Function Ascription Practices in Circuit Interventions

Presented as part of the PSA 2024 symposium

Circuits in Neuroscience: Exploring the Brain's Mesoscale

Caitlin Mace | cbm49@pitt.edu

April 2025 (preprint)

Abstract: Novel tools have allowed researchers to intervene into circuits at the mesoscale. The results of these interventions are often explained by appeal to functions. How are functions ascribed to circuit parts experimentally? I identify two kinds of function ascription practices in circuit interventions. Analysis of these practices shows us that function ascriptions are challenging due to a lack of interventive control and insufficient constraints on the class of candidate functions to discriminate in practice. One kind of function ascription practice—subtractive analysis—fares better at addressing these challenges.

Main Text:

Neural circuits are thought to be the relevant functional unit in the brain for explaining cognition and behavior. Neuroscientists loosely refer to projections of neurons as 'circuits', but more precise characterizations of circuits appeal to anatomical connections of neurons that perform various functions. More precisely, neural circuits are anatomical connections of neurons along which signals flow or are processed (Ross & Woodward 2024). Interventions into circuits at various spatial and temporal scales provide a means to functionally decompose circuits into various parts to determine what those parts do.

But the long project of ascribing functions to circuit parts in the brain has been fraught with controversy since its origins (McCaffrey 2023). A central issue is that brain parts are pluripotential, so engaged in various cognitive processes (Poldrack, 2006; Figdor, 2010), as well as have multiple functional competences, meaning that a part may be competent to perform many functions (Barack 2024). Even worse, multiplexing of various functions means that they are not easily decomposable (Cao 2022). Problems with function ascriptions and

multifunctionality of brain parts have motivated contextualism about functions in and of the brain (Klein 2012; McCaffrey 2015; Burnston 2016, 2021), a view that lacks a satisfying story for how functions are ascribed other than by stipulation. Some have thrown in the towel and suggested that functions not be ascribed at all but rather be reduced to their causal roles (Klein 2017). But functions are more than causal roles. The functions often ascribed by neuroscientists and of interest here involve a dispositional capacity to produce some outcome *for* the behaving organism (Cummins 1975). The success of the reductionist project, then, depends on whether teleology can be reduced to causal roles without remainder (Craver 2012). But insofar as we are interested in functions, we can't avoid addressing what the brain part or its activity is for (Mitchell 2003). As for scientific inquiry, function ascriptions serve a broader purpose in guiding the search for brain areas or parts (Bechtel & Mundale 1999; Hardcastle & Stewart 2002) and mechanisms (Craver 2012; Chirimuuta 2024) that are causally implicated in experimental tasks. This at least explains why neuroscientists continue to ascribe functions to circuits and their parts, despite the challenges.

Why is it so challenging to ascribe functions to neural circuits experimentally? The answer to this question is fairly well understood—multifunctionality of brain parts and multiplexing of functions is a serious challenge for experimentalists. But this isn't the whole story. And such challenges are not insurmountable—we might think that technological and conceptual progress could make brain functions experimentally tractable. Such progress requires answering the question: how *can* we ascribe function experimentally to parts of the brain? Answering this question is my task. I argue that certain methods are more successful for ascribing functions, in part because they target function directly. In developing this argument, I identify two additional challenges for ascribing functions.

I focus here on interventions into circuit parts that are done in order to determine what that part does for the circuit. I identify one challenge to be a lack of interventive control over the whole circuit in Section 1. In Section 2, I argue that neuroscientists circumvent this issue by ascribing functions according to two methods: by appeal to background research or via subtractive analysis. While the latter is promising as a way to target function directly, an appeal to background research is unavoidable for function ascriptions, or so I argue in Section 3. Appeals to background research are a second challenge for ascribing functions. This is just because, for any given circuit part, there likely already exists a healthy disjunction of possible functions that the part may be performing, many of which are compatible with the same experimental results. While there need not be a single, unique function for each circuit part, usually only one function needs to be implicated in any explanation of a cognitive or behavioral capacity. The challenge of multiple candidate functions is overcome by focusing on a manageable selection of candidate functions at a time via subtractive analysis. I conclude by taking stock and suggesting strategies to strengthen function ascription practices.

I. Functional Decomposition of Circuits

Circuits exist at various scales in the brain. The hippocampus is part of a mesoscale circuit beginning and ending at the entorhinal cortex (EC) that is called the ‘trisynaptic’ circuit, so-called for its three circuit parts: granule cells in the dentate gyrus, pyramidal cells in CA3, and pyramidal cells in CA1. The hippocampal trisynaptic circuit sits within larger circuits, one of which is sometimes called the ‘fear pathway’ as the pathway also runs through the amygdala and is implicated in fear conditioning. The parts of the trisynaptic circuit also contain local microcircuits. Figuring out what the parts of the trisynaptic circuit (or the larger fear pathway) do is the project of much cognitive neuroscience. I focus on these mesoscale circuits as they are

particularly apt for intervention. At the mesoscale, there is both stability and plasticity of circuit structure. Stable circuit structure is thought to imply stability of function, allowing for generalizations both over time and across organisms, unlike at the microscale. Plasticity allows for interventions to elicit changes to the circuit, which, unlike at the macroscale, can be transient perturbations. These features of mesoscale neural circuits make them a desirable target of intervention, as opposed to circuits at micro- and macroscales. With circuit dissection tools like optogenetics and chemogenetics, researchers have intervened into mesoscale circuits to functionally decompose the circuit and ascribe functions to circuit parts.

Optogenetics and chemogenetics are well-suited for the task of investigating mesoscale circuits as they provide the spatial specificity and temporal precision required to target particular causal steps in a circuit. These causal steps are the activity or inactivity of neurons (i.e., circuit parts) along the circuit (Ross 2021). Using optogenetics and chemogenetics, researchers isolate the causal step by genetically modifying the target cells so they can be intervened on by light or chemicals, respectively. Intervening on these causal steps directly activates (or inhibits) the circuit to cause (or prevent) behavior. An experiment along these lines ideally shows researchers, first, that the targeted neuronal activity is indeed a causal step in the circuit, and second, what that causal step does for the circuit. What the causal step does for the circuit—i.e., the causal role of the step within the circuit—is determined by investigation of circuit dynamics, or the processes that occur along the circuit. Circuit dynamics play an explanatory role in neuroscientific inquiry (Woodward 2023).

With circuit dissection techniques, circuits are decomposed into causal steps and functions are localized to those causal steps. Localization tells us what is carrying out the function. Decomposing systems into their component parts and localizing functions to those

parts is a common research strategy in biology and neuroscience (Bechtel and Richardson 2010). This strategy works for systems that are decomposable (or nearly decomposable), meaning that the component parts work in at least partial isolation from each other over the short term (see Simon 1962). Even if the brain is more of a dynamical system than a decomposable one—meaning that the activities of neurons, circuits, and other parts are highly interdependent and do not operate in isolation—decomposition and localization approaches are thought to be a good approximation for studying complex systems (Bechtel and Richardson 2010). In circuit dissection experiments, localization is done not to explain the circuit part but rather to explain the behavior of an experimental group of model organisms as compared to a control group (Cummins 1975; Mitchell 1993).

Localization of function to a part requires positing an interlevel relation by which the *fulfillment* of some function—usually processing of some stimulus input or some behavioral output—shares spatiotemporal properties with some brain area or neurons performing some activity (Wimsatt 2006). Determining which function to ascribe is part of functional analysis. On a well-received view, "to ascribe a function to something is to ascribe a capacity to it which is *singled out* by its role in an analysis of some capacity of a containing system" (Cummins 1975, 765, my emphasis). This *singling out* of the relevant capacity is of issue here. I follow causal role theories of functions in conceiving of the fulfillment of functions as having a causal effect. Although functions are often understood as purposive (evolutionarily-backed or goal-driven), causal role functions instead derive their teleology from being situated within a containing system for which they perform the relevant role (Craver 2012). In cognitive neuroscience, the containing system is often the behaving organism. Failure to perform the relevant role counts as dysfunction. Importantly, particular functions of circuit parts need not be cognitive; noncognitive

functions—those for maintaining or repairing biological systems, for example—will be appropriate ascriptions to circuit parts in many cases (Haueis 2018). I set non-cognitive functions aside here as ascribing these functions will often look more like explaining the circuit part or activity rather than explaining experimental behavior.

The challenge is to determine which one of various candidate cognitive functions to ascribe to the circuit part. Consider the diagram in Figure 1. Suppose researchers are interested in explaining memory, and to do so need to find the components storing the memory. The researchers target the trisynaptic circuit of the hippocampus and intervene on the dentate gyrus, which we might say is C_2 in the diagram. The distal input to the circuit is information about some stimulus or task variable, and the output of the circuit will be measured by a distal output, behavior. (Proximal inputs and outputs to the circuit, although targets of inquiry themselves, are idealized as information about the stimulus input and behavioral output.) Based on the intervention on the dentate gyrus alone, it should prove methodologically impossible to sort out the contribution of the dentate gyrus within the circuit and for the system without already knowing the roles of the other components. First, the input has already been processed by upstream components. While an intervention on C_2 should screen off the influence of C_1, we must already know what C_1 did in the circuit to know what input it provides to C_2. Second, the functional output is downstream and involves processing by the other components. Inferring back to the role of C_2, then, requires sorting out what contributions the other parts are making. Crucially, researchers lack interventive control to hold activity at these other causal steps fixed. How can we determine whether the dentate gyrus is performing the relevant role?

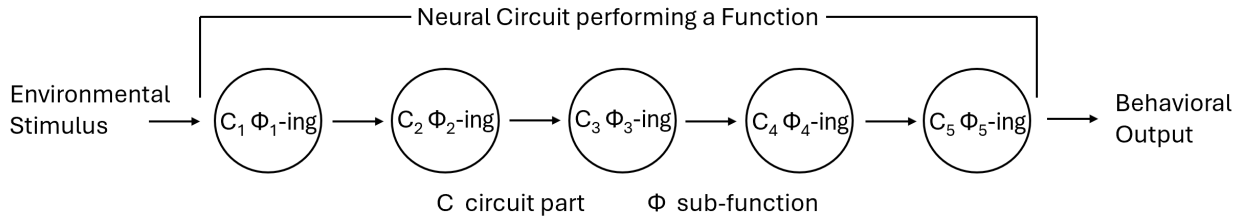

Figure 1 Adapted from (Craver, Glennan, and Povich 2021)

In the next section, I describe two kinds of function ascription practices, highlighting the challenges and shortcomings of each.

II. Function Ascription Practices

Standard function ascription practice is guided by consideration of the behavior observed in experimental results, the experimental paradigm used, and authoritative research informing the interpretation of experimental results. For example, authoritative research has it that the amygdala is specialized for negative reward learning. This background knowledge provides a target of intervention for researchers investigating fear as well as a simplifying strategy for interpreting the results of interventions on the amygdala. The presumption of function localization—i.e., that the amygdala functions to process fear—guides interpretation of experiments results (Hardcastle & Stewart 2002). Moreover, fear conditioning paradigms are designed to elicit fear behaviors in experimental animals according to standard protocols. In particular experiments that aim to target the function of a part for cognition or behavior, functions are ascribed in two ways: either by appeal to function ascriptions in other research or by subtractive analysis. I consider each in turn, focusing on function ascriptions to the dentate gyrus.

First, consider function ascriptions made by appeal to background authoritative research. Researchers in the Tonegawa lab optogenetically manipulated cells in the dentate gyrus that were active while mice were in a novel context (Liu et al. 2012; Ramirez et al. 2013). As stated in the research report, the dentate gyrus was chosen because previous research had established the existence of memory engrams—biologically stored memories—in the dentate gyrus (see also Robins 2018). In the experiment, the targeted cells were activated during the formation of a fear conditioning memory in a different context. Researchers found that when mice were placed back in the original context, they exhibited freezing behavior, which under natural conditions would be unexpected since fear conditioning did not take place in that context. To explain their results, the researchers claimed to have found and manipulated memory engrams for the original context that were falsely associated with fear conditioning. The ascribed function of the circuit part in the dentate gyrus, then, is memory storage, based on previous research and the success of their experiment. The same procedures were successfully conducted using chemogenetics (Garner et al. 2012). Importantly, engram function is ascribed to the neurons in the dentate gyrus because researchers were looking for the physical substrate that performs that function. As Robins says it, "there is a commitment to the idea that there is an engram, which is prior to any discoveries about the engram" (2023, 1). Researchers are confident they found engrams because their interventions took place where other researchers claim to have found them. Despite some worries about looming circularity (see, e.g., Rothman 2002), such circularity is pervasive and often benign, so I set this aside here.

Circuit interventions follow this basic inference structure: First, researchers identify some phenomenon they are interested in, such as memory storage. Some influential research has previously identified a circuit component—say neurons in the dentate gyrus—that is implicated

in behaviors that are a measure of memory storage. In this case, freezing behavior is a measure of memory storage in the contextual fear conditioning paradigm. So, neurons in the dentate gyrus are targeted. When the neurons are activated in a contextual fear conditioning paradigm, fear related behavior occurs. Thus, the reasoning goes, the neurons function to store memory. In short, the idea is that previous research has shown that this particular brain region BR functions to Φ, so BR must be Φ-ing here. Much of this inferential structure rests on previous localizations, that the experimental paradigm elicits the appropriate function, and that the behavior is a correct measure of the function.

Second, consider function ascription by subtractive analysis. Subtractive analysis refers to three kinds of analyses. The first kind of subtractive analysis, which I call *function-eliminating* subtractive analysis, is similar to the appeal to authoritative research. With this method, researchers have candidate functions in mind and attempt to determine which function the part is performing. The second is *experiment-based* subtractive analysis. As a canonical example, researchers correlate spike rates with object features by subtracting features of the object—i.e., by making the object round instead of square or by shifting the orientation of a bar—to rule out features of the object that may affect spiking. (Examples of this include Hubel &Wiesel's (1959) work in the visual cortex and Tanaka et al.'s (1991) work in the inferotemporal cortex.) For example, if a circuit part responds to a blue ball, researchers might determine whether the part also responds to a blue cube or a red ball to determine whether the part is responding to color or shape. (This example is simple enough to be illustrative, but we might doubt that function is being targeted in this case. I'll give an example below in which function is targeted directly in experiment-based subtractive analysis.) The third, which I call *target-based* subtractive analysis, paradigmatically occurs in fMRI research. Researchers analyze voxel activity by, so to speak,

subtracting voxel activity that is common across task conditions (Roskies 2010). Differences in voxel activity between task conditions are informative about where in the brain the relevant function is being performed. I'll examine each kind of subtractive analysis in turn.

Function-eliminating subtractive analysis is usually completed over numerous experiments in numerous labs for circuit interventions. In an ideal single experiment, researchers would hold some circuit parts stable while inhibiting others to determine which function is fulfilled or is no longer fulfilled. But such interventive control is impossible at present. Another issue, which I will return to below, is that the behavioral measures in experimental paradigms are often too coarse-grained to discriminate between all the various functions a part can perform. This leaves function-eliminating subtractive analysis with indeterminacies that seem irresolvable with present technology. To make function-eliminating subtractive analysis tractable in single experiments, researchers must arbitrarily constrain the pool of candidate functions to those they are interested in.

To illustrate this, consider the candidate functions of the dentate gyrus described in section II. Researchers used a contextual fear conditioning paradigm, which involves mice learning to fear a particular context based on foot shocks received in that context. Now a number of capacities are required for mice to learn to fear a particular context, including perception of the context and foot shocks, the ability to store information about each and their association, as well as, for researchers' purposes, the ability to behaviorally express fear. Background knowledge is certainly required to constrain the possible capacities performed by the dentate gyrus, or else researchers would be overwhelmed with complexity. But the dentate gyrus is thought to perform more than seven functions, many of which are compatible with each other (Borzello et al. 2023). As it happens, at least six putative functions are consistent with the

experimental results described in section II, and two of these functions—storage of memory and indexing for cortically stored memories—are incompatible with each other. So it is hypothesized, based on a wealth of research but also quite arbitrarily, that the dentate gyrus performs some memory role. Even with such constraints, the determinate function in the above experiments could be storage of the complete memory, storage of part of the memory, indexing of a cortically stored memory for retrieval, pattern separation to distinguish memories, pattern completion to retrieve memories or generalize them, or some less central, modulatory role so that another structure can perform these functions. Researchers often need to begin with an experimentally tractable number of candidate functions to constrain the inference space. For example, Bernier et al. (2017) were interested in determining whether the dentate gyrus plays a role in memory acquisition or memory retrieval. These researchers optogenetically inhibited the dentate gyrus during different aspects of the contextual fear conditioning paradigm to differentiate between acquisition and retrieval roles. In particular, they inhibited the dentate gyrus during the training phase of the paradigm to make inferences about acquisition, finding that the dentate gyrus is necessary for memory acquisition. Surprisingly, inhibitions of the dentate gyrus during retrieval in the relevant fear conditioning context did not prevent memory retrieval.

The other two kinds of subtractive analysis similarly require a stipulation of candidate functions. *Experiment-based subtractive analysis* involves manipulation of the experimental paradigm while holding fixed the part and candidate functions. Bernier et al. (2017) used this kind of subtractive analysis as well. Initially, inhibition of the dentate gyrus while experimental mice were in the conditioning context after training seemed to suggest that it did not play a role in memory retrieval. This is because experimental mice did not exhibit freezing behavior indicative of memory retrieval at the level that would be expected if mice successfully

remembered receiving foot shocks in that context. But Bernier et al. manipulated the paradigm to discriminate between memory retrieval and another related function: context discrimination. Context discrimination is a kind of pattern separation the dentate gyrus putatively does to distinguish between contextual memories for two different contexts. Context discrimination is a more precise form of memory retrieval. Researchers pre-exposed some mice to a neutral context similar to the one in which they were fear conditioned. Mice that were not pre-exposed to this neutral context exhibited more freezing behavior than controls *in the neutral context* when the dentate gyrus was inhibited. In contrast, mice that were pre-exposed to the neutral context had reduced freezing when the dentate gyrus was inhibited *in the conditioning context*. Bernier et al. interpret this difference as indicative of a failure to discriminate the contexts upon retrieval when the dentate gyrus is inhibited. In other words, the mice pre-exposed to the neutral context overgeneralized the neutral context, and the mice who were not pre-exposed to the neutral context overgeneralized the conditioning context. This difference between exposure groups indicated that the dentate gyrus plays a more particular retrieval role in context discrimination.

Target-based subtractive analysis aims to identify the parts that are performing various functions. For example, Sun et al. (2020) used optogenetics to show that subpopulations of memory engrams in the dentate gyrus have distinct functions. In this study, researchers drew on background research to propose memory generalization and discrimination as finer functions of memory engrams in the dentate gyrus. They then dissociated two subpopulations in dentate gyrus engrams by their distinct genetic markers and proposed that each population performs a subfunction. Here, these researchers are ascribing a function to each of these two subpopulations based on dissociated lower-level properties (akin to differences in voxel activity) and behavioral

outputs (akin to different task conditions) for each population. Optogenetic treatment of the sub-ensembles suggested evidence in favor of their proposal.

One benefit of subtractive analysis over appeals to authoritative research is that it targets functions directly. But challenges remain for constraining the space of candidate functions to discriminate in experimental analyses. While subtractive analysis can be used to discriminate among various candidate functions, where those candidate functions come from is a different matter. Often the space of candidate functions is arbitrarily constrained according to researchers' interests, past research, and current controversies. Science then proceeds by working out which functions ascriptions fare better for certain parts given the experimental paradigm, results of intervention, and previous research. There may be better or worse ways to constrain the space of candidate functions, an investigation of which cannot be undertaken here. For now, subtractive analysis may be better described as a heuristic for managing uncertainty about functions.

III. Taking Stock of Function Ascription Challenges

In the previous section, I argued that subtractive analysis is a means of explicitly targeting function in experimentation. But making subtractive analysis tractable requires arbitrarily constraining the class of candidate functions to analyze. This makes function ascription by appeal to background research on the relevant circuit part unavoidable. That is, when researchers ask what a circuit part does or what function it performs for the system, part of the answer must be already assumed rather than determined experimentally. The reason for this is that functions are not observable or directly measurable; they must be inferred through the kind of eliminative induction described here. And the space of candidate functions for eliminative induction is already partly determined by authoritative research and what the researchers are interested in explaining. This is what makes ascribing functions so challenging.

In diagnosing the challenges of function ascription and localization, philosophical accounts have generally pointed to multifunctionality of brain parts and context-dependent functions, as well as multiplexing of functions in the brain. Outside of fMRI research, there has been little work explicitly done to analyze the strategies that researchers adopt for making such properties of brain functions experimentally tractable despite the lack of interventive control and observability. Doing this work with other kinds of experiments here has provided us with a refined understanding of the challenges facing function ascriptions. Here I identified two additional challenges. The first is a lack of interventive control over non-targeted parts of the circuit which would allow inference from experimental behavior to the part's function to be more exact. Second, researchers need to appeal to background research to constrain the pool of candidate functions their experiment will discriminate among. Yet, these constraints are often insufficient because there are many candidate functions to choose among. These are challenges because functions are not observable or directly measurable. Researchers must make inferences about the function of a certain part based on the experimental paradigm, authoritative background research, and experimental results.

These inferences are best done with subtractive analysis. With subtractive analysis, although functions are not observable, they are directly targeted in the experiment. Research uses subtractive analysis to discriminate among a class of candidate functions experimentally. While subtractive analysis cannot dissolve the challenges I identified above, it allows researchers to make progress with function ascriptions. Some may be pessimistic about function ascriptions as a result, heralding the arbitrary and stipulative constraints that are required to achieve experimental tractability for investigating functions. At its strongest, the claim might be that our current function ascriptions are at present unwarranted. Others may be more optimistic,

appealing to intuitions that these practices constitute everyday science that continues to be successful. So long as experiments are conducted in order to discriminate between various functions, such experiments may be part of an iterative process of refining the ascriptions given to particular circuits parts. I will not take sides here, as I am undecided myself. But I think there are many reasons to be optimistic.

Instead, I will make a proposal for strengthening subtractive analyses. I suspect that many failures to discriminate functions are the result of impoverished experimental paradigms. These paradigms are simple enough to execute consistently and standardized to allow findings to generalize across laboratories. Both are certainly virtues of these paradigms. But these virtues are at the cost of the kind of complexity that will make subtractive analysis valuable. Altering features of the paradigm as well as when and how interventions occur can provide means for distinguishing various candidate roles of any brain part intervened into. Rather than standardization of paradigms for generalization, manipulations of paradigms may make progress in justifying function ascriptions.

IV. Conclusion

To understand cognition and behavior, neuroscientists often investigate circuits at various scales. Relatively recent innovations have allowed researchers to intervene into mesoscale circuits to investigate the function of these circuits and their parts. Yet with this new methodology, function ascriptions remain challenging. Aside from multifunctionality of circuit parts and multiplexing of circuit functions, I have identified two additional challenges for researchers: a lack of interventive control over other parts of the circuit and insufficient constraints on the class of candidate functions to discriminate them experimentally. All of these challenges are due to the fact that circuit functions are not directly measurable. Function

ascriptions are made on the basis of a number of inferences. I argued that these inferences are stronger as a result of subtractive analyses.

Works Cited

Barack, David. 2024. "Context and Neural Function." *Philosophy of Science* 00: 1–19.

Bechtel, William, and Robert Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. The MIT Press.

Bechtel, William, and Jennifer Mundale. 1999. "Multiple Realizability Revisited: Linking Cognitive and Neural States." *Philosophy of Science* 66 (2): 175–207.

Bernier, Brian, Anthony Lacagnina, Adam Ayoub, Francis Shue, Boris Zemelman, Franklin Krasne, and Michael Drew. 2017. "Dentate Gyrus Contributes to Retrieval as well as Encoding: Evidence from Context Fear Conditioning, Recall, and Extinction." *The Journal of Neuroscience* 37 (26): 6359–6371.

Borzello, Mia, Steve Ramirez, Allessandro Treves, Inah Lee, Helen Scharfman, Craig Stark, James Knierim, and Lara Rangel. 2023. "Assessments of Dentate Gyrus Function: Discoveries and Debates." *Nature Reviews Neuroscience* 24: 502–517.

Burnston, Daniel. 2016. "A Contextualist Approach to Functional Localization in the Brain." *Biology and Philosophy* 31: 527–50.

Burnston, Daniel. 2021. "Getting Over Atomism: Functional Decomposition in Complex Neural Systems." *The British Journal for the Philosophy of Science* 72 (3): 743–772.

Cao, Rosa. 2022. "Multiple Realizability and the Spirit of Functionalism." *Synthese* 200 (506): 1–31.

Chirimuuta, M. 2024. *The Brain Abstracted*. MIT Press.

Craver, Carl. 2012. "Functions and Mechanisms: A Perspectivalist View." In *Functions: Selection and Mechanisms*, edited by Philippe Huneman, 133–158. Dordrecht: Springer.

Craver, Carl, Stuart Glennan, and Mark Povich. 2021. "Constitutive Relevance & Mutual Manipulability Revisited." *Synthese* 199: 8807–8828.

Cummins, Robert. 1975. "Functional Analysis." *Journal of Philosophy* 72: 741–765.

Figdor, Carrie. 2010. "Neuroscience and the Multiple Realization of Cognitive Functions." *Philosophy of Science* 77 (3): 419–456.

Garner, Aleena, David Rowland, Sang Youl Hwang, Karsten Baumgaertel, Bryan Roth, Cliff Kentros, and Mark Mayford. 2012. "Generation of a Synthetic Memory Trace." *Science* 335: 1513–1516.

Griffiths, P. E. 1993. "Functional analysis and proper functions." *British Journal for the Philosophy of Science* 44: 409–422.

Hardcastle, Valerie, and C. Matthew Stewart. 2002. "What Do Brain Data Really Show?" *Philosophy of Science* 69 (S3): S72–S82.

Haueis, Philip. 2018. "Beyond cognitive myopia: a patchwork approach to the concept of neural function." *Synthese* 195: 5373–5402.

Hubel, D., and T. Wiesel. 1959. “Receptive Fields of Single Neurones in the Cat’s Striate Cortex.” *Journal of Physiology* 148: 574–591.

Klein, Colin. 2012. “Cognitive Ontology and Region- versus Network-Oriented Analyses.” *Philosophy of Science* 79: 952–960.

Klein, Colin. 2017. “Brain Regions as Difference Makers.” *Philosophical Psychology* 30 (1–2): 1–20.

Liu, Xu, Steve Ramirez, Petti Pang, Corey Puryear, Arvind Govindarajan, Karl Deisseroth, and Susumu Tonegawa. 2012. “Optogenetic Stimulation of a Hippocampal Engram Activates Fear Memory Recall.” *Nature* 484 (7394): 381–385.

Machamer, Peter, Lindley Darden, and Carl Craver. 2000. “Thinking About Mechanisms.” *Philosophy of Science* 67 (1): 1–25.

McCaffrey, Joseph. 2015. “The Brain’s Heterogeneous Functional Landscape.” *Philosophy of Science* 82: 1010–1022.

McCaffrey, Joseph. 2023. "Evolving Concepts of Functional Localization." *Philosophy Compass* 18 (e12914): 1–17.

Mitchell, Sandra. 1993. “Dispositions or Etiologies? A Comment on Bigelow and Pargetter.” *The Journal of Philosophy* 90 (5): 249–259.

Mitchell, Sandra. 2003. *Biological Complexity and Integrative Pluralism.* Cambridge University Press.

Poldrack, Russ. 2006. “Can Cognitive Processes be Inferred from Neuroimaging Data?” *Trends in Cognitive Sciences* 10 (2): 59–63.

Ramirez, Steve, Xu Liu, Pei-Ann Lin, Junghyup Suh, Michele Pignatelli, Roger Redondo, Tomás Ryan, and Susumu Tonegawa. 2013. “Creating a False Memory in the Hippocampus.” Science 341 (6144): 387–391.

Rathkopf, Charles. 2013. "Localization and Intrinsic Function." Philosophy of Science 80 (1): 1–21.

Robins, Sarah. “Memory and Optogenetic Intervention: Separating the Engram from the Ecphory.” *Philosophy of Science* 85 (5): 1078–1089.

Robins, Sarah. 2023. “The 21st Century Engram.” *WIREs Cognitive Science* 14 (5): 1–15.

Roskies, A. 2010. “Saving Subtraction: A Reply to Van Orden and Paap.” The British Journal for the Philosophy of Science 61 (3): 635–665.

Ross, Lauren. 2021. “Causal Concepts in Biology: How Pathways Differ from Mechanisms and Why It Matters.” *The British Journal for the Philosophy of Science* 72 (1): 131–158. https://doi.org/10.1093/bjps/axy078

Ross, Lauren, and James Woodward. “Circuits and Circuit Manipulations in Neuroscience.” Paper presented at the Philosophy of Science Biennial Meeting, New Orleans, LA, November 15, 2024.

Rothman, Stephen. 2002. *Lessons from the Living Cell: The Limits of Reductionism.* McGraw-Hill.

Simon, Herbert. 1962. “The architecture of complexity.” *Proceedings of the American Philosophical Society* 106 (6): 467–482.

Sun, Xiaochen, Max J. Bernstein, Meizhen Meng, ..., Xiaohui Zhang, Polina O. Anikeeva, Yingxi Lin. 2020. “Functionally Distinct Neuronal Ensembles within the Memory Engram.” *Cell* 181: 410–423.

Tanaka, K., H. Saito, Y. Fukada, and M. Moriya. 1991. “Coding Visual Images of Objects in the Inferotemporal Cortex of the Macaque Monkey.” *Journal of Neurophysiology* 66 (1): 170–189.

Warren, Brandon, Michael Mendoza, Fabio Cruz, Rodrigo Leao, Daniele Caprioli, F. Javier Rubio, Leslie Whitaker, Kylie McPherson, Jennifer Bossert, Yavin Shaham, and Bruce Hope. 2016. “Distinct Fos-Expressing Neuronal Ensembles in the Ventromedial Prefrontal Cortex Mediate Food Reward and Extinction Memories.” *Journal of Neuroscience* 36 (25): 6691–6703.

Wimsatt, William. 2006. "Reduction and Its Heuristics: Making Methodological Reductionism Honest." *Synthese* 151: 445–475.

Woodward, James. 2023. "Networks, Dynamics and Explanation." UNSPECIFIED. URL: https://philsci-archive.pitt.edu/id/eprint/22694 (accessed 2024-08-10).