



# HHS Public Access

Author manuscript

*Curr Opin Neurobiol.* Author manuscript; available in PMC 2019 April 01.

Published in final edited form as:

*Curr Opin Neurobiol.* 2018 April ; 49: 84–94. doi:10.1016/j.conb.2018.01.004.

## The promise and perils of causal circuit manipulations

Steffen B. E. Wolff<sup>1</sup> and Bence P. Ölveczky<sup>1</sup>

<sup>1</sup>Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, Cambridge MA, USA 02138

### Abstract

The development of increasingly sophisticated methods for recording and manipulating neural activity is revolutionizing neuroscience. By probing how activity patterns in different types of neurons and circuits contribute to behavior, these tools can help inform mechanistic models of brain function and explain the roles of distinct circuit elements. However, in systems where functions are distributed over large networks, interpreting causality experiments can be challenging. Here we review common assumptions underlying circuit manipulations in behaving animals and discuss the strengths and limitations of different approaches.

### Introduction

A major goal of systems neuroscience is to explain the neural underpinnings of behavior. We believe a necessary first step is to describe experimentally tractable behaviors in terms of the algorithms and computations that underlie them [1,2]. The neural implementations of these processes can then be interrogated by recording and manipulating activity in targeted circuits and cell types, and by relating the experimental results to mechanistic models of how the probed functionality could be implemented in neural hardware [3].

Neural recordings play an important role in this endeavor by providing descriptions of brain dynamics that can reveal correlations with various aspects of behavior or sensory stimuli. The neural representations or coding schemes thus discovered can be evocative and inspire theories about the function of the probed circuits (e.g. ‘place cells’ in hippocampus [4], ‘mirror neurons’ in cortex [5]). However, the extent to which such correlations reflect *causal* contributions to the process under study is harder to gauge [6].

The concept of causality (i.e. cause and effect) in neuroscience is perhaps most intuitively grasped by thinking of the brain as a mechanistic system in which independent components with distinct functions interact to generate behavioral output [7] (Box 1). Besides offering tangible explanations, mechanistic models also speak to how a system can be controlled and

#### CONFLICTS OF INTEREST

NONE

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

manipulated [8] - an important aspect for many neuroscientists for whom the end game is to intervene in neurobiological processes gone awry [9,10].

### BOX 1

#### The pained but consequential nomenclature of causality

The notion of causality and the interpretations of causality experiments can quickly become ambiguous and muddled by vague terminology. While superficially a semantic problem, imprecise or misconceived terms can interfere with our thinking, models, and experimental design.

The concept of *causality* itself has been the subject of intense debates in biology, physics, philosophy, and other fields [8]. In neuroscience, causality is most commonly described in terms of the relationship between neuronal activity and behavior: i.e. activity patterns in a distinct part of the brain (cells, circuits, brain areas) directly control an aspect of behavior. However, what is meant by such causal statements is often unclear. Is a neuron that merely relays commands to downstream circuits causal to the resulting behavior or not? More differentiated statements require additional classifications, like the distinction between *necessary* and *sufficient* which has been thoroughly explored in epidemiology, oncology, and genetics [134–136].

In genetics, *loss-* and *gain-of-function* experiments to identify *necessary* (function not possible without) and *sufficient* (can trigger function by itself) genes for a given function are commonplace. *Rescue experiments* are typically done to test for the specificity of destructive manipulations. In neuroscience, silencing and activating certain parts of the brain and determining the effects on behavior are often considered tests of necessity and sufficiency respectively, but rescue experiments are difficult and rarely done (e.g. activating axon terminals in a presumed target area while silencing the cell bodies). As discussed in the text, the interpretation of such experiments may not be straightforward, and will depend on the complexity of the studied system and the nature of the manipulation.

The distinction between *instructive* and *permissive* provides another perspective on causality [137,138]. Instructive neural elements are those whose activities contribute uniquely to computations crucial to the function of interest. Activity in a permissive part of the brain, on the other hand, is acutely necessary for a specific function, but does not contribute unique information or signal processing for it (i.e. the system can recover from loss of permissive activity without adding any new information to it).

In face of all these different aspects of causality, having clearly defined terminology is important to avoid misconceptions and ambiguities and to allow for meaningful discussions, interpretations, and conclusions of causality experiments in neuroscience.

Establishing and verifying mechanistic models requires demonstrating causal links between neural activity and behavior (Figure 1). Thankfully, optogenetics and other recently developed circuit manipulation tools are making it possible to control the activity of circuits and cell types with increasing specificity, something that has catalyzed an exciting new era

in systems neuroscience [11–14]. But in embracing these new methods, we must recognize that they are sharp tools that should be used with caution [15,16]. Indeed, how specific perturbations of neural activity affect the function of complex and interconnected neural networks, and how such experiments inform mechanistic models of brain function are often far from obvious. While there are no universal rules for designing and interpreting causality experiments in neuroscience, being cognizant of the complexities involved and explicit about the assumptions underlying particular experiments is likely to improve the utility of these remarkable technologies. The purpose of our review is to highlight some issues we believe are important and relevant for neural circuit manipulations in behaving animals.

## Are neural systems decomposable into functional modules?

The goal of most causality experiments is to attribute functions to specific neural components, be they circuits or cell types. The underlying assumption is that the brain is a modular system with localized functions [17,18]. If this is indeed a good approximation, we should be able to study individual components of the brain in isolation, determine their respective functions, and then explain how complex behaviors and processes are implemented in terms of how the components interact [19]. This widely adopted approach has been successful in explaining how relatively simple and specialized systems operate [20–22].

Take, for example, the frog's prey-capture behavior [23]. It relies on a fast and largely feed-forward circuit, with sensory processing 'modules' (retina/tectum) extracting important features of the sensory environment (is there prey?). The brainstem receives the results of these computations and makes a decision whether to trigger a pattern generator for prey capture. The system is well approximated by a box and arrow model where each box relates to an anatomically confined circuit with a specific function, and the arrows denote the (unidirectional) causal interactions between the modules (Figure 2, left and middle).

Models of this kind are satisfying because they align with our mechanistic worldview and mirror the design of our own man-made systems, whether machines or electronic circuits. However, there is a growing recognition that more complex and less specialized neural systems may not be easily decomposed into discrete modules executing sequential and causally linked operations [24,25]. Indeed, in circuits characterized by recurrence, feedback, and dense interconnectivity, system components are often interdependent to a degree that makes localization of function and the notion of causality less intuitive and explanatory (Figure 2, right).

In complex systems, neural and otherwise, functional principles may reveal themselves not by studying parts of the system in isolation, but rather by considering the rules that govern their dynamic interactions [26–29]. These interactions can give rise to emergent properties that must be considered to avoid mereological fallacies, i.e. ascribing to a part what only applies to the whole [30]. While there is no easy prescription for how to deal with emergent phenomena in neuroscience, dynamical systems theory [31] offers a framework for relating neural activity patterns to behavior in a way that abstracts from individual nodes in the network and describes the time-varying behavior of the system as a whole [32,33]. Recently,

Churchland and colleagues described the control function of motor cortex from a dynamical systems perspective [34], resolving apparent contradictions in studies taking a more neuron-centric view. In limiting their analysis to motor cortex, the implication was that this control functionality is, in large part at least, anatomically localized. This allowed the large scale box-and-arrow model of the motor system to be preserved, with one of the boxes (here: motor cortex) described as a dynamical system (Figure 2, middle).

The above discussion reminds us that the brain can be interrogated, understood, and modeled at different levels [18], and that the explanatory or descriptive framework most appropriate for one level may not be the best approximation for another [35]. Note that mechanistic explanations and dynamical systems descriptions are neither exclusive nor incompatible [24]; they simply represent complementary ways of analyzing a system (as do other types of models, including statistical and network-based models). Which of the various formalisms serve as the most useful abstraction in any given case will depend on a variety of factors, including the nature of the explanation sought, the complexity of the system etc.

However, when it comes to circuit dissection experiments in which the function of a circuit or cell type is queried, there is really no substitute for having a plausible and concrete mechanistic model (or a set of models) of how the probed algorithm could be implemented. Such models allow us to hypothesize putative functions of the manipulated neural elements in ways that relate them to other parts of the circuit and ultimately to behavior. We can then test these hypotheses with causal manipulations, thus furthering our mechanistic understanding. If, for reasons discussed above, the system cannot be adequately reduced to a meaningful mechanistic model, experimental outcomes may remain phenomenological descriptions of what happens when a circuit is perturbed in a particular way, with the bridge to conceptual understanding awaiting further theoretical or experimental advances.

## Acute versus chronic manipulations

While the development of modern circuit dissection tools is revolutionizing neuroscience [9,12–14], the design and interpretation of these cutting-edge experiments are often informed by foundational studies using far ‘cruder’ perturbations – most notably lesions [36]. While it is expedient to think of acute and reversible neuronal silencing [37] (e.g. optogenetics, pharmacology) as a spatiotemporally precise version of lesions, there is a growing realization that acute perturbations and chronic silencing probe neural circuit function in fundamentally different ways [38–40]. Appreciating this distinction may clear up some confusion and allow meaningful comparisons across different manipulations, thus also building a better understanding of neural circuit function.

## Acute perturbations

While treating a complex neural system as a collection of distinct functional modules can be a convenient and often necessary simplification (Figure 2), we must not lose sight of the fact that the brain is also a non-linear dynamical system defined by the interactions of its many components [41,42]. This is particularly important to remember when performing sudden activity perturbations since their effects, however targeted, are likely to ripple through the interconnected brain and influence activity patterns also in remote circuits [39,43,44]

(Figure 3). It is this difficulty in confining the effects of specific perturbations to the targeted circuit that makes it hard to probe circuit function based on behavioral effects alone. This is particularly true if the consequence of a perturbation is the non-specific cessation of ongoing behavior (i.e. loss of function).

An unexpected and sudden perturbation could, for example, switch the brain into a dynamical state distinct from the one associated with the probed behavior [45], thus ‘activating’ other neural systems that generate competing behaviors. For example, if you are in the middle of knitting a sweater when perceiving a tingling sensation in your leg, you may stop the knitting to investigate what caused the sudden sensation. If it was the consequence of artificially perturbing neural activity in the leg region of your somatosensory cortex, it doesn’t follow that this part of your brain plays an essential role in knitting.

More generally, suddenly inhibiting (or exciting) a neural circuit or cell type can change how downstream circuits operate, including how they process inputs from other non-targeted areas [39,43,46] (Figure 3). In line with this, we recently found that inactivating sensorimotor nucleus Nif in songbirds disrupts vocal behavior by acutely suppressing activity in nucleus HVC, a premotor region downstream of Nif that controls the temporal progression of song [47,48]. However, after Nif was permanently inactivated, HVC spontaneously (i.e. without the bird practicing its song) recovered its pattern generator function and produced the same neural activity sequence (and song) as before the lesions. We concluded that normal activity in Nif is ‘permissive’ (Box 1), i.e. it is acutely required for HVC (and the song control system) to express its function but it is not necessary long-term [39]. Interestingly, the spontaneous recovery of HVC function after Nif lesions could be explained by homeostatic processes bringing HVC neurons back to their original firing rate set-point [39,49].

In contrast to the ‘permissive’ role we ascribe to Nif, neural activity in a circuit can be ‘instructive’ (Box 1) for a behavior if it contributes essential information or computation(s) [39]. Distinguishing whether activity in a circuit is permissive or instructive (or indeed irrelevant) is a prerequisite for a mechanistic understanding of how neural circuits implement algorithms and computations, yet this distinction can be difficult to make with acute activity perturbations alone (Figure 3, Box 1).

### Chronic perturbations

In contrast to fast reversible manipulations, acute behavioral effects of lesions are rarely considered (though see, for example, [39,50]). Rather, subjects are typically evaluated once the ‘non-specific’ effects have subsided [46], which can be days, weeks or even months (as in the case of humans) after the lesion. A significant fraction of these transient ‘non-specific’ effects are likely due to changes in network dynamics caused by the sudden silencing of a part of the brain, i.e. the very same effects described and analyzed during fast reversible manipulations [39]. By allowing time for homeostatic and other processes to bring the brain back to dynamic equilibrium, lesion studies probe the steady-state contributions of a brain area (Figure 3). However, potential off-target effects must be considered also for lesions. For example, structural and functional reorganization (homeostatic and otherwise) that follow

lesions could end up altering circuits remote from the lesioned area, thus potentially compromising their original function ('connectional diaschisis' [46]).

But the major reason why conventional lesions have ceded to acute and reversible manipulations as the favored method for interrogating circuit function may be the difficulties in interpreting a lack of behavioral effects after lesions [38,51]. That a particular behavior can be implemented without a brain area days or weeks after it was lesioned does not necessarily mean that it was not acutely instructive. This is because the brain can have redundancies, degeneracies, and the capacity for plasticity (Figure 3). Redundancy assures that if an essential part of the system fails a back-up stands ready to prevent system failure. While there is redundancy *within* specific brain circuits that make their function robust to the fractional loss of neurons experienced in degenerative disease [52,53], the extent to which functionally distinct circuits can cover for each other (degeneracy, [54]) is less clear and should not be generally assumed [55–58].

Besides redundancies and degeneracies, the brain can also compensate for lost functionality through experience-dependent plasticity [59] (Figure 3). Spared neural circuits could plausibly assume some of the functions of the lesioned one(s) or alternative solutions to a behavioral task could be learned. For example, after partial motor cortex lesions, other parts of motor cortex can rewire to 'cover' for the lost area [60]. Further, if most or all of motor cortex is gone, tasks that depended on it, such as skilled reaching, can be solved in new, albeit slower and less dexterous, ways [61], likely by engaging subcortical control circuits to a larger degree [62]. Importantly, such experience-dependent plasticity requires renewed engagement with the task [63]. One way to rule it out as an explanation for post-lesion recovery is to prevent lesioned animals from 'practicing' the probed behavior [39,64], though this may not always be possible, especially for innate and habitually expressed behaviors.

The above discussion highlights the strengths and weaknesses of both acute perturbations and chronic lesions (Figure 3). Lesion studies ask how a brain that has reached post-lesion equilibrium performs with a piece missing. But gleaning the functionality of the missing piece by studying the behavioral deficit is complicated by the fact that other parts of the brain could have compensated for functions initially carried by the lesioned area. While acute perturbations leave little time for such compensation, their interpretive difficulty lies in the brain being in an altered and unusual dynamical state with unknown consequences on overall network function.

Thus, rather than relying exclusively on one method, our quest to understand brain function may benefit from combining both acute and chronic manipulations with careful behavioral monitoring and electrophysiological recordings of network-wide consequences. The ultimate goal should be to explain the often disparate experimental outcomes [38–40] within a unifying mechanistic model. With the recent development of targeted chronic manipulations (e.g. by genetically- or anatomically-defined expression of Tetanus or Diphtheria toxin [65,66]), it is now possible to manipulate the very same neural circuit elements both acutely and chronically, making comparisons across the different types of perturbations more meaningful.



Interestingly, issues relating to the peculiarities of acute and chronic manipulations have cropped up in other fields of biology. In genetics, for example, there is an ongoing discussion about the use of mutants (chronic genetic manipulations) versus morpholinos (acute manipulations: transient antisense gene knockdown) in zebrafish, *Xenopus*, and other organisms [67–70]. While both approaches probe the role of targeted genes, interpreting the outcomes of such experiments is complicated by possible compensation and off-target effects. However, the field has seemingly embraced the strengths and weaknesses of each approach and acknowledged that they can be used in complementary ways to dissect the roles of specific genes and gene networks [67,71].

Every so often acute and chronic perturbations alter (without completely disrupting) a behavior in similar ways, and these cases may be particularly informative. An illustrative example comes from songbirds, where silencing the outflow nucleus of the song-specialized basal ganglia (LMAN) with lesions [72] and acute inactivations [73] produces the same outcome: a reduction in song variability. Given our discussion above, one might ask why the sudden silencing of LMAN does not cause larger and more disruptive effects on both network function and behavior. We believe this is because the system is tuned to expect - and function in - a scenario where LMAN activity is suddenly suppressed, which happens when the male sings to a female [74]. LMAN neurons provide excitatory drive to the downstream song-control region RA, which also receives excitatory input from HVC, the temporal pattern generator. However, balanced feed-forward excitation and feed-forward inhibition from LMAN to RA ensures that the overall excitability of RA projection neurons is robust to acute changes in LMAN activity, thus allowing the motor pathway (HVC-RA) to produce song both with and without LMAN active [75].

## Disruptive vs physiological perturbations

The above example emphasizes the power of using physiologically realistic and behaviorally relevant manipulations to probe neural circuit function. While sudden and unnatural perturbations of complex non-linear systems are often difficult to interpret (Figure 3), coaxing the system into dynamic regimes it naturally visits is likely to be less disruptive and more informative (Figure 4). Also, manipulation experiments are meant to probe ‘what-happens-if’ scenarios, and the most useful ones to explore are those that occur under normal conditions [44]. This is where modern manipulation techniques have the potential to change the game. Spatiotemporally precise, bi-directional manipulations of neuronal activity allow us, in principle at least, to reproduce behaviorally relevant activity patterns in specific circuits [76,77]. Exploring the brain’s natural dynamic regimes in controlled ways could allow for real gain-of-function experiments.

Several recent studies have shown the power of this approach to reveal causal relationships and inform mechanistic models. Given technical limitations, most successful examples thus far have involved neural systems in which behavioral variables are encoded by the overall firing rate of more-or-less homogenous neuronal populations (Figure 4). Perhaps the best understood examples of such systems are ‘command neurons’, classically identified in invertebrates and small vertebrates [78–80].

Bidirectional manipulations, now possible with optogenetics and chemogenetics, can dial up or down the activity of such neurons and explore the consequences on behavior. Studies in nematodes [81], flies [82–84], and zebrafish [85,86] have established the value of such experiments. A prominent example in rodents is the manipulation of the MLR, a brainstem nucleus controlling locomotion. Building on classic electrical stimulation experiments [87], recent optogenetic manipulations of MLR activity could increase and decrease the speed of locomotion in predictable ways [88,89]. Besides demonstrating a causal relationship between MLR activity and locomotion, the authors also identified the involved neuronal MLR subpopulation. Finally, they performed electrophysiological recordings from optogenetically identified neurons, showing that the effects of their manipulations matched the physiological activity of these neurons under normal conditions [89] (Figure 4). Other systems in which similar experiments have been performed include the freeze and flight systems of the central amygdala [90–92] and periaqueductal gray [93], amygdala-centered anxiety networks [94–96], and the hypothalamic circuits controlling aggression [97] and feeding [98,99]. There have also been encouraging developments in systems where the activity of specific neural populations contributes to computations without triggering immediate behavioral output. Examples include dissecting the role of interneuron subtypes in gating plasticity during fear learning [100,101], revealing the arithmetic of the dopamine prediction error circuit [102], and identifying VTA GABA projection neurons as key in communicating saliency during associative learning [103].

Acute circuit manipulations have also contributed to elucidating how memories are formed and stored [104–107]. Combining activity-dependent expression techniques with optogenetic and chemogenetic manipulations has allowed neuronal populations ‘active’ during a previous learning experience to be reactivated [108–113] or inhibited [114,115]. Such manipulation can trigger or suppress learned sensorimotor associations, indicating that the manipulated neurons are part of a memory network, or ‘engram’ [104–106]. However, because these manipulations do not mimic physiological activity patterns, interpretations of how the behavioral responses relate to the memory being probed remain debated [104–107].

More generally, simply increasing or decreasing the firing rate of neurons may not adequately probe circuits whose functions rely on precise spike-timing and complex network dynamics [77,116] (Figure 4). Such manipulations will generate non-physiological dynamics and likely lead to disrupted processing (i.e. loss of function), limiting the insights that can be gained (see discussion above). Performing behaviorally relevant perturbations in such systems has to start with describing, through neural recordings, the natural dynamical regimes of the system. Modern manipulation techniques could then be used to coax the system into making transitions between states by replaying physiological activity patterns at the right times. Due to uncertainties in each individual experiment, e.g. the number of neurons to manipulate, side-, off-target- and network-effects etc. [15,44], online monitoring of the induced effects will be crucial for interpreting these experiments. The promises and perils of manipulating neural systems within their natural regimes have been explored in a recent review [44] which gives some pointers on how to move forward.

Changing activity patterns in complex neural systems within their natural regimes can, in certain cases, be effectively accomplished with more ‘low-tech’ methods. Manipulating



temperature, for example, is a powerful way to probe how small changes in the speed of signal propagation in a network affect behavior – something not easily done with other manipulation techniques. It has been successfully used in songbirds to reveal the relationships between the tempo of neural activity sequences and behavior [117,118], and could be relevant also for a variety of other systems [119,120].

## Conclusion

New tools for manipulating neural activity in behaving animals have enabled researchers to control nervous system activity with unprecedented specificity and ease, transforming systems neuroscience in the process. Fulfilling the promise of these techniques requires us to tread carefully and mold our experiments and interpretations to the realities of the complex systems we are studying. While there are no fool-proof prescriptions for how to move forward, we believe that success will be contingent on carefully describing the behaviors we study in terms of the algorithms and computations that underlie them (Figure 1). Our current understanding of the system can then inform mechanistic models of how these algorithms are implemented in neural circuitry, models that can then be tested by manipulating activity patterns in different parts of the brain (Figure 2). Though these reductionist models often assume localization of function and independence of its various modules, in interpreting experimental outcomes of circuit manipulations we must be cognizant of the interconnected nature of the brain and the fact that its components interact over multiple temporal and spatial scales [39,40] (Figure 3). For this and other reasons, the most effective and informative manipulations may be those that mimic behaviorally relevant neural dynamics (Figure 4). While this is now being achieved in relatively simple systems, where our ability to control the overall firing rates of neurons comes close to replicating behaviorally relevant scenarios, as we venture into systems with more complex dynamics and coding schemes, there will be a need for more sophisticated monitoring and manipulation tools. Ongoing research to achieve more selective expression of effectors like opsins and DREADDs [14,121,122], to clarify their mechanisms [37,123], to engineer opsins with non-overlapping spectra [37,124,125], to miniaturize light guides or LEDs with integrated recording sites [126–128], and to develop patterned and holographic illumination [129–133]), will bring us closer to mimicking ‘normal’ neural activity in complex systems. Concurrently, there must also be an emphasis on developing analytic tools, theoretical frameworks, and mechanistic models to help us make sense of the emerging data and extract principles of function. Combining these efforts and approaches will allow us to probe our systems from disparate but complementary angles, thus advancing our understanding of how the brain generates behavior.

## Acknowledgments

We thank Sean Escola, James Fitzgerald, Grigori Guitchounts, Johannes Letzkus, Philip Tovote and the whole Ölveczky lab for helpful comments on the manuscript. This work was supported by a grant from NINDS (R01NS099323) to BPÖ and a HFSP fellowship to S.B.E.W.

## References

- 1\*\*. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*. 2017; 93:480–490. This review argues that detailed descriptions and analysis of behavior are crucial for understanding brain function and necessary for interpreting causal experiments. [PubMed: 28182904]
2. Marr, D. *Vision*. MIT Press; 1982.
3. Craver CF. When mechanistic models explain. *Synthese*. 2006; 153:355–376.
4. O'Keefe J, Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res*. 1971; 34:171–175. [PubMed: 5124915]
5. Rizzolatti G, Fadiga L, Gallese V, Fogassi L. Premotor cortex and the recognition of motor actions. *Cogn Brain Res*. 1996; 3:131–141.
6. Silvanto J, Pascual-Leone A. Why the Assessment of Causality in Brain–Behavior Relations Requires Brain Stimulation. *J Cogn Neurosci*. 2012; 24:775–777. [PubMed: 22264196]
7. Craver, CF., Darden, L. *In Search of Mechanisms*. 2013.
8. Woodward, J. *Making Things Happen A Theory of Causal Explanation - Oxford Scholarship*. 2004.
- 9\*. Rajasethupathy P, Ferenczi E, Deisseroth K. Targeting Neural Circuits. *Cell*. 2016; 165:524–534. This is an overview about different forms of modern circuit manipulation techniques, strategies for their targeting and their possible therapeutic use. [PubMed: 27104976]
10. Song C, Knöpfel T. Optogenetics enlightens neuroscience drug discovery. *Nat Rev Drug Discov*. 2015; doi: 10.1038/nrd.2015.15
11. Boyden ES. Optogenetics and the future of neuroscience. *Nat Neurosci*. 2015; 18:1200–1201. [PubMed: 26308980]
12. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nat Neurosci*. 2015; 18:1213–1225. [PubMed: 26308982]
- 13\*. Park HG, Carmel JB. Selective Manipulation of Neural Circuits. *Neurotherapeutics*. 2016; 13:311–324. A comprehensive overview of circuit manipulation techniques and targeting strategies. [PubMed: 26951545]
14. Sternson SM, Roth BL. Chemogenetic Tools to Interrogate Brain Functions. *Annu Rev Neurosci*. 2014; 37:387–407. [PubMed: 25002280]
- 15\*. Allen BD, Singer AC, Boyden ES. Principles of designing interpretable optogenetic behavior experiments. *Learn Mem*. 2015; 22:232–238. Considerations about the technical and interpretative challenges of optogenetic experiments, especially focusing on side effects. [PubMed: 25787711]
16. Häusser M. Optogenetics: the age of light. *Nat Methods*. 2014; 11:1012–1014. [PubMed: 25264778]
17. Kanwisher N. Functional specificity in the human brain: A window into the functional architecture of the mind. *Proc Natl Acad Sci*. 2010; 107:11163–11170. [PubMed: 20484679]
18. Meunier D, Lambiotte R, Fornito A, Ersche K, Bullmore ET. Hierarchical modularity in human brain functional networks. *Front Neuroinformatics*. 2009:3.
19. Miller EK, Phelps EA. Current Opinion in Neurobiology—Cognitive Neuroscience 2010. *Curr Opin Neurobiol*. 2010; 20:141–142. [PubMed: 20399637]
20. Borst A. Fly visual course control: behaviour, algorithms and circuits. *Nat Rev Neurosci*. 2014; 15:590–599. [PubMed: 25116140]
21. Heiligenberg, W. *Neural Nets in Electric Fish*. MIT Press; 1991.
22. Naumann EA, Fitzgerald JE, Dunn TW, Rihel J, Sompolinsky H, Engert F. From Whole-Brain Data to Functional Circuit Models: The Zebrafish Optomotor Response. *Cell*. 2016; 167:947–960e20. [PubMed: 27814522]
23. Ewert J-P. Neuroethology of releasing mechanisms: Prey-catching in toads. *Behav Brain Sci*. 1987; 10:337–368.
24. Tognoli E, Kelso JAS. The Metastable Brain. *Neuron*. 2014; 81:35–48. [PubMed: 24411730]
25. Uttal, W. *The New Phrenology*. MIT Press; 2001.
26. Anderson PW. More Is Different. *Science*. 1972; 177:393–396. [PubMed: 17796623]

27. Berdahl A, Torney CJ, Ioannou CC, Faria JJ, Couzin ID. Emergent Sensing of Complex Environments by Mobile Animal Groups. *Science*. 2013; 339:574–576. [PubMed: 23372013]
28. Gu M, Weedbrook C, Perales Á, Nielsen MA. More really is different. *Phys Nonlinear Phenom*. 2009; 238:835–839.
29. Marsh GE. The Demystification of Emergent Behavior. 2009 ArXiv09071117 Phys.
30. Bennett, MR., Hacker, PMS. *Philosophical Foundations of Neuroscience*. Wiley; 2003.
31. Katok, A., Hasselblatt, B. *Introduction modern theory dynamical systems | Differential and integral equations, dynamical systems and control*. Cambridge University Press; 1996.
32. Breakspear M, Jirsa VK. *Neuronal Dynamics and Brain Connectivity*. Handbook of Brain Connectivity. 2007
33. Ermentrout B. Neural networks as spatio-temporal pattern-forming systems. *Rep Prog Phys*. 1998; 61:353.
34. Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV. Neural population dynamics during reaching. *Nature*. 2012; 487:51–56. [PubMed: 22722855]
35. Tononi G, Sporns O, Edelman GM. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci*. 1994; 91:5033–5037. [PubMed: 8197179]
36. Damasio, H., Damasio, AR. *Lesion Analysis in Neuropsychology*. Oxford University Press; 1989.
37. Wiegert JS, Mahn M, Prigge M, Printz Y, Yizhar O. Silencing Neurons: Tools, Applications, and Experimental Constraints. *Neuron*. 2017; 95:504–529. [PubMed: 28772120]
38. Lomber SG. The advantages and limitations of permanent or reversible deactivation techniques in the assessment of neural function. *J Neurosci Methods*. 1999; 86:109–117. [PubMed: 10065980]
- 39\*\*. Otchy TM, Wolff SBE, Rhee JY, Pehlevan C, Kawai R, Kempf A, Gobes SMH, Ölveczky BP. Acute off-target effects of neural circuit manipulations. *Nature*. 2015; 528:358–363. This paper shows in both songbirds and rats that acute and chronic manipulations of neural activity in the same brain area can have very different effects and that such discrepancies can be explained by acute off-target effects of the manipulations. [PubMed: 26649821]
40. Goshen I, Brodsky M, Prakash R, Wallace J, Gradinaru V, Ramakrishnan C, Deisseroth K. Dynamics of Retrieval Strategies for Remote Memories. *Cell*. 2011; 147:678–689. [PubMed: 22019004]
41. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U. Complex networks: Structure and dynamics. *Phys Rep-Rev Sect Phys Lett*. 2006; 424:175–308.
42. Pillai AS, Jirsa VK. Symmetry Breaking in Space-Time Hierarchies Shapes Brain Dynamics and Behavior. *Neuron*. 2017; 94:1010–1026. [PubMed: 28595045]
43. Allen WE, Kauvar IV, Chen MZ, Richman EB, Yang SJ, Chan K, Gradinaru V, Deverman BE, Luo L, Deisseroth K. Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*. 2017; 94:891–907e6. [PubMed: 28521139]
- 44\*\*. Jazayeri M, Afraz A. Navigating the Neural Space in Search of the Neural Code. *Neuron*. 2017; 93:1003–1014. An important review dealing with the difficulties of interpreting neural activity manipulations and provides helpful pointers on how to conduct and interpret physiologically realistic circuit manipulations. [PubMed: 28279349]
45. Froehlich F, Sejnowski TJ, Bazhenov M. Network Bistability Mediates Spontaneous Transitions between Normal and Pathological Brain States. *J Neurosci*. 2010; 30:10734–10743. [PubMed: 20702704]
46. Carrera E, Tononi G. Diaschisis: past, present, future. *Brain*. 2014; 137:2408–2422. [PubMed: 24871646]
47. Hahnloser RH, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*. 2002; 419:65–70. [PubMed: 12214232]
48. Long MA, Jin DZ, Fee MS. Support for a synaptic chain model of neuronal sequence generation. *Nature*. 2010; 468:394–399. [PubMed: 20972420]
49. Keck T, Bener MH, Bonhoeffer T. Interactions between synaptic homeostatic mechanisms: an attempt to reconcile BCM theory, synaptic scaling, and changing excitation/inhibition balance. *Curr Opin Neurobiol*. 2017; 43:87–93. [PubMed: 28236778]

50. Newsome WT, Pare EB. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J Neurosci*. 1988; 8:2201–2211. [PubMed: 3385495]
51. Rorden C, Karnath H-O. Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nat Rev Neurosci*. 2004; 5:812–819.
52. Gómez-Isla T, Price JL, McKeel DW, Morris JC, Growdon JH, Hyman BT. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J Neurosci Off J Soc Neurosci*. 1996; 16:4491–4500.
53. Savioz A, Leuba G, Vallet PG, Walzer C. Contribution of neural networks to Alzheimer disease's progression. *Brain Res Bull*. 2009; 80:309–314. [PubMed: 19539730]
54. Tononi G, Sporns O, Edelman GM. Measures of degeneracy and redundancy in biological networks. *Proc Natl Acad Sci U S A*. 1999; 96:3257–3262. [PubMed: 10077671]
55. Betley JN, Cao ZFH, Ritola KD, Sternson SM. Parallel, Redundant Circuit Organization for Homeostatic Control of Feeding Behavior. *Cell*. 2013; 155:1337–1350. [PubMed: 24315102]
56. Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. *Nat Rev Neurosci*. 2015; 16:159–172. [PubMed: 25697159]
57. Glassman RB. An hypothesis about redundancy and reliability in the brains of higher species: Analogies with genes, internal organs, and engineering systems. *Neurosci Biobehav Rev*. 1987; 11:275–285. [PubMed: 3684057]
58. Li N, Daie K, Svoboda K, Druckmann S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature*. 2016; 532:459–464. [PubMed: 27074502]
59. Kleim JA, Jones TA. Principles of Experience-Dependent Neural Plasticity: Implications for Rehabilitation After Brain Damage. *J Speech Lang Hear Res*. 2008; 51:S225–S239. [PubMed: 18230848]
60. Nudo RJ, Wise BM, SiFuentes F, Milliken GW. Neural Substrates for the Effects of Rehabilitative Training on Motor Recovery After Ischemic Infarct. *Science*. 1996; 272:1791–1794. [PubMed: 8650578]
61. Whishaw IQ. Loss of the innate cortical engram for action patterns used in skilled reaching and the development of behavioral compensation following motor cortex lesions in the rat. *Neuropharmacology*. 2000; 39:788–805. [PubMed: 10699445]
62. Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, Kampff AR, Ölveczky BP. Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. *Neuron*. 2015; 86:800–812. [PubMed: 25892304]
63. Nudo RJ, Milliken GW. Reorganization of movement representations in primary motor cortex following focal ischemic infarcts in adult squirrel monkeys. *J Neurophysiol*. 1996; 75:2144–2149. [PubMed: 8734610]
64. Maldonado MA, Allred RP, Felthaus EL, Jones TA. Motor Skill Training, but not Voluntary Exercise, Improves Skilled Reaching After Unilateral Ischemic Lesions of the Sensorimotor Cortex in Rats. *Neurorehabil Neural Repair*. 2008; 22:250–261. [PubMed: 18073324]
65. Murray AJ, Sauer J-F, Riedel G, McClure C, Ansel L, Cheyne L, Bartos M, Wisden W, Wulff P. Parvalbumin-positive CA1 interneurons are required for spatial working but not for reference memory. *Nat Neurosci*. 2011; 14:297–299. [PubMed: 21278730]
66. Zhan C, Zhou J, Feng Q, Zhang J, Lin S, Bao J, Wu P, Luo M. Acute and Long-Term Suppression of Feeding Behavior by POMC Neurons in the Brainstem and Hypothalamus, Respectively. *J Neurosci*. 2013; 33:3624–3632. [PubMed: 23426689]
67. Blum M, De Robertis EM, Wallingford JB, Niehrs C. Morpholinos: Antisense and Sensibility. *Dev Cell*. 2015; 35:145–149. [PubMed: 26506304]
68. Kok FO, Shin M, Ni C-W, Gupta A, Grosse AS, van Impel A, Kirchmaier BC, Peterson-Maduro J, Kourkoulis G, Male I, et al. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell*. 2015; 32:97–108. [PubMed: 25533206]
69. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, Stainier DYR. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*. 2015; 524:230–233. [PubMed: 26168398]

70. Stainier DYR, Kontarakis Z, Rossi A. Making sense of anti-sense data. *Dev Cell*. 2015; 32:7–8. [PubMed: 25584794]
71. Schulte-Merker S, Stainier DYR. Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. *Development*. 2014; 141:3103–3104. [PubMed: 25100652]
72. Kao MH, Doupe AJ, Brainard MS. Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*. 2005; 433:638–643. [PubMed: 15703748]
73. Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol*. 2005; 3:e153. [PubMed: 15826219]
74. Kao MH, Wright BD, Doupe AJ. Neurons in a Forebrain Nucleus Required for Vocal Plasticity Rapidly Switch between Precise Firing and Variable Bursting Depending on Social Context. *J Neurosci*. 2008; 28:13232–13247. [PubMed: 19052215]
75. Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol*. 2011; 106:386–397. [PubMed: 21543758]
- 76\*. Grosenick L, Marshal JH, Deisseroth K. Closed-Loop and Activity-Guided Optogenetic Control. *Neuron*. 2015; 86:106–139. This review explores the opportunities as well as the technical and theoretical challenges of optogenetic closed loop stimulation experiments, which bring us closer to behaviorally relevant circuit manipulations. [PubMed: 25856490]
- 77\*\*. Phillips EA, Hasenstaub AR. Asymmetric effects of activating and inactivating cortical interneurons. *eLife*. 2016; 5:e18383. This paper shows that bidirectional acute optogenetic manipulations of specific circuit elements (here different interneurons) do not necessarily have bidirectional effects on circuit activity. Depending on the parameters, manipulations may push neuronal populations out of their physiological range of activity and have effects which do not reflect the physiological role of the probed circuit elements. [PubMed: 27719761]
78. Edwards DH, Heitler WJ, Krasne FB. Fifty years of a command neuron: the neurobiology of escape behavior in the crayfish. *Trends Neurosci*. 1999; 22:153–161. [PubMed: 10203852]
79. Hedwig B. Control of Cricket Stridulation by a Command Neuron: Efficacy Depends on the Behavioral State. *J Neurophysiol*. 2000; 83:712–722. [PubMed: 10669487]
80. Kupfermann I, Weiss KR. The command neuron concept. *Behav Brain Sci*. 1978; 1:3–10.
81. Nagel G, Brauner M, Liewald JF, Adeishvili N, Bamberg E, Gottschalk A. Light Activation of Channelrhodopsin-2 in Excitable Cells of *Caenorhabditis elegans* Triggers Rapid Behavioral Responses. *Curr Biol*. 2005; 15:2279–2284. [PubMed: 16360690]
82. Clyne JD, Miesenböck G. Sex-Specific Control and Tuning of the Pattern Generator for Courtship Song in *Drosophila*. *Cell*. 2008; 133:354–363. [PubMed: 18423205]
83. Lima SQ, Miesenböck G. Remote Control of Behavior through Genetically Targeted Photostimulation of Neurons. *Cell*. 2005; 121:141–152. [PubMed: 15820685]
84. Suh GSB, Ben-Tabou de Leon S, Tanimoto H, Fiala A, Benzer S, Anderson DJ. Light Activation of an Innate Olfactory Avoidance Response in *Drosophila*. *Curr Biol*. 2007; 17:905–908. [PubMed: 17493811]
85. Douglass AD, Kraves S, Deisseroth K, Schier AF, Engert F. Escape Behavior Elicited by Single, Channelrhodopsin-2-Evoked Spikes in Zebrafish Somatosensory Neurons. *Curr Biol*. 2008; 18:1133–1137. [PubMed: 18682213]
86. Wyart C, Bene FD, Warp E, Scott EK, Trauner D, Baier H, Isacoff EY. Optogenetic dissection of a behavioural module in the vertebrate spinal cord. *Nature*. 2009; 461:407–410. [PubMed: 19759620]
87. Shik ML, Severin FV, Orlovskii GN. Control of walking and running by means of electric stimulation of the midbrain. *Biofizika*. 1966; 11:659–666. [PubMed: 6000625]
88. Lee AM, Hoy JL, Bonci A, Wilbrecht L, Stryker MP, Niell CM. Identification of a Brainstem Circuit Regulating Visual Cortical State in Parallel with Locomotion. *Neuron*. 2014; 83:455–466. [PubMed: 25033185]
- 89\*\*. Roseberry TK, Lee AM, Lalive AL, Wilbrecht L, Bonci A, Kreitzer AC. Cell-Type-Specific Control of Brainstem Locomotor Circuits by Basal Ganglia. *Cell*. 2016; 164:526–537. This paper is an important example of how careful manipulations and recordings can be used to dissect



neural circuits which encode behavioral variables by the overall firing rate of a neuronal population. [PubMed: 26824660]

90. Ciocchi S, Herry C, Grenier F, Wolff SBE, Letzkus JJ, Vlachos I, Ehrlich I, Sprengel R, Deisseroth K, Stadler MB, et al. Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature*. 2010; 468:277–282. [PubMed: 21068837]
91. Fadok JP, Krabbe S, Markovic M, Courtin J, Xu C, Massi L, Botta P, Bylund K, Müller C, Kovacevic A, et al. A competitive inhibitory circuit for selection of active and passive fear responses. *Nature*. 2017; 542:96–100. [PubMed: 28117439]
92. Li H, Penzo MA, Taniguchi H, Kopec CD, Huang ZJ, Li B. Experience-dependent modification of a central amygdala fear circuit. *Nat Neurosci*. 2013; 16:332–339. [PubMed: 23354330]
93. Tovote P, Esposito MS, Botta P, Chaudun F, Fadok JP, Markovic M, Wolff SBE, Ramakrishnan C, Fenno L, Deisseroth K, et al. Midbrain circuits for defensive behaviour. *Nature*. 2016; 534:206–212. [PubMed: 27279213]
94. Botta P, Demmou L, Kasugai Y, Markovic M, Xu C, Fadok JP, Lu T, Poe MM, Xu L, Cook JM, et al. Regulating anxiety with extrasynaptic inhibition. *Nat Neurosci*. 2015; 18:1493–1500. [PubMed: 26322928]
95. Felix-Ortiz AC, Beyeler A, Seo C, Leppla CA, Wildes CP, Tye KM. BLA to vHPC inputs modulate anxiety-related behaviors. *Neuron*. 2013; 79:658–664. [PubMed: 23972595]
96. Tye KM, Prakash R, Kim S-Y, Fenno LE, Grosenick L, Zarabi H, Thompson KR, Gradinaru V, Ramakrishnan C, Deisseroth K. Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature*. 2011; 471:358–362. [PubMed: 21389985]
97. Lin D, Boyle MP, Dollar P, Lee H, Lein ES, Perona P, Anderson DJ. Functional identification of an aggression locus in the mouse hypothalamus. *Nature*. 2011; 470:221–226. [PubMed: 21307935]
98. Aponte Y, Atasoy D, Sternson SM. AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat Neurosci*. 2011; 14:351–355. [PubMed: 21209617]
99. Atasoy D, Betley JN, Su HH, Sternson SM. Deconstruction of a neural circuit for hunger. *Nature*. 2012; 488:172–177. [PubMed: 22801496]
100. Letzkus JJ, Wolff SBE, Meyer EMM, Tovote P, Courtin J, Herry C, Lüthi A. A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature*. 2011; 480:331–335. [PubMed: 22158104]
101. Wolff SBE, Gründemann J, Tovote P, Krabbe S, Jacobson GA, Müller C, Herry C, Ehrlich I, Friedrich RW, Letzkus JJ, et al. Amygdala interneuron subtypes control fear learning through disinhibition. *Nature*. 2014; 509:453–458. [PubMed: 24814341]
102. Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*. 2015; 525:243–246. [PubMed: 26322583]
103. Brown MTC, Tan KR, O'Connor EC, Nikonenko I, Muller D, Lüscher C. Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. *Nature*. 2012; 492:452–456. [PubMed: 23178810]
- 104\*. Josselyn SA, Köhler S, Frankland PW. Finding the engram. *Nat Rev Neurosci*. 2015; 16:521–534. This paper gives an important overview of the 'engram' field. [PubMed: 26289572]
105. Poo M-M, Pignatelli M, Ryan TJ, Tonegawa S, Bonhoeffer T, Martin KC, Rudenko A, Tsai L-H, Tsien RW, Fishell G, et al. What is memory? The present state of the engram. *BMC Biol*. 2016; 14:40. [PubMed: 27197636]
106. Tonegawa S, Liu X, Ramirez S, Redondo R. Memory Engram Cells Have Come of Age. *Neuron*. 2015; 87:918–931. [PubMed: 26335640]
- 107\*. Eichenbaum H. Still searching for the engram. *Learn Behav*. 2016; 44:209–222. This paper reflects upon the progress of the 'engram' field regarding declarative memory. [PubMed: 26944423]
- 108\*\*. Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S. Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature*. 2012; 484:381–385. This is the first paper using modern acute manipulation techniques to identify 'engram' neurons. [PubMed: 22441246]
109. Garner AR, Rowland DC, Hwang SY, Baumgaertel K, Roth BL, Kentros C, Mayford M. Generation of a synthetic memory trace. *Science*. 2012; 335:1513–1516. [PubMed: 22442487]



110. Redondo RL, Kim J, Arons AL, Ramirez S, Liu X, Tonegawa S. Bidirectional switch of the valence associated with a hippocampal contextual memory engram. *Nature*. 2014; 513:426–430. [PubMed: 25162525]
111. Ramirez S, Liu X, MacDonald CJ, Moffa A, Zhou J, Redondo RL, Tonegawa S. Activating positive memory engrams suppresses depression-like behaviour. *Nature*. 2015; 522:335–339. [PubMed: 26085274]
112. Ramirez S, Liu X, Lin P-A, Suh J, Pignatelli M, Redondo RL, Ryan TJ, Tonegawa S. Creating a false memory in the hippocampus. *Science*. 2013; 341:387–391. [PubMed: 23888038]
113. Kim J, Kwon J-T, Kim H-S, Josselyn SA, Han J-H. Memory recall and modifications by activating neurons with elevated CREB. *Nat Neurosci*. 2014; 17:65–72. [PubMed: 24212670]
114. Denny CA, Kheirbek MA, Alba EL, Tanaka KF, Brachman RA, Laughman KB, Tomm NK, Turi GF, Losonczy A, Hen R. Hippocampal Memory Traces Are Differentially Modulated by Experience, Time, and Adult Neurogenesis. *Neuron*. 2014; 83:189–201. [PubMed: 24991962]
115. Tanaka KZ, Pevzner A, Hamidi AB, Nakazawa Y, Graham J, Wiltgen BJ. Cortical representations are reinstated by the hippocampus during memory retrieval. *Neuron*. 2014; 84:347–354. [PubMed: 25308331]
- 116\*. Srivastava KH, Holmes CM, Vellema M, Pack AR, Elemans CPH, Nemenman I, Sober SJ. Motor control by precisely timed spike patterns. *Proc Natl Acad Sci*. 2017; 114:1171–1176. This paper shows that precise spike timing is important for the control of birdsong. [PubMed: 28100491]
117. Aronov D, Fee MS. Analyzing the dynamics of brain circuits with temperature: design and implementation of a miniature thermoelectric device. *J Neurosci Methods*. 2011; 197:32–47. [PubMed: 21291909]
118. Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*. 2008; 456:189–194. [PubMed: 19005546]
119. Mello GBM, Soares S, Paton JJ. A scalable population code for time in the striatum. *Curr Biol CB*. 2015; 25:1113–1122. [PubMed: 25913405]
120. Wang J, Narain D, Hosseini E, Jazayeri M. Flexible control of speed of cortical dynamics. *bioRxiv*. 2017; doi: 10.1101/155390
121. Fenno LE, Mattis J, Ramakrishnan C, Hyun M, Lee SY, He M, Tucciarone J, Selimbeyoglu A, Berndt A, Grosenick L, et al. Targeting cells with single vectors using multiple-feature Boolean logic. *Nat Methods*. 2014; 11:763–772. [PubMed: 24908100]
122. Urban DJ, Roth BL. DREADDs (Designer Receptors Exclusively Activated by Designer Drugs): Chemogenetic Tools with Therapeutic Utility. *Annu Rev Pharmacol Toxicol*. 2015; 55 null.
123. Gomez JL, Bonaventura J, Lesniak W, Mathews WB, Sysa-Shah P, Rodriguez LA, Ellis RJ, Richie CT, Harvey BK, Dannals RF, et al. Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science*. 2017; 357:503–507. [PubMed: 28774929]
124. Chuong AS, Miri ML, Busskamp V, Matthews GAC, Acker LC, Sørensen AT, Young A, Klapoetke NC, Henninger MA, Kodandaramaiah SB, et al. Noninvasive optical inhibition with a red-shifted microbial rhodopsin. *Nat Neurosci*. 2014; 17:1123–1129. [PubMed: 24997763]
125. Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, Morimoto TK, Chuong AS, Carpenter EJ, Tian Z, et al. Independent optical excitation of distinct neural populations. *Nat Methods*. 2014; 11:338–346. [PubMed: 24509633]
126. Buzsáki G, Stark E, Berényi A, Khodagholy D, Kipke DR, Yoon E, Wise KD. Tools for probing local circuits: high-density silicon probes combined with optogenetics. *Neuron*. 2015; 86:92–105. [PubMed: 25856489]
127. Royer S, Zemelman BV, Barbic M, Losonczy A, Buzsáki G, Magee JC. Multi-array silicon probes with integrated optical fibers: light-assisted perturbation and recording of local neural circuits in the behaving animal. *Eur J Neurosci*. 2010; 31:2279–2291. [PubMed: 20529127]
128. Wu F, Stark E, Ku P-C, Wise KD, Buzsáki G, Yoon E. Monolithically Integrated  $\mu$ LEDs on Silicon Neural Probes for High-Resolution Optogenetic Studies in Behaving Animals. *Neuron*. 2015; 88:1136–1148. [PubMed: 26627311]
129. Emiliani V, Cohen AE, Deisseroth K, Häusser M. All-Optical Interrogation of Neural Circuits. *J Neurosci Off J Soc Neurosci*. 2015; 35:13917–13926.

130. Packer AM, Peterka DS, Hirtz JJ, Prakash R, Deisseroth K, Yuste R. Two-photon optogenetics of dendritic spines and neural circuits. *Nat Methods*. 2012; 9:1202–1205. [PubMed: 23142873]
131. Packer AM, Russell LE, Dalgleish HWP, Häusser M. Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nat Methods*. 2015; 12:140–146. [PubMed: 25532138]
132. Papagiakoumou E, Anselmi F, Bègue A, de Sars V, Glückstad J, Isacoff EY, Emiliani V. Scanless two-photon excitation of channelrhodopsin-2. *Nat Methods*. 2010; 7:848–854. [PubMed: 20852649]
133. Linghu C, Tanese D, Boyden ES, Papagiakoumou E, Ronzitti E, Piatkevich K, Shemesh OA, Emiliani V, Zampini V. Temporally precise single-cell-resolution optogenetics. *Nat Neurosci*. 2017; 20:1796. [PubMed: 29184208]
134. Katsanis N. The continuum of causality in human genetic disorders. *Genome Biol*. 2016; 17:233. [PubMed: 27855690]
135. Marian A. Causality in Genetics: The Gradient of Genetic Effects and Back to Koch's Postulates of Causality. *Circ Res*. 2014; 114:e18–e21. [PubMed: 24436434]
136. Woodman SE, Mills GB. Are oncogenes sufficient to cause human cancer? *Proc Natl Acad Sci U S A*. 2010; 107:20599–20600. [PubMed: 21084639]
- 137\*. Calcott B. Causal specificity and the instructive–permissive distinction. *Biol Philos*. 2017; 32:481–505. This paper deals with the distinction between instructive and permissive causes in the context of development. It tests a model for the causal structure of biological development and suggests a hierarchical analysis approach.
138. Woodward J. Causation in biology: stability, specificity, and the choice of levels of explanation. *Biol Philos*. 2010; 25:287–318.

Mechanistic understanding of the brain requires assigning functions to neural modules

New sophisticated tools for manipulating neural activity advance this quest

Localizing functions to brain areas can be difficult in distributed neural systems

Acute and chronic activity manipulations probe circuit function in different ways

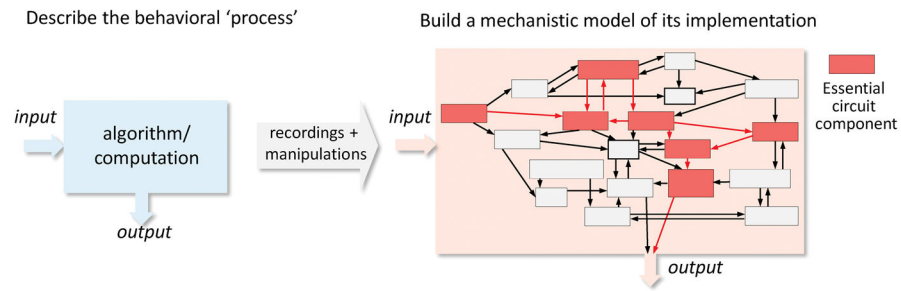
Future developments should focus on manipulations that mimic physiological activity

Author Manuscript

Author Manuscript

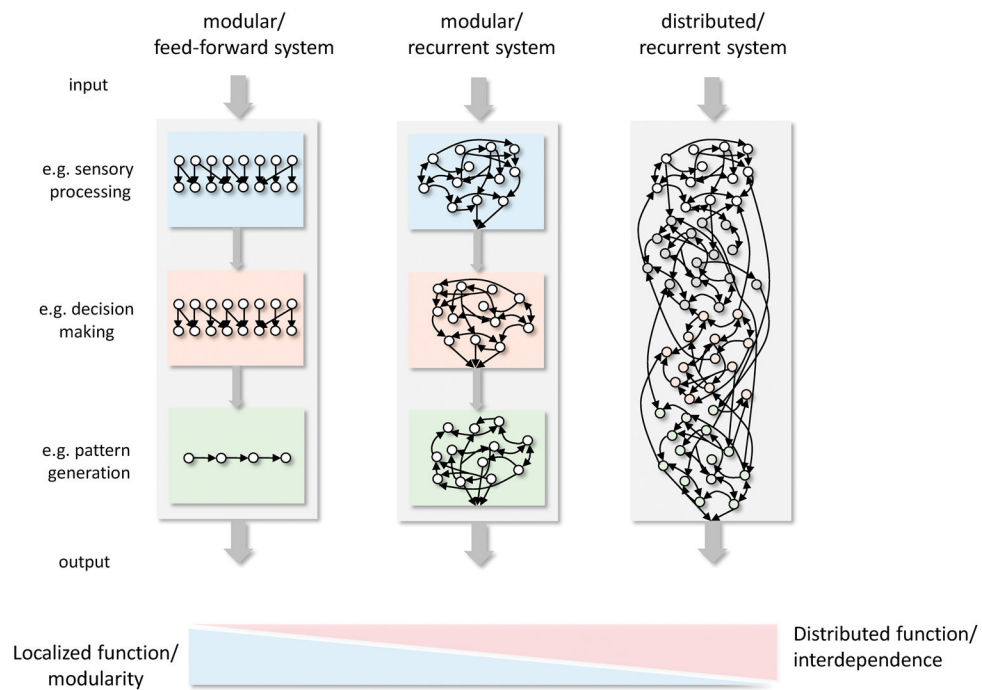
Author Manuscript

Author Manuscript



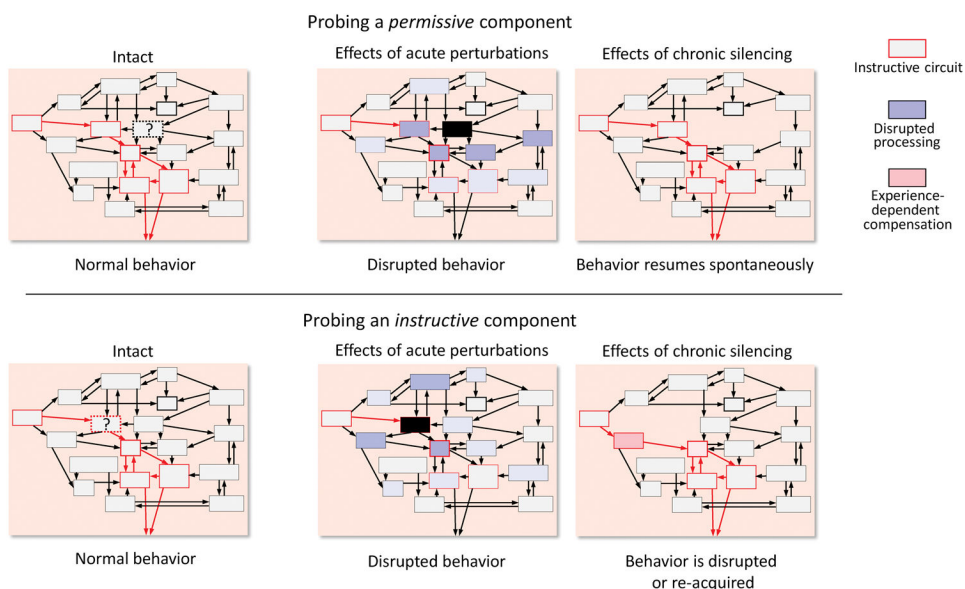
**Figure 1.**

Addressing the neural basis of behavior requires defining the behavioral processes to be understood in terms of the underlying algorithms and computations (left). The implementations can then be probed by recording and manipulating activity in targeted neural circuits during behavior. Typical intermediate goals are: first, to identify the elements in the brain ‘essential’ for a given process (red squares – ‘the causal network’), second, to describe the functions of the different parts, and third, to elucidate the logic by which these functions are implemented. The ultimate goal is to account for the behavioral ‘process’ in terms of how the various ‘parts’ (neurons and brain areas) interact, and to extract general rules that govern these interactions. This review discusses how different types of circuit manipulations can help us achieve this grander goal, and some of their pitfalls.



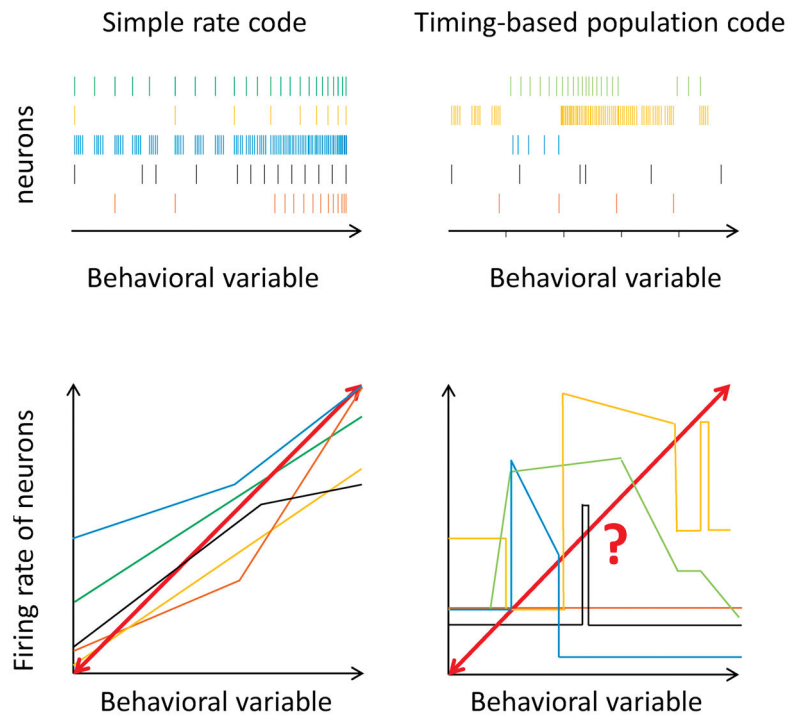
**Figure 2.**

The conceptual and practical utility of mechanistic descriptions depend on the complexity, dynamics, and topology of the underlying network. Certain systems (e.g. simple reflex arcs), can be represented as sequentially connected modules with distinct localized functions (left). In more distributed and complex information processing systems with recurrent connectivity and feedback loops, localization of function and causality are not as well defined, making mechanistic models less intuitive and explanatory. Certain approximations aim to localize functions to specific recurrent circuits (middle), while others treat the system as a complex interconnected dynamical system (right).

**Figure 3.**

Acute and chronic manipulations of neural activity can lead to different behavioral effects and conclusions regarding the role of the targeted circuit elements. When probing a permissive part of the system using acute perturbations (top row, see Box 1), sudden changes to the dynamics of downstream brain areas (blue) could plausibly disrupt their function, leading to behavioral deficits that do not reflect the computations of the targeted circuit element (top middle). Importantly, the outcome would be similar to targeting an essential, or instructive, part of the system (bottom middle). In contrast, after chronic removal of a permissive part of the targeted circuit, the system could spontaneously (i.e. without adding new information through practice) settle into a new dynamical equilibrium and resume its normal function (top right). This same outcome can be had if targeting an essential part of the system as long as there are experience-dependent mechanisms that allow other elements to compensate for lost function (bottom right, see text). Thus behavioral effects of acute manipulations could overestimate the role of a circuit element, while chronic perturbation could under-estimate it.





**Figure 4.** The activity of neuronal populations and the states of a behavioral variable can be related in different ways. Homogeneous bidirectional activity manipulations in neurons using similar rate codes (left) is more likely to mimic behaviorally relevant scenarios than the same manipulations in circuits with more complex population codes (right). Perturbing a system along a behaviorally relevant axis (red lines) allows for more direct tests of causality.