DOI: 10.1017/psa.2025.10133
This is a manuscript accepted for publication in *Philosophy of Science*.
This version may be subject to change during the production process.

Thinking About Neuron Doctrines

Mara McGuire

Department of History and Philosophy of Science
University of Pittsburgh
maramcguire@pitt.edu

Abstract. Philosophers of science have overlooked the role of theory in neuroscience, resulting in a somewhat surprising naïveté regarding the nature and function of neuroscientific theories. Here I provide a framework that identifies and begins to characterize what we need to know about neuroscientific theories so as to improve our epistemic standing. I argue that we need an account of the structural, interpretive, and functional aspects of neuroscientific theories, using the *neuron doctrine* as an illustrative case study. I introduce the novel metaphor of *theoretical infrastructure* as a guide for making sense of neuroscientific theories and their place within neuroscientific practice.

Acknowledgements. I want to express my gratitude to John Bickle whose work initially sparked my interest in these issues and whose mentorship extends far beyond the many fruitful conversations, and friendly disagreements, we have had about the place of theory in neuroscience. For their instructive written comments and advice on various aspects of the paper's content and framing I would like to thank Ken Aizawa and Edouard Machery. I am very grateful to Sameera Singh, Madeleine Potoskie, Gal Ben-Porath, Dzintra Ullis, Caitlin Mace, Helen McGuire, and Joshua Kramer for their insightful and meticulous feedback, written and otherwise. Finally, I greatly benefited from feedback and suggestions that I received from audiences at meetings of the Philosophy of Science Association, the Society for the Metaphysics of Science, and the Deep South Philosophy of Neuroscience Workgroup.

Funding Information/Declarations. None to declare.

1. Introduction

If asked to catalogue historically neglected topics in the philosophy of science, one would probably not put "theory" at the top of their list. Some might not even put "theory" *on* their list. John Bickle, for one, would fall in this latter camp. In recent work on tool development and other "extra-theory aspects" of neuroscience, Bickle lauds contemporary philosophy of science insofar as it has come to recognize that theory does not, in fact, exhaust scientific practice (Bickle 2022). He then laments the way in which extra-theory scientific practices are often still treated as supporting cast in service of the *real* goals of science, namely, theory progress, theory confirmation, etc. Both Bickle and the science-in-practice movement more widely reject this "theory centrism." The rejection is not a denial that theory has *any* place in neuroscience. It is "an attempt *to put theory in its proper place*" (Bickle 2022, 14; emphasis in original).

In response to such a statement, we may reasonably ask: "What *is* the 'proper place' of theory in neuroscience?" Save a passing remark to the effect that theory depends on tool development, Bickle does not provide us with an answer. Fair enough, his interests lie elsewhere. Still, we might expect to find answers from other, more theory-friendly folk. But this is not the case. Although much has been written about theory in the philosophy of science writ large, relatively little has been written about theory in neuroscience.[1] My intention in opening with Bickle's somewhat – perhaps even intentionally – provocative comments is not to criticize them in what follows. Rather, I find these comments constructively provocative: they prompt us to consider the largely overlooked nature and role of theory in neuroscience.

Contemporary philosophy of science finds itself, then, in the somewhat surprising position of being unable to put theory in its proper place because of a relative naïveté regarding neuroscientific theories. What needs to be known about theory in neuroscience to overcome this naïveté, such that we *can* consider the relation between theory and extra-theory aspects of neuroscience? It is my aim to address this question. To this end, I motivate the need for an account of the structural, interpretive, and functional aspects of theory (Section 2). I use the *neuron doctrine* as a case study to illustrate the questions an account of theory will need to address (Section 3). I conclude by proposing the metaphor of *theoretical infrastructure* as a guide for continuing to think about theory (Section 4).

---

[1] Some argue that theory is now neglected even in the physical sciences (Morrison 2007).

2. Theory Avoidance (And How *We* Might Avoid *It*)

Fussing over the neglect of neuroscientific theories is only reasonable if there are, in fact, theories in neuroscience. A theory-in-neuroscience skeptic might call attention to passages where philosophers write that theory plays "little role" (compared to models) in neuroscience (Bechtel and Richardson [2000] 2010, xvii) or characterize neuroscience as being "data rich and theory poor" (Churchland and Sejnowski 2016, 667). Even if neuroscience is "theory poor," sparse coffers are not the same as empty ones. If there are even *some* theories in neuroscience, it is reasonable to want an account of them. It is also reasonable to interrogate what notion of "theory" is being used when someone proclaims that theories are missing from neuroscience. There are many notions of "theory" (Sellars 1965). Neuroscience might appear more theoretically impoverished than it really is if we are working with a notion of "theory" ill-suited for thinking about theory within neuroscience. It would be foolish to dismiss this possibility out of hand. Philosophers of science have a time-honored tradition of using the physical sciences as a measure against which the biological sciences are compared. Adopting this approach has led to such conclusions as: the biological sciences don't *really* have laws because these purported laws do not align with our understanding of those in the physical sciences (Smart 1968). Alternative conclusions are certainly available. Perhaps drawing comparisons to the physical sciences is simply the wrong way to approach questions of knowledge or representation in the biological sciences (Mitchell 2000). Biology may have laws but may require a different framework than the one used for the physical sciences to account for such laws.

What philosophers have said about theory in neuroscience is of a similar, discipline-relative spirit. Ken Schaffner has argued that illustrations of theory in the biomedical sciences and neuroscience only partially overlap with those in the physical sciences, the latter exhibiting theories whose structure – i.e., the sort of entity that it is (Winther 2021) – might be accounted for on a *syntactic* conception of theory, where theories are formally specified, deductive axiomatic systems that are universal in scope (Schaffner 1993). Instead, we more often find instances of what Schaffner calls a "theory of the middle range" (Schaffner 1993, 321), a theory that is narrower in scope and whose structure is better captured by a *semantic* conception on which theories specify a class of models. More specifically, Schaffner characterizes theories as a "series of overlapping interlevel temporal models" (Schaffner 1993, 320) and, in doing so, highlights the existence of *nonformal* components. Subsequent work has continued to emphasize the nonformal components of theory in the biological sciences. Carl Craver, for example, has argued that many illustrations of theory in neuroscience are mechanistic in nature, where mechanistic theories are "abstract and idealized descriptions of a type of mechanism" – i.e., "mechanism schemata" (Craver 2002, 69).

Both Schaffner and Craver can be read as advocating a type of pluralism. Theories may have formal components, nonformal components, or both. While their arguments for pluralism and the inclusion of nonformal components are well taken, I will later argue that plausible instances of neuroscientific theories fall outside the scope of either Schaffner or Craver's accounts of structure. Even if my arguments fall flat, structure is only one aspect of theory. A further aspect of theory, one that is commonly discussed for theories in the physical sciences (Boyd and Bogen 2021) but that neither Schaffner nor Craver explicitly address, is theory interpretation.

One understanding of what theory interpretation *is* likens it to the translation of a foreign text (Dewar 2023). A theory is imbued with meaning when we put the components of the theory into correspondence with symbols whose meaning has already been specified. This way of thinking about interpretation is based on the way this process works in the physical sciences, where theories are often formally specified in a way that the syntactic or semantic conceptions of theory may capture. As indicated above, nonformal components are likely more characteristic of theories in the biological sciences, including neuroscience. Is there a comparable need for theory interpretation when considering theories with nonformal components? A formally specified theory is a calculus whose symbols lack significance prior to interpretation (Dewar 2023). But the symbols on a mechanistic diagram, for example, do seem to carry significance – e.g., "⟶" represents direction, "Na^+" represents sodium ions, etc.

Even if some nonformal components come equipped with already specified meanings, I argue that a need for interpretation still remains. One function of interpretation is to specify the commitments of a theory (Dewar 2023). If the commitments of neuroscientific theories are not completely transparent, there remains a plausible role for an interpretive project. Of course, such differences in the types of theory components may require a new consideration of what interpretation *is* in this context. For example, some might want to treat the interpretation of nonformal components as a type of interpretation*. So be it. What matters is whether there is *some* plausible notion of interpretation – or interpretation* – that remains relevant. In the next section, I show that the neuron doctrine admits of a variety of interpretations, making the case that the interpretive aspect of theory is still of central importance for neuroscience.

The reader may recall that my goal is to identify those aspects of theory that need accounting for in order to be able to put theory in its place – i.e., situate theory in relation to extra-theory aspects of neuroscience. My inclusion of an explicit interpretive aspect is an improvement over previous treatments that focus more exclusively on structure, yet this addition is not sufficient to accomplish the task at hand. An account of the functional role of theory in neuroscience is also needed.

Thinking about the function of various aspects of scientific practice – e.g., scientific definitions (Knuuttila and Loettgers 2017) – is widespread. But the function of theory in particular has received more minimal explicit attention, typically from those who advocate a pragmatist view of theory (Winther 2021). Some of this lack of attention may be due to an assumption that an account of the function of scientific theories is a misguided endeavor. Drawing a parallel to arguments against the plausibility of accounting for the uses of *language* on the grounds that there is an infinite number of uses to which language can be put (Cappelen 2018), we might think that theories likely have so many potential uses that to account for them is a pointless exercise. It is likely the case that scientific theories can be used for all sorts of things and that trying to catalogue all these uses would be pointless. But, taking inspiration from those who have not been dissuaded from thinking about linguistic function (Thomasson 2024), this myriad of uses does not entail that *any* account of theory function is misguided. Given a notion of function on which the function of some *x* is not equivalent to "any and every use to which some *x* may be put," we may have some success in specifying a range of functions that neuroscientific theories can play.

3. Meet the Neuron Doctrine(s)

I now turn to the neuron doctrine as a case study for further illustrating how we should be thinking about neuroscientific theories.[2] All statements of the doctrine portray the neuron, or neuronal activity, as some fundamental structure and/or functional unit. Yet, as I will discuss at some length, there are varied interpretations as to what this means. Consider the following

[2] Unless otherwise specified, all subsequent references to the "doctrine" refer to the neuron doctrine.

examples: “The neuron doctrine represents nerve cells as polarized structures that contact each other at specialized (synaptic) junctions and form the developmental, functional, structural, and trophic units of the nervous system” (Guillery 2005, 1281); the doctrine claims that the neuron is “the structural and functional unit of computation” (Saxena and Cunningham 2019, 103); the doctrine claims that neurons are “the relevant functional units in the brain” (Cao 2014, 891); the doctrine claims that “the only functionally relevant aspect of the brain, what gives it its behavioral capacities and gives us our mental states, is neuronal activity” (Cao 2022, 3). These interpretations are not equivalent. Being the unit of computation (Saxena and Cunningham 2019) does not necessarily entail that only neuronal activity is functionally relevant to our behavior and mental states (Cao 2022). Charting the dimensions along which interpretations vary is one goal of this section. A briefer discussion of the structure of the doctrine and potential functions bookends this interpretive analysis.

There are two reasons why I have chosen the neuron doctrine in particular to illustrate the limitations of current treatments of neuroscientific theories, and the sorts of questions that we will encounter as we try to improve and expand upon these accounts. The first reason is the recognized centrality and prominence of the doctrine within neuroscience. The neuroscientist Marina Bentivoglio writes, “Will there ever be another doctrine as fundamental as the neuron doctrine in what are now the neurosciences and cell biology in general?” (Bentivoglio 2016, xvii). Similarly, the neuroscientist Rafael Yuste writes, “For over a century, the neuron doctrine…has provided a conceptual foundation for neuroscience” (Yuste 2015, 487). The second reason is that, despite its prominence, the doctrine remains largely overlooked by philosophers. The most notable exception is Rosa Cao, who has argued that it is time for neuroscience to leave the doctrine behind (Cao 2014, 2022).

To emphasize once more the nature and scope of my aims, an evaluative project like Cao’s is not the kind of project I am engaged in here. Whether the doctrine is a theory worth keeping around is a different question from whether the doctrine is a mechanism schemata. It is likely unnecessary to have a full account of the structure, interpretation, and function of the doctrine in view before we embark on an evaluative project, yet some understanding is necessary. Structure influences the relevant notion of interpretation and a central role of interpretation is to specify the commitments of a theory. These commitments are certainly relevant to an evaluative project. Just as we need a better understanding of theory before we can situate it in relation to extra-theory aspects of science, we need a better understanding of the doctrine before we can evaluate it. After a brief discussion of structure, I return to the interpretive issue.

Treatments of theory structure in neuroscience have tended toward pluralism and have emphasized nonformal theory components. It seems plausible that some neuroscientific theories are a series of “overlapping interlevel temporal models” that bear a resemblance relation to each other and are narrow in scope, as Schaffner suggests. Or, that some theories are abstracted and idealized descriptions of mechanisms, as Craver suggests. But the doctrine does not appear to be the type of theory that either Schaffner or Craver had in mind when they developed their accounts of structure. For one thing, the doctrine is not narrow in scope. Its scope is not restricted to specifying a type of structure-function relationship that only holds for some organisms but that bears resemblance to a structure-function relationship found in other organisms. For another thing, the doctrine is not a mechanism, abstract or otherwise: it is not interlevel; it lacks temporal stages; there are no set-up or termination conditions (Machamer, Darden, and Craver 2000). I do not want to entirely rule out any possibility of these accounts being modified such that they could account for the doctrine. The main takeaway is that, even if this is possible, the necessary work has yet to be done.

Returning to the analysis of interpretation, I suggest that interpretations of the doctrine can vary along the following dimensions: (1) structure *versus* function, (2) lower- *versus* higher-functions, and (3) causal *versus* methodological notions of fundamentality.

Structure versus Function: Interpretations of the doctrine differ in their emphasis on its structural or functional components; these differences potentially change the theory’s commitments. For example, the introduction and initial interpretation of the doctrine near the end of the nineteenth century was predominantly structural, a claim about the tissue in the nervous system being composed of discrete nerve cells (Shepherd 2016). The Spanish anatomist and histologist Santiago Ramón y Cajal (1852-1934) was among the most vocal proponents of this theory, defending it against the so-called *reticular theory* on which nervous tissue is composed of a continuous network of neurons, the view championed by Cajal’s rival, the Italian histologist Camillo Golgi (1843-1926) (Ramón y Cajal 1906). Interpretations of the doctrine that foreground structural concerns, like Cajal’s, have functional implications (Guillery 2005). In contrast, interpreting the doctrine as committed to the neuron being the unit of computation in the nervous system (Saxena and Cunningham 2019), for example, foregrounds functional commitments in a way that the interpretation in the Cajal-Golgi debate does not. This leads to the second dimension along which interpretations may vary, the nature of the functions in question.

Lower- versus Higher-Functions: At the risk of stating a triviality, the nervous system is thought to be subserving an enormous range of functions, from the motor to the sensory, the cognitive to the conative, the organizational and integrative to the nutritive, etc. Philipp Haueis draws a useful distinction between *cognitive* and *non-cognitive* functions supported by neural activity, these notions identifying the ends of a continuum of neural functioning rather than discrete categories (Haueis 2018). The former pertains to more complex "functions that process behaviorally relevant information," the latter to functions that "do not process behaviorally relevant information but maintain and repair neural and other systems of the body" (Haueis 2018, 5373). While this is a rough distinction in need of further explication, it is sufficient to note that different interpretations of the doctrine appeal to different notions of function. For example, Cao appeals to behavioral and mental functions (Cao 2022), Horace Barlow to sensory and perceptual functions (Barlow 1972), Sir Charles S. Sherrington to lower-order – what Haueis might consider "non-cognitive" – functions (Sherrington 1906).

Distinguishing interpretations along the dimensions of structure *versus* function or lower- *versus* higher-functions does not, however, illuminate the notion of *fundamentality* at play. While specifications of structure-function relationships are widespread in neuroscience, the structure-function relationship in the doctrine is unique insofar as it is regarded as fundamental in some (largely uninterpreted) sense. To put it glibly, no one is championing an *amygdala doctrine* as the "conceptual foundation" of neuroscience. Multiple interpretations of fundamentality are available. Here, I will limit myself to discussing two: *causal* and *methodological*. While such distinctions have been drawn in other contexts – e.g., adaptationism (Godfrey-Smith 2001) – they have not been applied to the doctrine.

Interpretations of Fundamentality: On the *causal* interpretation of fundamentality, the doctrine involves an empirical commitment to the structure of the biological world being a certain way, a commitment to the neuron as the most causally important unit in the nervous system for the relevant function(s). On the *methodological* interpretation of fundamentality, the doctrine involves a commitment to a neuron-centric approach to the study of the nervous system, a way of structuring or organizing investigations within neuroscience, rather than an empirical claim about the causal structure of the nervous system.[3]

As Peter Godfrey-Smith notes in his application of similar distinctions to the adaptationist debate, these interpretations are logically distinct (Godfrey-Smith 2001). Still, in actual practice, the causal interpretation might have some bearing on the methodological one. The latter may involve some implicit empirical commitments as to the causal structure of the nervous system, as this causal structure must have some relevance to one's evaluation of whether a neuron-centric approach is worthwhile. At a minimum, the epistemic payoffs of a neuron-centric approach seem partially dependent on facts about the nervous system. Even so, the empirical commitments on the methodological interpretation are more minimal than those on the causal interpretation. The methodological interpretation may also take into consideration epistemic aims and interests, cognitive and technological limitations, etc. in a way that the causal interpretation does not.

To illustrate, consider the sort of evidence drawn upon in criticisms of the doctrine. In recent decades, findings from experimental work and computational modeling indicate that astrocytes (a type of non-neuronal brain cell) and populations of neurons, respectively, may make greater contributions to information-processing in the nervous system than previously thought (Cao 2014, 2022; Yuste 2015). Growing evidence suggests that astrocytes play an important role in synaptic transmission via their ability to both release neurotransmitters into, and reuptake them from, the synaptic cleft (Araque 2008; Araque et al. 1999). Neural network models that analyze neural dynamics distributed across an entire population of neurons challenge the idea that a *single* neuron is the fundamental unit of structure and function (Barack and Krakauer 2021; Ebitz and Hayden 2021; Yuste 2015). Criticisms like these are relevant to, and problematic for, some interpretations of the doctrine – e.g., a causal interpretation of fundamentality where the relevant functions of a neuron are taken to include higher-order,

---

[3] While I cannot discuss it here, the *ontological* interpretation of fundamentality involves a commitment to the neuron as the smallest entity within a hierarchy of entities that is still *neuroscientific* (rather than, say, chemical or physical). I thank Ken Aizawa for suggesting this notion of fundamentality to me.

cognitive and behavioral ones (Cao 2022). Alternative interpretations of the doctrine, such as the methodological interpretation, may not be equally affected. Suppose astrocytic processes are centrally involved in nervous system functioning. This does not preclude the possibility that organizing our investigations around the neuron may still best facilitate our (to-be-specified) aims.

Even so, one might question whether the doctrine under the methodological interpretation should still be considered a type of theory. One avenue through which to approach this question is a consideration of theory function. Can the doctrine, under the methodological interpretation, play the sort of functional role that we expect of neuroscientific theories? To answer this, we would need an account of theory function. As with structure and interpretation, I aim to explore the sorts of questions that we will encounter as we develop a framework for thinking about the function of neuroscientific theories.

The reader may recall that whatever notion of theory function we settle on cannot be a notion on which the function of theory is equivalent to how a theory is used on any given day by any given scientist. This is compatible with aiming to develop an account that captures a plurality of roles that theory may be thought to play – e.g., explanation, prediction, control, representation, unification, posing research questions, identifying mechanisms, experimental design, fitting models, etc. (Morrison 2007; Winther 2012, 2021). A list of functions is far from illuminating. A more promising approach is to try and identify higher-order categories of functions, or macro-functions, which may reveal underlying commonalities and relations among them. As a first pass, we might try to distinguish between *backward*- and *forward-facing* macro-functions. The former, backward-facing macro-function might include such sub-functions as figuring in explanations or supporting predictions. To play these sorts of roles, a theory must have relatively rich content. The latter, forward-facing macro-function might include such sub-functions as organizing or structuring current and future investigations, perhaps even ones that are more exploratory in nature. In this role, theories might be thought of as analogous to directives, with the sort of content required being more minimal.

Returning to the discussion of the doctrine through which we entered the topic of function, it would seem that, if the doctrine under the methodological interpretation is going to support any category of function, it will be the forward-facing kind – e.g., organizing investigation. The plausibility of this claim hinges on a much more elaborate characterization and defense of the various macro-functions that theory can play *and* the relation among them. With regard to the particular question at hand, this would necessarily involve an argument illustrating that the forward-facing macro-function is not dependent upon the features required for the backward-facing macro-function. If there is such a dependency, the doctrine under the methodological interpretation would likely be incapable of having any meaningful functional role because it lacks the richer content in virtue of which a theory can support a backward-facing macro-function. More work is needed to explore this question in neuroscience. However, these live questions about theory function already reveal how the choice of an interpretation of a theory in the first place affects one's commitments about the doctrine.

4. Conclusion: The Theoretical Infrastructure of Neuroscience

I began this paper by noting a relative naïveté regarding theory in neuroscience. In light of the subsequent analysis, are we any less naïve? I hope that the answer is affirmative. I have identified three aspects of theory – structure, interpretation, and function – that we need an account of if we want to make and evaluate claims like: neuroscientific theories are "*secondary to and dependent upon* new tool development, both temporally and epistemically" (Bickle 2022, 14; emphasis in original). I have discussed limitations of extant treatments of neuroscientific theories. And I have illustrated, and started to explore, some of the questions that we should consider when providing an account of the structure, interpretation, or function of the neuron doctrine in particular, and neuroscientific theories more generally. Are forward-facing functions independent of backward-facing ones? Can the doctrine under the methodological interpretation play a forward-facing functional role? If so, does this entail that the doctrine has significant epistemic value despite the criticisms that have been raised against it?

Rather than continuing to catalogue the questions that we will eventually need to answer, I want to conclude by briefly introducing the metaphor of *theoretical infrastructure* that I believe holds value as a guide for continuing to think about neuroscientific theories, and perhaps scientific theories more generally. While our stereotype of infrastructure likely features things like roadways, telecommunications networks, or bridges, infrastructure comes in two main varieties, *hard* and *soft*. Hard infrastructure refers to the sorts of things I listed just now, physical entities that are necessary for some enterprise to run smoothly. Soft infrastructure refers to the services, frameworks, rules and regulations, systems of governance, etc. that are sometimes less explicitly identifiable but whose implementation is also necessary for the smooth operation of some enterprise – e.g., health care services, social security, institutions for environmental protection, policies governing market competition, standards for safety inspections, etc. Combined, hard and soft infrastructure are the material and conceptual connections that organize and hold things together.

The metaphor of theoretical infrastructure is valuable for at least two reasons. First, the features and functions of soft infrastructure are useful guides for further developing the notion of a forward-facing macro-function, and the features needed to support this function. Thinking about soft infrastructure may help make explicit the legitimate, and perhaps independent, role of a framework or policy implemented for the purpose of directing, organizing, and structuring an enterprise. Second, the relation between hard and soft infrastructure suggests a way of thinking about the relation between theory – i.e., soft/theoretical infrastructure – and the extra-theory, material aspects of scientific practice – i.e., hard/material infrastructure.

As a very tentative sketch, the infrastructural metaphor may help us consider ways in which theory might be integrated into frameworks developed for thinking about extra-theory, material aspects of scientific practice. Hans-Jörg Rheinberger, for instance, has spent decades developing a framework for thinking about the biological sciences on which these sciences are oriented around an "experimental system," comprised of what Rheinberger calls "epistemic things" and "technical objects" (Rheinberger 1997). Epistemic things, the targets of scientific inquiry, are embedded within, and constrained by, a wider arrangement of "instruments, inscription devices, model organisms, and the gloating theorems or boundary concepts attached to them" (Rheinberger 1997, 29). Utilizing the infrastructural metaphor, these so-called technical objects might be repackaged as hard/material infrastructure. But hard/material infrastructure is not enough for complex systems to run smoothly, soft/theoretical infrastructure is also needed. Even on this breezy and preliminary discussion, the metaphor might be suggestive of a more nuanced view on which putting theory in its proper place means not minimizing it, but integrating it with other aspects of neuroscientific practice. How might this work? At present, we are still too naïve to say. But I hope to have shown how we can be *less* naïve about theory and its potential contributions to such infrastructure.

References.

Araque, Alfonso. 2008. "Astrocytes Process Synaptic Information." *Neuron Glia Biology* 4 (1): 3–10. https://doi.org/10.1017/S1740925X09000064

Araque, Alfonso, Vladimir Parpura, Rita P. Sanzgiri, and Philip G. Haydon. 1999. "Tripartite Synapses: Glia, the Unacknowledged Partner." *Trends in Neurosciences* 22 (5): 208–15. https://doi.org/10.1016/S0166-2236(98)01349-6

Barack, David L., and John W. Krakauer. 2021. "Two Views on the Cognitive Brain." *Nature Reviews Neuroscience* 22: 359–71. https://doi.org/10.1038/s41583-021-00448-6

Barlow, Horace B. 1972. "Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology?" *Perception* 1 (4): 371–94. https://doi.org/10.1068/p010371

Bechtel, William, and Robert C. Richardson. (2000) 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, Massachusetts: The MIT Press. https://doi.org/10.7551/mitpress/8328.001.0001

Bentivoglio, Marina. 2016. "Cajal, Golgi, and Ariadne's Thread." In *Foundations of the Neuron Doctrine*, edited by Gordon M. Shepherd, 25th ed., xvii–xx. New York: Oxford University Press. https://doi.org/10.1093/med/9780190259389.001.0001

Bickle, John. 2022. "Tinkering in the Lab." In *The Tools of Neuroscience Experiment*, edited by John Bickle, Carl Craver, and Ann-Sophie Barwich, 13–36. New York: Routledge. https://doi.org/10.4324/9781003251392

Boyd, Nora M., and James Bogen. 2021. "Theory and Observation in Science." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2021/entries/science-theory-observation/.

Cao, Rosa. 2014. "Signaling in the Brain: In Search of Functional Units." *Philosophy of Science* 81 (5): 891–901. https://doi.org/10.1086/677688

Cao, Rosa. 2022. "Multiple Realizability and the Spirit of Functionalism." *Synthese* 200 (6): 506. https://doi.org/10.1007/s11229-022-03524-1

Cappelen, Herman. 2018. *Fixing Language: Conceptual Engineering and the Limits of Revision*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198814719.001.0001

Churchland, Patricia S., and Terrence J. Sejnowski. 2016. “Blending Computational and Experimental Neuroscience.” *Nature Reviews Neuroscience* 17 (11): 667–68. https://doi.org/10.1038/nrn.2016.114

Craver, Carl F. 2002. “Structures of Scientific Theories.” In *The Blackwell Guide to the Philosophy of Science*, edited by Peter Machamer and Michael Silberstein, 55–79. Oxford: Blackwell Publishers Ltd. https://doi.org/10.1002/9780470756614

Dewar, Neil. 2023. “Interpretation and Equivalence; or, Equivalence and Interpretation.” *Synthese* 201 (4). https://doi.org/10.1007/s11229-023-04102-9.

Ebitz, R. Becket, and Benjamin Y. Hayden. 2021. “The Population Doctrine in Cognitive Neuroscience.” *Neuron* 109 (19): 3055–68. https://doi.org/10.1016/j.neuron.2021.07.011

Godfrey-Smith, Peter. 2001. “Three Kinds of Adaptationism.” In *Adaptationism and Optimality*, edited by Steven Hecht Orzack and Elliott Sober, 335–57. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511609084

Guillery, Rainer W. 2005. “Observations of Synaptic Structures: Origins of the Neuron Doctrine and Its Current Status.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360 (1458): 1281–1307. https://doi.org/10.1098/rstb.2003.1459

Haueis, Philipp. 2018. “Beyond Cognitive Myopia: A Patchwork Approach to the Concept of Neural Function.” *Synthese* 195 (12): 5373–5402. https://doi.org/10.1007/s11229-018-01991-z

Knuuttila, Tarja, and Andrea Loettgers. 2017. “What Are Definitions of Life Good For? Transdisciplinary and Other Definitions in Astrobiology.” *Biology & Philosophy* 32 (6): 1185–1203. https://doi.org/10.1007/s10539-017-9600-4

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. “Thinking About Mechanisms.” *Philosophy of Science* 67 (1): 1–25. https://doi.org/10.1086/392759

Mitchell, Sandra D. 2000. “Dimensions of Scientific Law.” *Philosophy of Science* 67 (2): 242–65. https://doi.org/10.1086/392774

Morrison, Margaret. 2007. “Where Have All the Theories Gone?” *Philosophy of Science* 74 (2): 195–228. https://doi.org/10.1086/520778

Ramón y Cajal, Santiago. 1906. “The Structure and Connexions of Neurons.” Nobel Lecture. December 12, 1906. https://www.nobelprize.org/prizes/medicine/1906/cajal/lecture/.

Rheinberger, Hans-Jörg. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.

Saxena, Shreya, and John P. Cunningham. 2019. "Towards the Neural Population Doctrine." *Current Opinion in Neurobiology* 55: 103–11. https://doi.org/10.1016/j.conb.2019.02.002

Schaffner, Kenneth F. 1993. "Theory Structure, Reduction, and Disciplinary Integration in Biology." *Biology and Philosophy* 8: 319–47. https://doi.org/10.1007/BF00860432

Sellars, Wilfrid. 1965. "Scientific Realism or Irenic Instrumentalism." In *Boston Studies in the Philosophy of Science*, edited by Robert S. Cohen and Marx W. Wartofksy, II:171–204. New York: Humanities Press.

Shepherd, Gordon M. 2016. *Foundations of the Neuron Doctrine*. 25th ed. New York: Oxford University Press. https://doi.org/10.1093/med/9780190259389.001.0001

Sherrington, Charles S. 1906. *The Integrative Action of the Nervous System*. New Haven, Connecticut: Yale University Press. https://doi.org/10.1037/13798-000

Smart, Jack. 1968. *Between Science and Philosophy: An Introduction to the Philosophy of Science*. New York: Random House.

Thomasson, Amie L. 2024. "How Should We Think about Linguistic Function?" *Inquiry* 67 (3): 840–71. https://doi.org/10.1080/0020174X.2022.2074886

Winther, Rasmus G. 2012. "Mathematical Modeling in Biology: Philosophy and Pragmatics." *Frontiers in Plant Science* 3: 102. https://doi.org/10.3389/fpls.2012.00102

Winther, Rasmus G. 2021. "The Structure of Scientific Theories." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/structure-scientific-theories/.

Yuste, Rafael. 2015. "From the Neuron Doctrine to Neural Networks." *Nature Reviews Neuroscience* 16 (8): 487–97. https://doi.org/10.1038/nrn3962