation. Therefore, behavior per-se is not sufficient to determine representation. However, if we discover that in the course of evolution, the food of the creature's ancestors tended to be closer to well-lit areas, we may conclude that its phototactic behavior has a function. If no analogous function for somatosensory-guided movement is found, we may say that the creature's receptors are

light receptors, while their response to touch is a "side-effect" of a mechanism whose purpose is to move the creature to well-lit areas. This insight thus required us to study both the creature's behavior and to speculate about the effect of this behavior throughout evolution [(Krakauer et al., 2017) argue similarly].

The argument presented here indicates that neuroscientific data alone may not suffice to resolve scientific disagreements about representation. How can such disagreements be resolved? We suggest that an explicit discussion of the considerations of function that play a part in the various stages of neuroscientific practice will help clarify disagreements. Such discussions can benefit from the philosophical literature that distinguishes between different descriptions of biological function, whether it is determined by natural-selection, other historical properties, or current contribution to goals (Wouters, 2005). Adopting a clear stance on biological function can also help come up with ways to address questions about function empirically, for example, by examining evolutionary changes (Cisek & Hayden, 2022), or how environmental change affects behavior on different time scales. Moreover, serious consideration of biological function is important to avoid telling just-so stories, which describe functions as evolutionarily selected without justification (Gould & Lewontin, 1979). We are hopeful that shining a spotlight on biological function as a central feature guiding neuroscientific decisions will help create more precise ways of thinking about representation in the brain.

Notes

1 For brevity, we discuss representation alone, but similar points can be made regarding computation.
2 We assume that neuronal representations are independent of scientific interests and explanations. For alternative views (Cao, 2022; Hacohen, 2022; Egan, 2014).
3 In contrast to these references, we aim to describe how representations can be identified (epistemology) and make no claims about what they are exactly (ontology).

Acknowledgments

References

Baker, B., Lansdell, B., & Kording, K. (2022). Three aspects of representation in neuroscience. Trends in cognitive sciences, 26, 942–958.
Cao, R. (2022). Putting representations to use. *Synthese*, *200*(2), 151.
Cisek, P., & Hayden, B. Y. (2022). Neuroscience needs evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1844), 20200518.

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

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. MIT Press.

Egan, F. (2014). How to think about mental content. *Philosophical Studies*, *170*(1), 115–135.

Fresco, N. (2021). Long-arm functional individuation of computation. *Synthese*, *199*(5), 13993–13916.

Gessell, B., Geib, B., & De Brigard, F. (2021). Multivariate pattern analysis and the search for neural representations. *Synthese*, *199*(5), 12869–12889.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London*, *205*, 581–598.

Grüsser, O.-J., & Hagner, M. (1990). On the history of deformation phosphenes and the idea of internal light generated in the eye for the purpose of vision. *Documenta Ophthalmologica*, *74*(1), 57–85.

Hacohen, O. (2022). What are neural representations? A Cummins functions approach. Philosophy of Science,89, 701-720.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302 LP–4311.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, *93*(3), 480–490.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science*, *56*(2), 288–302.

Neander, K. (2017). *A mark of the mental*. MIT Press.

Niv, Y. (2021). The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, *135*(5), 601–609.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, *60*(4), 389–443.

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

Saadon-Grosman, N., Loewenstein, Y., & Arzy, S. (2020). The 'creatures' of the human cortical somatosensory system. *Brain Communications*, *2*(1), fcaa003.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Tianwu, H., Watanabe, Y., Asai, M., Shimizu, K., Takada, S., & Mizukoshi, K. (1995). Effects of alcohol ingestion on vestibular function in postural control. *Acta Oto-Laryngologica*, *519*, 127–131.

Wouters, A. (2005). The function debate in philosophy. *Acta Biotheoretica*, *53*, 123–151.