



NEUROPSYCHOPHARMACOLOGY REVIEWS OPEN

Modeling brain, symptom, and behavior in the winds of change

David M. Lydon-Staley^{1,2}, Eli J. Cornblath^{1,3}, Ann Sizemore Blevins¹ and Danielle S. Bassett^{1,4,5,6,7,8}

Neuropsychopharmacology addresses pressing questions in the study of three intertwined complex systems: the brain, human behavior, and symptoms of illness. The field seeks to understand the perturbations that impinge upon those systems, either driving greater health or illness. In the pursuit of this aim, investigators often perform analyses that make certain assumptions about the nature of the systems that are being perturbed. Those assumptions can be encoded in powerful computational models that serve to bridge the wide gulf between a descriptive analysis and a formal theory of a system's response. Here we review a set of three such models along a continuum of complexity, moving from a local treatment to a network treatment: one commonly applied form of the general linear model, impulse response models, and network control models. For each, we describe the model's basic form, review its use in the field, and provide a frank assessment of its relative strengths and weaknesses. The discussion naturally motivates future efforts to interlink data analysis, computational modeling, and formal theory. Our goal is to inspire practitioners to consider the assumptions implicit in their analytical approach, align those assumptions to the complexity of the systems under study, and take advantage of exciting recent advances in modeling the relations between perturbations and system function.

Neuropsychopharmacology (2021) 46:20–32; <https://doi.org/10.1038/s41386-020-00805-6>

INTRODUCTION

In the Dancing Forest on the Curonian Spit in Kaliningrad Oblast, Russia, dozens of pine trees loop around in rings, spirals, hearts, and squiggles. Some inquisitive minds have suggested the work of shifting sands or of *Rhyacionia buoliana* caterpillars, whereas a naive observer might instead turn to the whims of a capricious wind to explain the dances seemingly frozen in time. And precisely how might wind perturb a tree? The answer depends upon the nature of the form “tree”. If a tree were simply a trunk, resting on the ground, then a gust of wind could quite easily raze the tree. If, instead, a tree were a trunk with roots grasping the dirt, then a gust of wind might require additional energy to raze the tree, disturbing the dirt in the process. Finally, if a tree were not standing alone but in a forest of other trees characterized by an intertwined root system, then a gust of wind might require even more energy to raze the tree, disturbing both the dirt and the surrounding trees.

The thought experiment of tree and wind may initially seem esoteric. But upon closer inspection, open questions about the Dancing Forest display marked similarities to questions of fundamental import to neuropsychopharmacology. The tree we care about is a human brain, or a human symptom of psychopathology, or a human behavior. A wind may perturb the tree, either to raze it or twist it, into something beautiful or something broken. In common parlance, that wind may be a change in environment or context, a life event, a healthy or altered neurodevelopmental process, a pharmacological treatment, or a

stimulation regimen. Our capacity to predict the beneficial or detrimental impacts of perturbations on brain, symptom, or behavior, depends in part upon the depth of our understanding of the nature of the form “brain”, or “symptom”, or “behavior”.

And what is the nature of the form “brain”? or “symptom”? or “behavior”? In some studies, we throw up our hands, admitting that we still cannot answer. We instead attempt simply to measure that nature. In other studies, we take a pragmatic tack by assuming a particular nature, thereby allowing us to subsequently make particular inferences. As the tree is assumed to be solely a trunk, a brain region is assumed to be an isolated volume with no connection to other brain regions; a symptom is assumed to be a self-reported experience with no dependence upon other experiences; and a behavior is assumed to be a response to a task with no relation to other responses. Such assumptions implicit in our experimental setup and analysis directly influence the type of inferences that we can draw from our results: the strength of the wind that we predict to be necessary to dance the tree.

Here we consider three common assumptions implicit in our analysis of empirical data in the spheres of brain, symptom, and behavior. We clarify how those assumptions correspond to existing formal models, each of increasing complexity: a commonly applied form of the general linear model (GLM) (tree trunk), impulse response models (tree trunk with roots), and network control models (forest). In each section, we first describe the model with brevity and simplicity. Next, we review how the

¹Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA; ²Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104, USA; ³Neuroscience Graduate Group, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁴Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁵Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁶Department of Electrical & Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁷Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA and ⁸The Santa Fe Institute, Santa Fe, NM 87501, USA
Correspondence: Danielle S. Bassett (dsb@seas.upenn.edu)

Received: 30 April 2020 Revised: 19 July 2020 Accepted: 22 July 2020
Published online: 28 August 2020

model has been used in recent literature to understand neural, symptom, and behavior systems. Finally, we address the strengths and weakness of each model. Our presentation motivates a discussion of future directions in both basic and clinical science, and a thoughtful examination of the relative merit of different points along the continuum from machine learning to computational modeling to theory. The nature of our contribution is partly organizational; we offer the reader an organized account of existing and emerging computational models generically relevant to open challenges in neuropsychopharmacology. Moreover, we provide readers with conceptual tools to reason about their underlying assumptions.

THE SYSTEMS RELEVANT FOR NEUROPSYCHOPHARMACOLOGY

As a field, our scientific investigations often seek to understand a complex system in light of perturbations to the system's structure or function. Such perturbations may be enacted by the environment, designed by laboratory personnel, or offered by trained clinicians. The systems upon which these perturbations impinge are myriad, but our remarks will largely center around the systems of brain, symptom, and behavior (Fig. 1). In what follows, we will use the term brain to refer to the cerebrum and all of its components; the term symptom to refer to a self-reported experience that may be especially intense in states of mental illness; and the term behavior to refer to a measurable (and measured) action that need not be altered in psychiatric or neurological disorders. Below we describe the components of each system and their interconnected nature, as well as the relevance of interconnectivity for perturbations. This discussion will allow us to move to a description of associated modeling efforts in the following sections.

THE NEURAL SYSTEM

Parts and relations

Neural systems are composed of computational and modulatory units that serve to receive, process, alter, and transmit information. Such units exist across multiple spatial scales, from

the microscale of individual cells to the mesoscale of cortical columns, and thereafter to the macroscale of nuclei and cytoarchitectonically distinct areas [1, 2]. Independent of the spatial scale, units are interconnected with one another in a heterogeneous pattern of physical links that support communication, coherence, and other sorts of functional interactions [3–5]. Physical links at the small scale comprise synapses and gap junctions [6], and at the large scale comprise large axonal bundles or white matter tracts [7–9]. The complete neural system thus comprises both parts and pathways, or conduits for relations.

Network representation

Early studies of macaque, cat, and *Caenorhabditis elegans* connectomes suggested the potential utility of representing neural systems as formal mathematical objects known as graphs or networks [10]. In its simplest instantiation, a network is composed of nodes which represent a system's units, and those nodes are connected in pairs by so-called edges which represent a system's inter-unit links, relations, or interactions [2]. From the pattern of edges between nodes, the structure, function, and dynamics of systems can in part be inferred, offering putative explanations for a system's manifest complexity [11]. The network representation has since proven useful in characterizing the interconnected architecture of neural systems across many organisms including humans [12], and the alteration of that architecture in conditions of mental illness and injury [13–15].

Network perturbations

The notion that neural systems are fundamentally network systems has changed the way we think about clinical interventions to support mental health. The most common mechanism of action of pharmacologic interventions for neuropsychiatric illness is the manipulation of distributed neurotransmitter systems. Neurotransmitter systems fall into two major categories: fast ionotropic neurotransmission and slow metabotropic neurotransmission [16]. The latter includes the frequently targeted serotonin [17–19], dopamine [20], and norepinephrine [17, 18] systems. Metabotropic neurotransmitters modulate ionotropic neurotransmission [21, 22], and as a result, the effects of perturbing neurotransmission unfold over multiple time scales. Varying metabotropic neurotransmitter input, also called neuromodulation, can drive groups of neurons to exhibit different patterns of coordinated firing [21], suggesting that neurotransmitters fundamentally influence network interactions rather than eliciting simple increases or decreases in firing.

Neuromodulatory medications cross the blood–brain barrier to reach every brain region, with effects depending on binding profiles of drugs as well as regional expression of neurotransmitter receptors [23]. The mechanism of neurotransmitter action also differs across the brain and depends on external demands [22, 24]. This spatial heterogeneity and state-dependence prevents us from understanding how pharmacotherapy perturbs individual regions in isolation, because a region of interest may be affected by changing inputs from its interacting partners. In our analogy, when there are winds coming from several directions, it is similarly hard to tell which gust had the biggest impact on the tree. Network representations resolve this complexity by allowing the brain to be studied as a whole, rather than as individual regions or even pairwise interactions. Indeed, manipulations of neurotransmitters shift interregional interactions in a way that can be parsimoniously characterized as affecting global network flexibility, segregation, and integration [25, 26]. Metabotropic neurotransmitter systems may contribute to separable components of large-scale fluctuations in cortical activity [27], which may allow the design of future therapeutics to target specific subsets of network interactions. Dynamical network models of neural systems provide a promising computational toolkit for testing

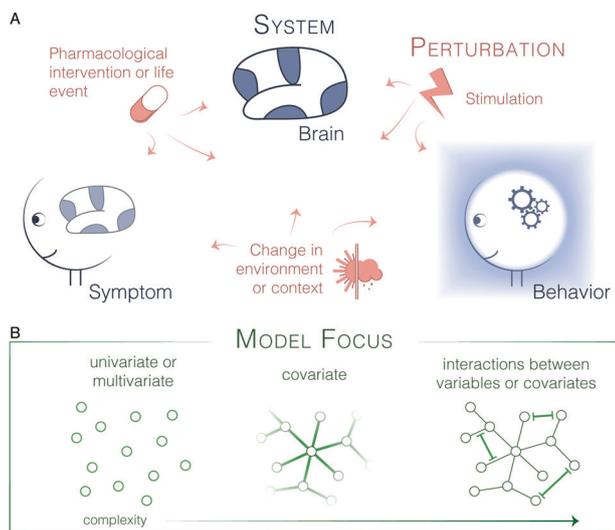


Fig. 1 Overview of systems, perturbations, and models. **a** Here we depict the three systems (blue) of brain, symptom, and behavior. Perturbations (rose) including stimulation, pharmacological interventions, life events, and changes in environment can affect each. **b** We can attempt to understand these systems and their responses to perturbations using computational models (green) of varying complexity.

hypotheses about the associations between neurotransmitter systems, brain activity, and behavior.

THE SYMPTOM SYSTEM

Parts and relations

Symptoms of mental disorders tend to co-occur. For example, individuals receiving treatment for psychiatric disorders experiencing sad mood are likely to also experience anhedonia [28]. Indeed, the formation of symptom clusters is reflected in diagnostic practices; a patient must present with five (or more) symptoms out of a list of nine to be diagnosed with major depressive disorder according to the *Diagnosics and Statistical Manual* [29]. Traditionally, this symptom co-occurrence is explained by assuming that an underlying, latent disease gives rise to groups of symptoms.

Network representation

Network perspectives of mental disorders provide an alternative account to explain symptom co-occurrence [30, 31]. A network approach to psychopathology suggests that symptoms are not simply interchangeable indicators of a latent disorder. Instead, individual symptoms are interesting in their own right and influence one another over time, forming causally connected networks [32, 33]. In line with this perspective, individual symptoms of disorders can vary in their time of onset, their severity, and even in their response to treatments [34–37]. Further, there exists a substantial body of literature that makes use of intensive repeated measures of symptom intensity from moment-to-moment and day-to-day to quantify time-lagged associations between symptoms; these studies provide strong evidence for symptom interplay across time [38–40].

Network perturbations

The network perspective provides a framework within which to capture the importance of individual symptoms as well as the intuitive notion that symptoms act to causally influence one another across time. By focusing on individual symptoms and their interplay, it also puts forward a new perspective for understanding the etiology of mental disorders. A key role is given to a phenomenon termed hysteresis [30, 41] whereby once a system shifts to an alternative state (e.g., a person becomes depressed), it tends to remain in the new state until the external input that led to the change is changed back to a much lower level than was needed to trigger the change in state. The maintenance of a state of psychopathology, consisting of high levels of symptom activity is theorized to occur only in strongly connected networks, consisting of many time-lagged associations between network components. When a symptom in such a strongly connected network is perturbed (e.g., the experience of a stressful life event), symptom activity spreads through causal associations in the network to other symptoms, resulting in a self-sustaining symptom network.

Beyond symptoms

One of the early principles of the network theory of mental disorders was that the components or symptoms in a psychopathology network correspond to problems that have been codified as symptoms in the past century and that are detailed in contemporary diagnostic manuals [30]. This focus on symptoms encoded in diagnostic manuals as the main components of a symptom network preserves the rich information available in diagnostic manuals, emerging from years of clinical observation [42]. Working from this principle, other aspects of human functioning not encoded in diagnostic manuals must (i) constitute a symptom (e.g., a brain-based realization of a symptom), (ii) constitute a symptom-symptom connection, or (iii) act as a variable in the external field that causes a symptom.

THE BEHAVIOR SYSTEM

Parts and relations

What are these extra-symptom aspects of human functioning? It may be reasonable to refer to some such aspects as behavior. Informed by the scientific method, one might define a behavior as a measurable (and measured) action: an interaction with the environment. A common example is a reaction time measured in response to a cognitive demand. A second kind of an extra-symptom aspect of human functioning is a person's emotional state. While distinct from behavior, emotional states can similarly be measured in response to a task manipulation. Both states and actions, however, are not independent but instead can interact. Humans may be more likely to move from a state of contempt to a state of fear than from a state of disgust to a state of joy [43]. Similarly, the measured speed of an action is typically inversely correlated with the measured accuracy of an action [44, 45]. Moreover, a given action might impact a subsequent action manifesting as an asymmetric switch cost [46–49]. Thus, both behavior and state are systems composed of units (actions; states) and their relations (either correlative or causal). This distinction we make between behavior and symptoms is practically useful, as we will see throughout this review; yet, in truth the two exist along a clear continuum of human experience.

Network representation

The network perspective so relevant for psychopathology can also be easily extended to behavior and state systems more broadly. Indeed, theorists, especially developmental scientists, have long considered persons as complex systems, with feelings, thoughts, and actions that are interconnected and that change over time [50–52]. In applications beyond mental disorders, networks consisting of emotions, social behaviors, and other psychological states have been created to realize this notion of persons as complex systems. In studies of non-human animals, recent work has focused on behavior alone, algorithmically coding extensive video footage and separating actions into behavioral units (network nodes) and the likelihood of transitioning between such units (edges) [53, 54]. Collectively, such network maps underscore the fact that domains of function are not binary or unitary. Instead, they are composed of a collection of units whose interactions give rise to the emergent modes of being [55].

Network perturbations

Human behavior is commonly impacted by changes in our environment or the particular social context in which we find ourselves [56–58]. For example, in a stressful circumstance the probability of transitioning from dinner to a drink may be high, whereas in a non-stressful circumstance the probability of transitioning from dinner to a run (or a good book) may be high. Understanding how perturbations to the environment or context lead to predictable changes in a complex system of interacting behavioral units could have far reaching consequences for social policy [57], workplace standards [58], and the practices of educational institutions [56]. Moreover, such an understanding could inform the development of environment interventions for patients with mental illness, consistent with recent work using theatre improvisation training as a recovery-oriented intervention for youths at clinical risk for psychosis [59].

SHARED MODELING NEEDS ACROSS THE THREE SYSTEMS

Across the three systems described above—brain, symptom, and behavior—there are many shared features that could be taken into account by models. First, all three systems are composed of parts that can be perturbed by external factors. Second, perturbations to one part of the system can impact neighboring parts. Third, a system's response to a perturbation may depend not only on the part and its neighbors but also on the relations

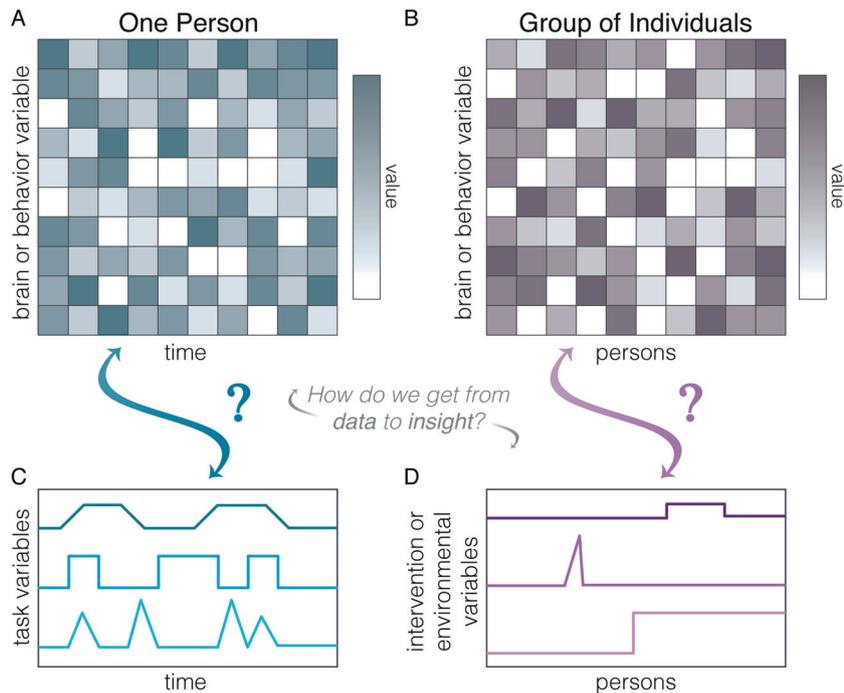


Fig. 2 Univariate or multivariate models to evaluate the effects of perturbations on system parts. **a** The systems that we study are often time-varying over multiple scales. **b** Some systems that we study are composed of canonical trait-like variables that vary across persons, possibly ordered by severity of condition, age, or even an identification number. How do we take these data and from them extract insight? **c** We can understand the impact of task variables on systems like those displayed in panel **a** using several approaches including a typical form of a general linear model. **d** We can also understand the impact of interventions or environmental variables on systems like those displayed in panel **b** using a similar general linear model.

among neighbors and the broader network. When we choose a modeling approach, we often implicitly choose which of these three facts we will acknowledge and investigate.

Like the tree in the wind, our models range from trunk, to trunk with roots, to the entire forest. For example, a commonly applied form of the GLM is used to assess how external perturbations affect the system's part(s). An impulse response model is used to assess how an external perturbation affects a given part of a system and travels into the network through that part's connections. Finally, a network control model is used to assess how an intrinsic or exogenous perturbation can be designed to drive a desired network-wide response. In the sections that follow, we describe each of these modeling approaches, review their use in studying brain, symptom, and behavior, and evaluate their relative strengths and weaknesses. This discussion will allow us to move to a description of future directions in the following section.

MODELING THE TREE AS SIMPLY A TRUNK

We will begin with one of the simplest and yet arguably most pervasive approaches to modeling complex systems that are perturbed by external variables. Specifically, we will consider a commonly applied form of a GLM that is typically used to evaluate the effect of an external variable (e.g., from the environment) on the activity or expression of parts of a system. The underlying assumption of this approach is that the system is composed of parts that can be perturbed (Fig. 2). Perhaps we have measured brain or behavior variables over time (Fig. 2a) and wish to understand how the perturbative drives of a task's structure affect those variables (Fig. 2c). Or perhaps we have measured brain or behavioral variables over persons (Fig. 2b) and wish to understand how the perturbative drives of interventions or environmental events affect those variables (Fig. 2d).

A rather simple form of a GLM allows us to answer these questions in a mathematically rigorous way with appropriate inferential statistics [60]. For a formal mathematical treatment, see for example [61, 62]. In general, these models take the form $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, and in this context, \mathbf{Y} is a vector measuring a perturbed variable over time, \mathbf{X} is a matrix measuring the magnitude of one or more perturbations over time, $\boldsymbol{\beta}$ is a vector of coefficients that estimate the static effect of each perturbation, and $\boldsymbol{\epsilon}$ is a vector of additive errors. In order to identify a unique solution to this equation, there must be more observations (rows) than the number of columns in \mathbf{X} , and no two columns in \mathbf{X} can be linearly dependent, though regularization methods can help to overcome this limitation [62].

Utility in neural, symptom, or behavior systems

To make knowledgeable interventions for the full human, ideally we would understand the relations among all three systems (brain, behavior, and symptoms) in detail. Yet, for practical reasons, many of our empirical studies are constrained to only one part of this multi-system organism. Here (and in the next few sections) we will therefore consider how each system is modeled, and return to their interdigitation towards the end of the review. We do acknowledge, however, that fundamental progress in neuropsychopharmacology will require us to understand how all three systems interact with one another.

Neural system

GLMs are frequently used in the analysis of task functional magnetic resonance imaging data to identify the neural correlates of cognitive functions. This inference is achieved by measuring the BOLD signal change elicited by a particular stimulus, or contrast between stimuli [61]. Classically, one assumes that a neural event elicits a slow increase in BOLD activity, defined by a sum of gamma functions known as the canonical hemodynamic response

function. A GLM is used to estimate the height of this increase independently at every voxel or brain region, thus inferring the extent to which individual regions engage in specific cognitive processes. Although many different types of GLMs can be specified, the most commonly used type does not assume any interactions between the system's parts.

The use of GLMs has helped to localize neural representations of stimulus attributes and identify areas involved in task-general and task-specific cognitive processes. Early work revealed that the fusiform gyrus exhibited apparently selective activation to images of faces. Other studies used GLMs to demonstrate activation of the amygdala and ventromedial prefrontal cortex while processing emotional salience of images [63]. Cognitively demanding tasks were found to elicit deactivation of posterior cingulate and medial prefrontal cortex with activation of dorsolateral prefrontal cortex [64], now known as the default mode system and frontoparietal system, respectively. These two opposing systems were subsequently identified from spontaneous blood oxygen level dependent (BOLD) signal fluctuations observed in resting state fMRI [65], suggesting that the signatures of task activation are etched into the intrinsic network architecture of the brain. Interestingly, while the default mode system is deactivated across many tasks, GLM-based studies have also revealed that default mode activity increases during tasks requiring autobiographical memory and theory of mind [66]. GLMs have provided critical insights into multiple cognitive processes, including but not limited to primary stimulus processing, internally- and externally directed attention, and self-referential thought.

In addition to revealing cognitive processes associated with brain areas, GLMs have served as a foundation for the subsequent development of more complex analytic techniques. For instance, GLMs are so commonly used in fMRI research that publicly available tools exist that allow for meta-analysis of more than 500,000 activation maps from over 14,000 studies [67], facilitating additional discoveries. Early GLM studies also laid the framework for the task fMRI protocols of large sample size studies such as the Human Connectome Project [68] and the Philadelphia Neurodevelopmental Cohort [69], which have subsequently revealed a great deal about predicting individual differences in cognition from brain activation [70–72]. Overall, GLM techniques have both revealed important associations between cognition and brain activation, and further motivated a rich literature on the neural substrates of individual differences in behavior.

Symptom and behavior systems

The GLM has been the workhorse of the psychologist's statistical toolbox when examining how mental illness, and behavior more generally, is impacted by changes in the environment or is targeted during intervention efforts. Examples relevant to neuropsychopharmacology are plentiful and include tests of the effects of behavioral interventions designed to reduce major depressive disorder [73], the effects of medication on nicotine withdrawal during smoking cessation attempts [74], and the effects of stressful life events on cognition [75]. GLM approaches treat other symptoms and behaviors that could plausibly affect the outcome variable of most interest as control variables. This approach is taken to identify, as cleanly as possible, the effect of the intervention or event of interest on the symptom or behavior of interest, independent of other possible sources of influence. Thus, the potentially interconnected nature of the symptoms of a disorder or the sets of behaviors within an organism's repertoire is typically ignored or viewed as a nuisance to be accounted for during statistical modeling rather than an attribute of interest.

Relative strengths and weaknesses

Restricting our attention to neural systems alone, it is worth mentioning a few strengths and limitations of the commonly used form of the GLM. While this model has been highly useful for

relating regional activation to cognition, it is nevertheless a mass univariate test as applied to fMRI data. As a result, the model is agnostic to the role of interactions between brain areas in cognition. This common form of the GLM and other trial-averaged measures may therefore fail to detect spatiotemporal waves of activation by virtue of modeling each region or voxel independently [76]. Additionally, traditional approaches for GLMs involve the assumption that neural activity in all brain areas and all individuals under study exhibit the same hemodynamic response profile, which is unlikely to be the case [77–79]. Issues of hemodynamic response variability aside, the field has continued to debate the appropriate method for assessing statistical significance of brain activation maps for almost two decades since the introduction of the technique [61, 80, 81]. For clinical use, resting state connectivity measures may indeed provide more reliability than task fMRI and the commonly applied form of the GLM, depending on the extent to which a lack of cognitive engagement in certain clinical groups acts as a confound [82].

Broadening out from neural systems alone, several relative strengths and weaknesses of the commonly used form of the GLM are shared across systems. A shared strength is that the approach can flexibly address univariate and multivariate processes; that is, it can go beyond the investigation of a single part of the system to acknowledge the fractionated, particulate, and decomposable nature of complex systems [83]. A shared weakness of the typical model is that it does not formally account for the embeddedness of parts in a complex whole; in other words, perturbations are assumed to affect the system's parts, but not in a way that depends systematically on neighboring parts. This limitation motivates the use of alternative models with a greater level of complexity. In the next section, we explore just such a model type, the impulse response model (Fig. 3), which serves to address the shared limitation of the typical GLM. Alternative models that we do not describe in this review include those that assess the statistical relationships between (i) a measure of connectivity between two regions and (ii) a measure of symptoms or behavior. Particularly good examples of such models include psychophysiological interaction, partial least squares, and canonical correlation analysis [84, 85].

ACKNOWLEDGING THE TREE'S ROOT SYSTEM

Unlike the commonly used GLM just described, network models explicitly encode the relations among a system's parts. Network models of brain, symptoms, and behavior provide insight into which set of system variables contains information that improves the prediction of another set of system variables. The models provide information about the individual associations among network nodes. Yet, in many applications, interest goes beyond this edge-central representation and towards an examination of how the system behaves as a whole. Capturing system-level functioning necessitates a tool capable of modeling the interplay between network units in a high-dimensional system.

Early efforts to capture such system-level notions focused on node centrality and network density. A common measure of node centrality is degree centrality, defined as the number of edges emanating from a node [86]. Here we will frame our discussion within the context of symptom systems, and broaden to other systems later. Node centrality has been identified as a potential index for interrogating networks to find symptoms that would facilitate the spread of behavior through the network, activating self-perpetuating sequences of symptoms across time [87, 88]. The complementary notion of density is defined as the fraction of possible edges that exist in a network. Network density has similarly been identified as a potential index for characterizing the likelihood that a network will support reverberating symptom activation [89, 90]. The process is analogous to the behavior of two sets of domino tiles: one in which the tiles are far apart and a

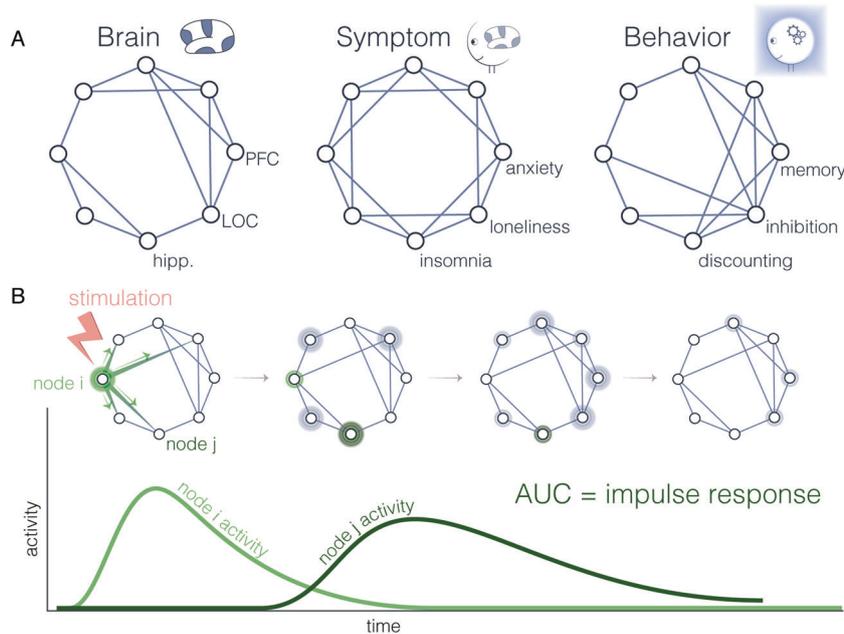


Fig. 3 Impulse response models. **a** (Left) A brain network encoding across-time or across-person relations between regions. (Middle) A symptom network encoding across-time or across-person relations between symptoms. (Right) A behavior network encoding across-time or across-person relations between cognitive, behavioral, or neuropsychological variables. **b** Illustration of the response of a network to an impulse (pink). The network response is shown across the top row; the bottom plot shows the response of the stimulated node (light green) and the response of a second node (dark green). The area under the curve is an impulse response metric.

second in which the tiles are close together. In a densely connected symptom network, when one tile falls it causes other tiles to topple and activity ripples through the system [91]. The use of both centrality and density are limited in their focus on direct connections, containing little information about how symptoms might affect each other indirectly as activity flows through the entire network [92].

Recent efforts have turned to impulse response analysis (see ref. [93, p. 51] for general mathematical treatment and Blaauw et al. [94] for treatment specific to symptom and behavior networks). Using a network model of the interactions between variables across time, impulse response analysis involves first simulating an instantaneous, exogenous impulse (referred to as a shock) to certain variables in the network. Then, one charts how this impulse propagates through the network, along the time-lagged edges. The impulse response function shows the hypothetical change in a variable in response to a simulated increase in one of the other variables over a horizon of several time points. Crucially, the activity in one node observed after a shock to another node in the system contains system-level information. The impulse felt by the non-shocked node is due solely to the flow of activation along the network's edges originally emanating from the shocked node. In this manner, impulse response models (unlike the commonly used form of the GLM) acknowledge the fact that a system's parts influence one another; moreover, that precise influence constrains the response of the system to a perturbation.

Utility in neural, symptom, or behavior systems

Symptom and behavior systems. Despite their marked advantages over simpler models, impulse response models have not been used extensively in the context of behavior and symptom systems. Notable exceptions are studies of depression [95], anxiety [96], and tobacco dependence [97]. Here we review each in turn before moving to a discussion of the use of impulse response models in neural systems.

In a sample of subclinically depressed individuals, Bos et al. [95] examined affect dynamics in daily life, and the extent to which

physical activity and experiences of stress impacted affect in participants with anhedonic symptoms versus those without. By completing affect, behavior, and cognition items via an electronic diary three times a day for 30 consecutive days, participants generated time series that were then subjected to vector autoregression (VAR) in which each variable was regressed on its own lagged values as well as the lagged values of the other variables. This process resulted in, for each participant, a network indicating the time-lagged associations between six variables: high-arousal positive affect, low-arousal positive affect, high-arousal negative affect, low-arousal negative affect, stress, and physical activity. Impulse response analysis was then applied to these VAR models to examine the impact of physical activity and stress on affect over a horizon of ten time points. When an increase in physical activity was simulated, the other variables only marginally changed in response. When an increase in stress was simulated, affect in non-anhedonic individuals showed a stronger increase relative to affect in anhedonic individuals. These findings ran contrary to expectations, highlighting the surprising findings that can emerge when the flow of activity from one behavior to another is allowed to play out in a complex, high-dimensional system.

In the preceding example, the effect of an individual behavior (e.g., physical activity) on an individual outcome (e.g., negative affect) was examined. By using impulse response analysis, the observed association between physical activity and negative affect took into account the broader system consisting of six variables. The effect of physical activity on negative affect consisted of the effect of a simulated increase in physical activity filtering through direct pathways to impact negative affect, but also via indirect pathways emerging from the complex interplay between physical activity and negative affect with other variables in the system (e.g., stress) across time. Other uses of impulse response analysis in the literature have focused on capturing system properties occurring at a higher level of topological analysis [1]. Using experience-sampling data assessing six symptoms of tobacco withdrawal (anxiety, craving, depressed mood, irritability, hunger, and

difficulty concentrating) in daily smokers during a 2-week smoking cessation attempt, Lydon-Staley et al. [97] examined the effects of cessation treatments on the reverberation of symptom activity across time. After constructing person-specific networks indicating the time-lagged associations between symptoms, impulse response analysis was applied to quantify how a simulated increase in one symptom affected other symptoms. Specifically, the time profile of a symptom of interest following perturbation of another symptom was simulated over 100 time steps. The time profile was then examined to identify the recovery time of the symptom, quantified as the number of time steps from perturbation to equilibrium (see also ref. [98]). The average time taken for a symptom to return to equilibrium after perturbation was faster in participants in an intensive combination treatment condition relative to participants in a matched placebo condition. Broadly, the study demonstrates the distinct insights that an impulse response analysis can provide on the question of how interventions can differentially impact the dynamic interplay between symptoms.

The preceding examples used self-reports of behavior, symptoms, and various states as network components and employed impulse response analysis to capture the effect of a change in a variable of interest (e.g., physical activity) on an outcome of interest [95] (e.g., affect) or to capture the system-level notion of hysteresis [97, 98]. In a striking example demonstrating the flexibility of both the network approach and impulse response analysis, Yang et al. [96] used a combination of the two to capture adolescent emotion system dynamics. Second-by-second psychophysiological time series, consisting of respiratory sinus arrhythmia (an indicator of parasympathetic nervous system activity [99]) and skin conductance level (an indicator of sympathetic nervous system activity [100]), and second-by-second self-report of distress measured using video-mediated recall were collected from 130 adolescents. In constructing a network of the time-lagged associations embedded in this tri-node system, notions of emotional concordance and emotion regulation were captured. Emotional concordance is the synchronized combination of psychophysiological, cognitive, and behavioral components of the emotion system [101]. Emotion regulation is the ability to produce a flexible, controlled emotional response to environmental events [102]. An impulse was given to each node separately and, for each simulation, changes in all three nodes were tracked until activity reached equilibrium. The value of node activity at this equilibrium level provided an index of system reactivity. In contrast to prior studies, the approach utilized here strongly underscores the interrelated nature of emotional states.

Neural system

Impulse response models and related techniques have also been applied to the study of neural systems. Early fMRI studies sought to characterize the impulse responses of primary visual cortex to single stimuli [103, 104], without estimating an underlying multicomponent network model. More recent studies have used activity flow, a variant of impulse response models, to estimate task activations from activity spread along the whole-brain resting state functional connectome [105, 106]. Work from this group has focused on average controllability, which quantifies the area under the curve of the impulse response of each brain region, using linear dynamics along structural networks obtained through diffusion-weighted imaging [107, 108]. Average controllability of frontolimbic regions increases with age [109, 110] is associated with impulsivity [110], and is reduced in patients with long-standing bipolar disorder [111]. Studying the effects of local perturbations in nonlinear systems [112, 113] is promising due to their ability to recapitulate oscillations characteristic of brain signals [114]. However, a validated toolbox of regional measures of impulse response does not exist for such systems, though one study suggests that linear impulse response metrics may extend

to nonlinear systems [115]. These findings suggest that impulse response metrics in neural systems capture behaviorally relevant information that may serve as useful biomarkers and theoretical testbeds for treatments in the future.

Relative strengths and weaknesses

Impulse response analysis provides a fitting analytic match to theories emphasizing the spread of activity through a network of connected components. It incorporates both the directionality and the sign of edges during computation, allowing insight into complex features such as positive and negative feedback loops. This sensitivity stands in contrast to approaches that take the absolute value of edges in brain or behavior networks before computing the density of a system to quantify the potential for activity spread through a network [89, 90] or those that decompose strength into in-strength and out-strength when identifying influential nodes [39, 116].

Alongside these strengths, a number of limitations are important to consider. In using impulse response analysis coupled with a VAR model, we make the assumption that the system and the interactions among units in the system are time invariant. It is this assumption that allows us to forecast how the system will react to a perturbation. Time-varying VAR models may be appropriate in cases in which system dynamics change over time [117] and would limit the use of impulse response analysis for predicting system activity following perturbation to periods of time where time-invariant VAR models provide a sufficient description of system dynamics. A related consideration associated with limiting one's focus to periods during which time-invariant processes may be assumed is that densely sampled time series data are often required to fit VAR models. Guidelines for the amount of data required remain to be established, will depend on the size of the system one is attempting to model, and will benefit from the transparent reporting of when time series length emerges as a barrier to VAR modeling (see ref. [98] for an example of such reporting). An additional consideration concerns the system units used to build a model upon which we simulate perturbation. The results of simulated perturbations will differ depending on the units we assume are relevant to a system. Yet, for even the most intensively studied psychological disorders (e.g., major depression) there exists substantial heterogeneity across measurement instruments as to which symptoms are relevant [118]. We note that these considerations are not specific to impulse response analysis but to network approaches more generally.

A consideration more specific to impulse response analysis relates to the time horizon chosen to examine system responses to perturbation. The examples of impulse response analysis in brain, behavior, and symptom networks described above flexibly used several different time horizons. For example, one used a time horizon of 10 time points following perturbation [95]; one used a backward search to identify the time taken for the system to return close to its starting value following perturbation [97]; and another identified the time at which the system settled at a point of equilibrium [96]. This difference across studies highlights the flexibility with which impulse response analysis can be used to capture the system responses of most interest. Yet, it also reveals a limitation of the method: chiefly, the difficulty of incorporating information about particular system end states that may be of theoretical interest. In the next section, we therefore turn to a model that addresses this limitation: namely, the network control model (Fig. 4).

EXPANDING TO THE FOREST

Network control models stipulate how the state of a system changes over time as a function of intrinsic dynamics and time-varying external inputs. Note that the impulse response model is a special case of network control models, which captures how the

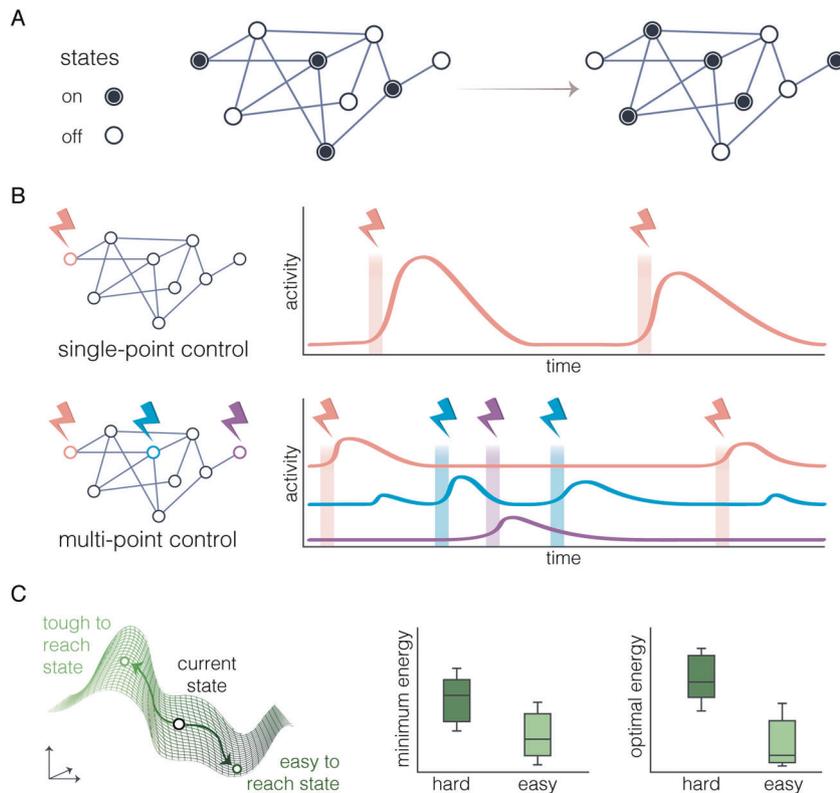


Fig. 4 Network control models. **a** Network control theory asks how to design interventions that move a system from one state (values on nodes) to another state (different values on nodes), by constraining the cause of state change to occur along edges. **b** Interventions can be enacted in a time-dependent manner from a single region (top) or from multiple regions (bottom). **c** (Left) Illustration of an energy landscape that is dictated by the architecture of the network and that in turn encodes the fact that some transitions are easier to drive than others. (Right) Two complementary metrics that quantify the ease of a state transition.

system evolves following a single initial input. Like impulse response models, network control models define a system of N interacting units, each of which has some measured property at time t represented in a vector $\mathbf{x}(t)$. For example, the N units may represent behavioral symptoms, neurons, or mesoscale brain structures. The values of $\mathbf{x}(t)$ might therefore contain measures of symptom intensity, firing rate, or BOLD fMRI signal, respectively. For a formal mathematical treatment, see [119].

Distinct from the special case of impulse response models, a generic network control model may contain time-varying external inputs into the system represented by the vector $\mathbf{u}(t)$. These inputs are filtered into the units and dimensions of \mathbf{x} via some transformation. The temporal progression of the system can then be described as $\dot{\mathbf{x}}(t) = f(\mathbf{x}, \mathbf{u}, t)$. This function posits that the change in state variables $\dot{\mathbf{x}}(t)$ is determined by some function of the current state of the system, intrinsic interactions between the N units, and filtered external inputs. In the case where $\mathbf{x}(t)$ represents symptoms, \mathbf{u} might represent an adverse life event, which, when filtered by some function, has a characteristic impact on the symptoms in $\mathbf{x}(t)$. In the case where $\mathbf{x}(t)$ represents neural activity, \mathbf{u} might represent the binding affinity of a drug, which has a characteristic impact on neural activity when filtered through spatial patterns of receptor expression. Network control models are particularly well-suited to the study of neural, symptom, and behavioral systems because they capture the continuous interplay between intrinsic dynamics and the external environment [120].

Utility in neural, symptom, or behavior systems

Neural system. Numerous studies have applied models of network control to mesoscale [111, 121–124] and microscale

[125–127] neural systems. Due to analytical and computational limitations in the study of nonlinear systems, most studies have used linear, time-invariant network models [128]. The most common form is $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$, where \mathbf{A} contains a set of measured connections between N neural units and \mathbf{B} is a linear input filter. A subdomain of network control theory, known as optimal control, allows one to solve for the external inputs $\mathbf{u}(t)$ needed to drive a linear system \mathbf{A} from a specified initial state to some desired target state [121–123]. The literature on optimal control is related to a sister literature on reinforcement learning and other continuous action policy search methods [129].

Using the control framework, the latent inputs driving brain activity across multiple cognitive tasks can be recovered [130, 131]. Data-driven network control models have been used to successfully decode mood fluctuations [132] and motor patterns underlying speech [133] from implanted electrode grids in patients with intractable epilepsy. A combination of open loop and optimal control techniques can explain how exogenous electrical stimulation drives the brain into activity states favorable for episodic memory recall [123]. Patients with schizophrenia were found to theoretically require stronger inputs to reach whole-brain fMRI activation patterns associated with working memory performance, while the strength of inputs required to maintain those states varied with estimated prefrontal D1 and D2 receptor expression [134]. Patients with chromosome 22q11.2 deletion syndrome, which confers genetic risk for schizophrenia, were found to spend more time in fMRI activation patterns requiring stronger inputs to maintain [135]. These initial studies lay the framework for merging neurotransmitter pharmacology and brain stimulation approaches with network control models of neural activity.

Symptom and behavior systems

Mental health professionals are increasingly interested in interrogating personalized behavior networks (broadly encompassing emotions, cognitive processes, symptoms, and actions) in order to inform treatment [136]. A proposed use of the estimated networks is the identification of nodes that drive the behavior of other nodes and that, thus, can be targeted early in treatment to arrive at a desired system state [137, 138]. Methods of identifying nodes with particularly strong downstream effects on other system nodes include patient–provider discussions of estimated networks to identify potential causal chains [138] and clinician examination of edges in estimated networks to identify edges with particularly large coefficients [137]. Nodes with relatively many and relatively strong out-going edges (high out-strength in the parlance of graph theory) are also thought to be important nodes for influencing other network nodes [139].

Network control models have recently been extended to psychological networks [140] to provide a statistical framework capable of guiding personalized interventions. Data-driven control models that have been applied to neural systems to explain how exogenous stimulation drives the brain into desired activity states [123] can be readily translated to symptom and behavior networks to identify optimal treatment targets and to simulate the theoretical efficacy of potential interventions. The translation of network control models from neural systems to symptom and behavioral networks is an important step with promising potential to overcome the limitations of relying on visual inspection and estimation of centrality indices of symptom and behavior networks to guide treatment.

Relative strengths and weaknesses

The mathematical sophistication of the network control models provides an elegant simplicity and validity in tackling the true complexity of neural systems and behaviors. Of course, we must always acknowledge that although more complex models may more closely align with the complexity of the true system, they can also hamper intuitions and be difficult to interpret as well as fit to data. Indeed, the practical clinical utility of network control models has yet to be realized due to this type of limitation. For example, the functions relating the state of the system and the inputs are often unknown in practice. In the case of neurotransmitter systems, this ignorance stems from an inability to link molecular and neuronal mechanisms with mesoscale measurements of brain activity. In the case of behavioral symptoms, it is unknown how life events reliably impact sets of symptoms over time. Studies of network control [123, 124, 134] have gleaned useful insights by solving for unfiltered, theoretical inputs that act directly upon the N units of $x(t)$. However, practical utility depends upon the link between real interventions (i.e., pharmacologic agents or brain stimulation) and brain state (i.e., $x(t)$). Data-driven approaches may prove useful for approximating these link functions as well as the underlying network interactions [130, 141, 142], but they will require careful validation with large amounts of data from single individuals before they can be used in the clinic. Gleaning interpretable and intuitive clinical decision support from complex models may require clinicians to have independent expertise, likely limiting their utility to highly specialized care settings.

From a practical point of view, the utility of network control models depends upon the certainty with which the network itself can be estimated, and the verity of the form of dynamics used in the model. When using network control models in the context of human structural brain networks, care should be taken to consider the impact of distinct diffusion imaging sequences that differ in their spatial resolution, scan duration, and number of diffusion directions sampled. Similarly, when using network control models in the context of symptom and behavior networks, formal tools to quantify uncertainty on each of the network edges linking

symptoms (or linking behaviors) will be important. We recognize that thorough methodological studies have yet to be done to define clear benchmarks for the amount and type of data needed to build accurate models to ensure a given level of confidence in statistical inference. We look forward to future work addressing these important issues.

FUTURE DIRECTIONS

The three modeling approaches that we have discussed here are motivated by the overarching goal of understanding how a network system (brain, symptom, behavior) responds to a beneficial or maleficent perturbation (Fig. 1). How does the tree bend in the wind? Each model type seeks that understanding by implementing a different level of complexity in the analytical assumptions and model formulation: the most commonly applied form of the GLM (tree as trunk), impulse response models (tree as trunk with roots), and network control models (forest with interdigitated root system). Each new model advances the state of the field by addressing the limitations of the previous model. While the process of model development and selection continues apace, many opportunities exist to take full advantage of the approaches available today.

In the realm of basic science, there is one troublesome question left open by our entire discussion. Precisely how are the three systems interlinked with one another? How exactly does one system impinge upon or support another? And even more: How can such interlinkage be modeled? The pines of Kaliningrad Oblast thrive in a broader ecosystem alongside diverse other flora. How might the approaches that we review be extended to deal with the presence of more than one system, with its own notion of a part (or network node) and its own notion of a relation (or network edge)? Might we need an entirely new approach all together? Perhaps it could prove useful to consider canonical correlation analysis as a means of deriving dependencies between two systems [143, 144]. Such dependencies could be used to develop multilayer network models of two or more systems [145], which in turn could be studied with the principles of network control [146]. Are such possibilities reasonable next steps? Or are more foundational studies needed first?

In considering clinical translation, perhaps the most critical challenge we face is understanding and parsing heterogeneity [147]. Differences in participant, treatment, and outcome characteristics challenge the development of experiments, the collection of necessary data, and the subsequent inferences that can be drawn. Implicit in understanding heterogeneity is understanding the time scales over which participant characteristics (brain, symptom, and behavior) are trait-like versus state-like, as this distinction can clarify whether differences are true or only apparent. Traits are often hailed as the gold standard in clinical medicine due to their reproducibility; that is, the fact that they remain unchanged over iterative measurement. Yet, complex systems typically display rich temporal dynamics in which the reproducible feature is a dynamical rule (fit by first, second, and sometimes even third derivatives) rather than by any single snapshot of variable values. By requiring reproducibility of a static measurement, rather than conscientiously seeking reproducible dynamical trajectories through a rich state space, are we hamstrung in truly understanding the intricacies of mental illness? What are the appropriate time scales over which to be measuring (and by extension modeling) these systems? And are those time scales the same as or different from the appropriate time scales over which to be measuring the effects of perturbations? How would the answers to these questions change the way we model?

Beyond basic and clinical science, the methodologists might ask whether statistical models (such as those described here) are truly the workhorse of the future. Will statistical models eventually be more formally complemented by machine learning approaches

and if so, how? It is true that some machine learning models have been criticized for lack of interpretability, and others for bias and lack of fairness in their respective algorithms [148, 149]. Yet, it is important to acknowledge that some computational models can also lack interpretability. In determining the appropriate balance between statistical models and machine learning models, it is important to clearly state the sort of understanding the investigator wishes to obtain. A black-box algorithm may be superior if the question is predicting treatment outcome based on a complex biological measure, but will do less well if the goal is to understand the computational role a specific brain region plays in a given behavior. Computational models also have the benefit of being written in the same language as generalizable formal theories, which in turn are structured in a way that humans can naturally reason about and with. A complementarity between statistical models and machine learning could take on several forms. Theory-agnostic algorithms may prove useful for making a first reasonable estimate of appropriate model parameters, or perhaps for determining which parts of a many-part system must be included in the model to canvas the space of possible dynamics appropriately. Future work is likely to better formalize the types of interdigitation between algorithm and model that may best advance the science toward our collective goals.

A key consideration in linking theory, modeling, and experiment is the utilization of an appropriate framework for causal inference [150]. Our discussion has focused on studying and discovering the effect of an intervention in a complex system. This effort can quite naturally lend itself to the identification of causes and mechanisms. The three methods we cover differ in their explicit account of causality. For example, the particular form of GLMs that we discuss seeks to measure the statistical variance in system function that is explained by a perturbation. This approach can provide evidence in support of a causal theory [151], but does not itself comprise a causal framework. In contrast, impulse response models and network control models apply an artificial perturbation *in silico* to predict the effects of a natural perturbation *in vivo*. Two causal frameworks that appear particularly relevant to these models are that of mechanisms and that of pathways [152, 153]. Future work bridging empirical science, theoretical science, and philosophy of science could seek to clarify particularly fruitful points of interdigitation between perturbative models and causal frameworks. Moreover, such efforts could delineate best practices in causal inference from observational studies with particularly abundant data versus from clinical studies with less data and carefully designed interventions.

Finally, it is worth posing the question of whether computational models are the end goal, or whether they are a humble stepping stone to a larger goal. In considering this question, we are faced with the fundamental differences and relations between modeling and theory. What we truly seek is understanding, and computational models (along with their associated analytical approaches) are useful tools with which to gain understanding. But a model is not the understanding itself. Models can instantiate theories, and understanding arises when theories are tested, and either proven or disproven. An important future direction, then, is to define the theories behind the models, and use the models in a more directed way to prove or disprove specific theories.

CONCLUSION

Together, our goal is to understand how, when, and where to intervene to divert mental illness and drive mental health. Frankly, we are reaching for the stars. While the stars may still be beyond our reach, we may attain some altitude by taking a path through the Dancing Forest, and acknowledging how the wind bends the trees. Recent advances in modeling the relations between perturbations and system function lay the groundwork for us to better understand how brain, symptom, and behavior

systems respond to exogenous changes both in health and disease.

CITATION DIVERSITY STATEMENT

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minorities are under-cited relative to the number of such papers in the field [154–159]. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, and other factors. We obtained predicted gender of the first and last author of each reference by using databases that store the probability of a name being carried by a woman [154, 160]. By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 12% woman(first)/woman(last), 14.5% man/woman, 19.7% woman/man, 53.8% man/man, and 0% unknown categorization. This method is limited in that (a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and (b) it cannot account for intersex, non-binary, or transgender people. We look forward to future work that could help us to better understand how to support equitable practices in science.

FUNDING AND DISCLOSURE

D.M.L.-S. was supported by a K01 from the National Institute on Drug Abuse (K01DA047417). E.J.C. was supported by an F30 from the National Institute of Mental Health (F30 MH118871-01). A.S.B. and D.S.B. acknowledge support from the John D. and Catherine T. MacArthur Foundation, the ISI Foundation, the Paul G. Allen Family Foundation, the Army Research Laboratory (W911NF-10-2-0022), the Army Research Office (Falk W911NF-18-1-0244, Bassett-W911NF-14-1-0679, Grafton-W911NF-16-1-0474), the National Science Foundation (BCS1631550, PHY-1554488, NCS-FO-1926829), the National Institute of Mental Health (2-R01-DC-009209-11, R01-MH112847, R01-MH107235, R21-MH-106799, R01-MH-116920), and the National Institute of Child Health and Human Development (1R01HD086888-01). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies. The authors declare no competing interests.

ACKNOWLEDGEMENTS

We thank Dr. Xiaosong He, Harang Ju, Sophie Loman, Sophia David, and Dr. Linden Parkes for helpful comments on an earlier version of this manuscript.

AUTHOR CONTRIBUTIONS

D.S.B. conceived the piece; D.S.B., D.M.L.-S., E.J.C., and A.S.B. designed the current instantiation of the conception and revised the manuscript and approved the final version; D.S.B., D.M.L.-S., and E.J.C. drafted the manuscript; A.S.B. constructed the figures.

ADDITIONAL INFORMATION

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Betzel RF, Bassett DS. Multi-scale brain networks. *Neuroimage*. 2017;160:73–83.
2. Bassett DS, Zurn P, Gold JI. On the nature and use of models in network neuroscience. *Nat Rev Neurosci*. 2018;19:566–78.
3. Avena-Koenigsberger A, Misisic B, Sporns O. Communication dynamics in complex brain networks. *Nat Rev Neurosci*. 2017;19:17–33.
4. Palmigiano A, Geisel T, Wolf F, Battaglia D. Flexible information routing by transient synchrony. *Nat Neurosci*. 2017;20:1014–22.

5. Srivastava P, Nozari E, Kim JZ, Ju H, Zhou D, Becker C, et al. Models of communication and control for brain networks: distinctions, convergence, and future outlook. *Netw Neurosci*. 2020. Advance publication. https://doi.org/10.1162/netn_a_00158.
6. Schafer WR. The worm connectome: back to the future. *Trends Neurosci*. 2018;41:763–5.
7. Butt AM, Fern RF, Matute C. Neurotransmitter signaling in white matter. *Glia*. 2014;62:1762–79.
8. Johansen-Berg H. Human connectomics—what will the future demand? *Neuroimage*. 2013;80:541–4.
9. Le Bihan D, Johansen-Berg H. Diffusion MRI at 25: exploring brain tissue structure and function. *Neuroimage*. 2012;61:324–41.
10. Bassett DS, Bullmore E. Small-world brain networks. *Neuroscientist*. 2006;12:512–23.
11. Mitchell M. *Complexity: a guided tour*. Oxford University Press; 2011.
12. van den Heuvel MP, Bullmore ET, Sporns O. Comparative connectomics. *Trends Cogn Sci*. 2016;20:345–61.
13. Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. *Nat Rev Neurosci*. 2015;16:159–72.
14. Braun U, Schaefer A, Betzel RF, Tost H, Meyer-Lindenberg A, Bassett DS. From maps to multi-dimensional network mechanisms of mental disorders. *Neuron*. 2018;97:14–31.
15. Douw L, van Dellen E, Gouw AA, Griffa A, de Haan W, van den Heuvel M, et al. Stam. The road ahead in clinical network neuroscience. *Netw Neurosci*. 2019;3:969–93.
16. Purves D, Fitzpatrick D, Katz LC, Lamantia AS, McNamara JO, Williams SM, et al. *Neuroscience*. Sinauer Associates; 2000.
17. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163:1905–17.
18. Koen N, Stein DJ. Pharmacotherapy of anxiety disorders: a critical review. *Dialogues Clin Neurosci*. 2011;13:423.
19. Stein DJ, Ipser JC, Seedat S, Sager C, Amos T. Pharmacotherapy for post traumatic stress disorder (PTSD). *Cochrane Database Syst Rev*. 2006;1. <https://doi.org/10.1002/14651858.CD002795.pub2>.
20. Buchanan RW, Kreyenbuhl J, Kelly DL, Noel JM, Boggs DL, Fischer BA, et al. The 2009 schizophrenia port psychopharmacological treatment recommendations and summary statements. *Schizophrenia Bull*. 2010;36:71–93.
21. Nusbaum MP, Beenhakker MP. A small-systems approach to motor pattern generation. *Nature*. 2002;417:343–50.
22. Avery MC, Krichmar JL. Neuromodulatory systems and their interactions: a review of models, theories, and experiments. *Front Neural Circuits*. 2017;11:108.
23. Turk E, Scholtens LH, van den Heuvel MP. Cortical chemoarchitecture shapes macroscale effective functional connectivity patterns in macaque cerebral cortex. *Hum Brain Mapp*. 2016;37:1856–65.
24. Shine JM, van den Brink RL, Hernaes D, Nieuwenhuis S, Poldrack. RA. Catecholaminergic manipulation alters dynamic network topology across cognitive states. *Netw Neurosci*. 2018;2:381–96.
25. Braun U, Schaefer A, Walter H, Erk S, Romanczuk-Seiferth N, Haddad L, et al. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc Natl Acad Sci USA*. 2015;112:11678–83.
26. Shine JM, Bissett PG, Bell PT, Koyejo O, Balsters JH, Gorgolewski KJ, et al. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron*. 2016;92:544–54.
27. Shine JM, Breakspear M, Bell PT, Martens KAE, Shine R, Koyejo O, et al. Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat Neurosci*. 2019;22:289–96.
28. Beard C, Millner AJ, Forgeard MJC, Fried EI, Hsu KJ, Treadway MT, et al. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychol Med*. 2016;46:3359–69.
29. Castillo RJ, Carlat DJ, Millon T, Millon CM, Meagher S, Grossman S, et al. *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association Press; 2007.
30. Borsboom D. A network theory of mental disorders. *World Psychiatry*. 2017;16:5–13.
31. Nuijten MB, Deserno MK, Cramer A, Borsboom D, et al. Mental disorders as complex networks: an introduction and overview of a network approach to psychopathology. *Clin Neuropsychiatry*. 2016;13:5.
32. Bringmann LF, Vissers N, Wichers M, Geschwind N, Kuppens P, Peeters F, et al. A network approach to psychopathology: new insights into clinical longitudinal data. *PLoS ONE*. 2013;8:e60188.
33. Schmittmann VD, Cramer AOJ, Waldorp LJ, Epskamp S, Kievit RA, Borsboom D. Deconstructing the construct: a network perspective on psychological phenomena. *N. Ideas Psychol*. 2013;31:43–53.
34. Bujarski S, Roche DJO, Sheets ES, Krull JL, Guzman I, Ray. LA. Modeling naturalistic craving, withdrawal, and affect during early nicotine abstinence: a pilot ecological momentary assessment study. *Exp Clin Psychopharmacol*. 2015;23:81.
35. Piper ME, Schlam TR, Cook JW, Sheffer MA, Smith SS, Loh W-Y, et al. Tobacco withdrawal components and their relations with cessation success. *Psychopharmacology*. 2011;216:569–78.
36. Hirschfeld RMA, Mallinckrodt C, Lee TC, Detke MJ. Time course of depression-symptom improvement during treatment with duloxetine. *Depress Anxiety*. 2005;21:170–7.
37. Conradi HJ, Ormel J, De Jonge P. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. *Psychol Med*. 2011;41:1165–74.
38. Lydon-Staley DM, Cleveland HH, Huhn AS, Cleveland MJ, Harris J, Stankoski D, et al. Daily sleep quality affects drug craving, partially through indirect associations with positive affect, in patients in treatment for nonmedical use of prescription drugs. *Addict Behav*. 2017;65:275–82.
39. Fisher AJ, Reeves JW, Lawyer G, Medaglia JD, Rubel JA. Exploring the idiographic dynamics of mood and anxiety via network analysis. *J Abnorm Psychol*. 2017;126:1044.
40. Groen RN, Snippe E, Bringmann LF, Simons CJP, Hartmann JA, Bos EH, et al. Capturing the risk of persisting depressive symptoms: a dynamic network investigation of patients' daily symptom experiences. *Psychiatry Res*. 2019;271:640–8.
41. Cramer AOJ, van Borkulo CD, Giltay EJ, van der Maas HLJ, Kendler KS, Scheffer M, et al. Major depression as a complex dynamic system. *PLoS ONE*. 2016;11:e0167490.
42. Suris A, Holliday R, North CS. The evolution of the classification of psychiatric disorders. *Behav Sci*. 2016;6:5.
43. Thornton MA, Tamir. DI. Mental models accurately predict emotion transitions. *Proc Natl Acad Sci USA*. 2017;114:5982–7.
44. Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S. The neural basis of the speed-accuracy tradeoff. *Trends Neurosci*. 2010;33:10–6.
45. Heitz RP. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front Neurosci*. 2014;8:150.
46. Schneider DW, Anderson JR. Asymmetric switch costs as sequential difficulty effects. *Q J Exp Psychol*. 2010;63:1873–94.
47. Arbuthnott KD. Asymmetric switch cost and backward inhibition: carryover activation and inhibition in switching between tasks of unequal difficulty. *Can J Exp Psychol*. 2008;62:91–100.
48. Samuel S, Roehr-Brackin K, Jelbert S, Clayton NS. Flexible egocentricity: asymmetric switch costs on a perspective-taking task. *J Exp Psychol Learn Mem Cogn*. 2019;45:213–8.
49. Verschooren S, Liefvooghe B, Brass M, Pourtois G. Attentional flexibility is imbalanced: asymmetric cost for switches between external and internal attention. *J Exp Psychol Hum Percept Perform*. 2019;45:1399–414.
50. Baltes PB, Lindenberger U, Staudinger UM. Life span theory in developmental psychology. In W. Damon (Ed.), *Handbook of child psychology*, Vol. 1 (pp. 1029–1143), 2007. New York, NY: John Wiley & Sons.
51. Ford DH, Lerner RM. *Developmental systems theory: an integrative approach*. Sage Publications, Inc.; 1992.
52. Magnusson D, Cairns, RB. Developmental science: toward a unified framework. In RB Cairns, GH Elder, EJ Costello (Eds), *Developmental science* (pp. 7–30). 1996. Cambridge University Press.
53. Berman GJ, Bialek W, Shaevitz JW. Predictability and hierarchy in *Drosophila* behavior. *Proc Natl Acad Sci USA*. 2016;113:11943–8.
54. Berman GJ, Choi DM, Bialek W, Shaevitz JW. Mapping the stereotyped behavior of freely moving fruit flies. *J R Soc Interface*. 2014;11:20140672.
55. Baum WM. Multiscale behavior analysis and molar behaviorism: an overview. *J Exp Anal Behav*. 2018;110:302–22.
56. Maitland C, Stratton G, Foster S, Braham R, Rosenberg M. A place for play? The influence of the home physical environment on children's physical activity and sedentary behaviour. *Int J Behav Nutr Phys Act*. 2013;10:99.
57. Brown TM, Fee E. Social movements in health. *Annu Rev Public Health*. 2014;35:385–98.
58. Allan J, Querret D, Banas K, de Bruin M. Environmental interventions for altering eating behaviours of employees in the workplace: a systematic review. *Obes Rev*. 2017;18:214–26.
59. Tang SX, Seelaus KH, Moore TM, Taylor J, Moog C, O'Connor D, Burkholder M, et al. Theatre improvisation training to promote social cognition: a novel recovery-oriented intervention for youths at clinical risk for psychosis. *Early Interv Psychiatry*. 2020;14:163–71.
60. Mardia KV, Kent JT, Bibby JM. *Multivariate analysis*. Academic Press; 1979.
61. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994;2:189–210.

62. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media; 2009.
63. Vuilleumier P, Armony JL, Driver J, Dolan RJ. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*. 2001;30:829–41.
64. Anticevic A, Cole MW, Murray JD, Corlett PR, Wang X-J, Krystal JH. The role of default network deactivation in cognition and disease. *Trends Cogn Sci*. 2012;16:584–92.
65. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA*. 2005;102:9673–8.
66. Spreng RN, Mar RA, Kim AS. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J Cogn Neurosci*. 2009;21:489–510.
67. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 2011;8:665.
68. Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, et al. Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*. 2013;80:169–89.
69. Satterthwaite TD, Elliott MA, Ruparel K, Loughhead J, Prabhakaran K, Calkins ME, et al. Neuroimaging of the Philadelphia neurodevelopmental cohort. *Neuroimage*. 2014;86:544–53.
70. Murphy AC, Bertolero MA, Papadopoulos L, Lydon-Staley DM, Bassett DS. Multimodal network dynamics underpinning working memory. *Nat Commun*. 2020;11:3035.
71. Satterthwaite TD, Wolf DH, Erus G, Ruparel K, Elliott MA, Gennatas ED, et al. Functional maturation of the executive system during adolescence. *J Neurosci*. 2013;33:16249–61.
72. Lamichhane B, Westbrook A, Cole MW, Braver T. Exploring brain-behavior relationships in the n-back task. *Neuroimage*. 2020;212:116683.
73. Edel M-A, Blackwell B, Schaub M, Emons B, Fox T, Tornau F, et al. Antidepressive response of inpatients with major depression to adjuvant occupational therapy: a case-control study. *Ann Gen Psychiatry*. 2017;16:1.
74. West R, Baker CL, Cappelleri JC, Bushmakin AG. Effect of varenicline and bupropion SR on craving, nicotine withdrawal symptoms, and rewarding effects of smoking during a quit attempt. *Psychopharmacology*. 2008;197:371–7.
75. Tschanz JT, Pfister R, Wanzek J, Corcoran C, Smith K, Tschanz BT, et al. Stressful life events and cognitive decline in late life: moderation by education and age. The Cache County Study. *Int J Geriatr Psychiatry*. 2013;28:821–30.
76. Alexander DM, Jurica P, Trengove C, Nikolae AR, Gepshtein S, Zvyagintsev M, et al. Traveling waves and trial averaging: the nature of single-trial and averaged brain responses in large-scale cortical signals. *Neuroimage*. 2013;73:95–112.
77. Steffener J, Tabert M, Reuben A, Stern Y. Investigating hemodynamic response variability at the group level using basis functions. *Neuroimage*. 2010;49:2113–22.
78. Aguirre GK, Zarahn E, D'Esposito M. The variability of human, bold hemodynamic responses. *Neuroimage*. 1998;8:360–9.
79. Handwerker DA, Ollinger JM, D'Esposito M. Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*. 2004;21:1639–51.
80. Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA*. 2016;113:7900–5.
81. Kessler D, Angstadt M, Sripada CS. Reevaluating “cluster failure” in fMRI using nonparametric control of the false discovery rate. *Proc Natl Acad Sci USA*. 2017;114:E3372–E3373.
82. Fox MD, Greicius M. Clinical applications of resting state functional connectivity. *Front Syst Neurosci*. 2010;4:19.
83. Simon HA. The organization of complex systems. In: *Models of discovery*, Vol. 54. Dordrecht: Springer; 1977.
84. Grellmann C, Bitzer S, Neumann J, Westlye LT, Andreassen OA, Villringer A, et al. Comparison of variants of canonical correlation analysis and partial least squares for combined analysis of MRI and genetic data. *Neuroimage*. 2015;107:289–310.
85. Wang H-T, Smallwood J, Mourao-Miranda J, Xia CH, Satterthwaite TD, Bassett DS, et al. Finding the needle in a high-dimensional haystack: canonical correlation analysis for neuroscientists. *Neuroimage*. 2020;216:116745.
86. Newman, MEJ. *Networks: an introduction*. Oxford University Press; 2010.
87. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are ‘good’ depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *J Affect Disord*. 2016;189:314–20.
88. Hasmi L, Drukker M, Guloksuz S, Menne-Lothmann C, Decoster J, van Winkel R, et al. Network approach to understanding emotion dynamics in relation to childhood trauma and genetic liability to psychopathology: replication of a prospective experience sampling analysis. *Front Psychol*. 2017;8:1908.
89. Lydon-Staley DM, Xia M, Mak HW, Fosco GM. Adolescent emotion network dynamics in daily life and implications for depression. *J Abnorm Child Psychol*. 2019;47:717–29.
90. Pe ML, Kircanski K, Thompson RJ, Bringmann LF, Tuerlinckx F, Mestdagh M, et al. Emotion-network density in major depressive disorder. *Clin Psychol Sci*. 2015;3:292–300.
91. Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol*. 2013;9:91–121.
92. Bringmann LF, Elmer T, Epskamp S, Krause RW, Schoch D, Wichers M, et al. What do centrality measures measure in psychological networks? *J Abnorm Psychol*. 2019;128:892.
93. Lutkepohl H. *New introduction to multiple time series analysis*. Springer Science & Business Media; 2005.
94. Blaauw FJ, van der Krieke L, Emerencia AC, Aiello M, de Jonge P. Personalized advice for enhancing well-being using automated impulse response analysis—AIRA. 2017. Preprint at <https://arxiv.org/abs/1706.09268>.
95. Bos FM, Blaauw FJ, Snippe E, Van der Krieke L, De Jonge P, Wichers M. Exploring the emotional dynamics of subclinically depressed individuals with and without anhedonia: an experience sampling study. *J Affect Disord*. 2018;228:186–93.
96. Yang X, Ram N, Loughheed JP, Molenaar P, Hollenstein T. Adolescents’ emotion system dynamics: network-based analysis of physiological and emotional experience. *Dev Psychol*. 2019;55:1982.
97. Lydon-Staley DM, Leventhal AM, Piper M, Schnoll RA, Bassett DS. More than the sum of its parts: a network perspective on tobacco withdrawal. 2019. Preprint at <https://psyarxiv.com/t2a9q>.
98. Yang X, Ram N, Gest SD, Lydon-Staley DM, Conroy DE, Pincus AL, et al. Socio-emotional dynamics of emotion regulation and depressive symptoms: a person-specific network approach. *Complexity* 2018;2018:5094179.
99. Allen JJB, Chambers AS, Towers DN. The many metrics of cardiac chronotropy: a pragmatic primer and a brief comparison of metrics. *Biol Psychol*. 2007;74:243–62.
100. Dooren Mvan, Janssen JH, et al. Emotional sweating across the body: comparing 16 different skin conductance measurement locations. *Physiol Behav*. 2012;106:298–304.
101. Hollenstein T, Lanteigne D. Models and methods of emotional concordance. *Biol Psychol*. 2014;98:1–5.
102. Helion C, Krueger SM, Ochsner KN. Emotion regulation across the life span. *Handb Clin Neurol*. 2019;163:257–80.
103. de Zwart JA, Silva AC, van Gelderen P, Kellman P, Fukunaga M, Chu R, et al. Temporal dynamics of the bold fMRI impulse response. *Neuroimage*. 2005;24:667–77.
104. Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human v1. *J Neurosci*. 1996;16:4207–21.
105. Cole MW, Ito T, Bassett DS, Schultz DH. Activity flow over resting-state networks shapes cognitive task activations. *Nat Neurosci*. 2016;19:1718.
106. Ito T, Kulkarni KR, Schultz DH, Mill RD, Chen RH, Solomyak LI, et al. Cognitive task information is transferred between brain regions via resting-state network topology. *Nat Commun*. 2017;8:1–14.
107. Karrer TM, Kim JZ, Stiso J, Kahn AE, Pasqualetti F, Habel U, et al. A practical guide to methodological considerations in the controllability of structural brain networks. *J Neural Eng*. 2020;17:026031.
108. Gu S, Pasqualetti F, Cieslak M, Telesford QK, Alfred BY, Kahn AE, et al. Controllability of structural brain networks. *Nat Commun*. 2015;6:1–10.
109. Tang E, Giusti C, Baum GL, Gu S, Pollock E, Kahn AE, et al. Developmental increases in white matter network controllability support a growing diversity of brain dynamics. *Nat Commun*. 2017;8:1252.
110. Cornblath EJ, Tang E, Baum GL, Moore TM, Adebimpe A, Roalf DR, et al. Sex differences in network controllability as a predictor of executive function in youth. *Neuroimage*. 2019;188:122–34.
111. Jeganathan J, Perry A, Bassett DS, Roberts G, Mitchell PB, Breakspear M. Frontolimbic dysconnectivity leads to impaired brain network controllability in young people with bipolar disorder and those at high genetic risk. *Neuroimage Clin*. 2018;19:71–81.
112. Gollo LL, Roberts JA, Cocchi L. Mapping how local perturbations influence systems level brain dynamics. *Neuroimage*. 2017;160:97–112.
113. Papadopoulos L, Lynn CW, Battaglia D, Bassett DS. Relations between large scale brain connectivity and effects of regional stimulation depend on collective dynamical state. 2020. Preprint at <https://arxiv.org/abs/2002.00094>.
114. Breakspear M. Dynamic models of large-scale brain activity. *Nat Neurosci*. 2017;20:340.
115. Muldoon SF, Pasqualetti F, Gu S, Cieslak M, Grafton ST, Vettel JM, et al. Stimulation-based control of dynamic brain networks. *PLoS Comput Biol*. 2016;12:e1005076.

116. Lutz W, Schwartz B, Hofmann SG, Fisher AJ, Husen K, Rubel JA. Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: a methodological proof-of-concept study. *Sci Rep.* 2018;8:1–9.
117. Bringmann LF, Ferrer E, Hamaker EL, Borsboom D, Tuerlinckx F. Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. *Multivar Behav Res.* 2018;53:293–314.
118. Fried EI. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J Affect Disord.* 2017;208:191–7.
119. Kim JZ, Bassett DS. Linear dynamics and control of brain networks. In: He B, editor. *Neural engineering.* Springer Nature; 2019.
120. Tang E, Bassett DS. Control of dynamics in brain networks. *Rev Mod Phys* 2018;90:031003.
121. Betzel RF, Gu S, Medaglia JD, Pasqualetti F, Bassett DS. Optimally controlling the human connectome: the role of network topology. *Sci Rep.* 2016;6:30770.
122. Gu S, Betzel RF, Mattar MG, Cieslak M, Delio PR, Grafton ST, et al. Optimal trajectories of brain state transitions. *Neuroimage.* 2017;148:305–17.
123. Stiso J, Khambhati AN, Menara T, Kahn AE, Stein JM, Das SR, et al. White matter network architecture guides direct electrical stimulation through optimal state transitions. *Cell Rep.* 2019;28:2554–66.
124. Cornblath EJ, Ashourvan A, Kim JZ, Betzel RF, Ciric R, Adebimpe A, et al. Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Commun Biol.* 2020;3:261.
125. Yan G, Vertes PE, Towilson EK, Chew YL, Walker DS, Schaefer WR, et al. Network control principles predict neuron function in the *Caenorhabditis elegans* connectome. *Nature.* 2017;550:519–23.
126. Ju H, Kim JZ, Bassett DS. Network structure of cascading neural systems predicts stimulus propagation and recovery. 2018. Preprint at <https://arxiv.org/pdf/1812.09361.pdf>.
127. Kim JZ, Soffer JM, Kahn AE, Vettel JM, Pasqualetti F, Bassett DS. Role of graph architecture in controlling dynamical networks with applications to neural systems. *Nat Phys.* 2018;14:91
128. Towilson EK, Vertes PE, Yan G, Chew YL, Walker DS, Schaefer WR, et al. *Caenorhabditis elegans* and the network control framework-FAQs. *Philos Trans R Soc Lond B Biol Sci.* 2018;373:1758.
129. Sigaud O, Stulp F. Policy search in continuous action domains: an overview. *Neural Netw.* 2019;113:28–40.
130. Becker CO, Bassett DS, Preciado VM. Large-scale dynamic modeling of task-fMRI signals via subspace system identification. *J Neural Eng.* 2018;15:066016.
131. Ashourvan A, Pequito S, Bertolero M, Kim JZ, Bassett DS, Litt B. A dynamical systems framework to uncover the drivers of large-scale cortical activity. 2019. Preprint at <https://www.biorxiv.org/content/10.1101/638718v1>.
132. Sani OG, Yang Y, Lee MB, Dawes HE, Chang EF, Shanechi MM. Mood variations decoded from multi-site intracranial human brain activity. *Nat Biotechnol.* 2018;36:954–61.
133. Anumanchipalli GK, Chartier J, Chang EF. Speech synthesis from neural decoding of spoken sentences. *Nature.* 2019;568:493–8.
134. Braun U, Harnett A, Pergola G, Menara T, Schaefer A, Betzel RF, et al. Brain state stability during working memory is explained by network control theory, modulated by dopamine d1/d2 receptor function, and diminished in schizophrenia. 2019. Preprint at <https://arxiv.org/abs/1906.09290>.
135. Zoller D, Sandini C, Schaefer M, Eliez S, Bassett DS, Van De Ville D. Structural control energy of resting-state functional brain states reveals inefficient brain dynamics in psychosis vulnerability. 2019. Preprint at <https://www.biorxiv.org/content/10.1101/703561v1>.
136. Cornblath EJ, Lydon-Staley DM, Bassett DS. Harnessing networks and machine learning in neuropsychiatric care. *Curr Opin Neurobiol.* 2019;55:32–39.
137. Fisher AJ, Boswell JF. Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment.* 2016;23:496–506.
138. Kroeze R, van der Veen DC, Servaas MN, Bastiaansen JA, Voshaar RO, Borsboom D, Riese H. Personalized feedback on symptom dynamics of psychopathology: a proof-of-principle study. *J Pers Oriented Res.* 2017;3:1–11.
139. Beck ED, Jackson JJ. Consistency and change in idiographic personality: a longitudinal ESM network study. *J Pers Soc Psychol.* 2019;118:1080–100.
140. Henry T, Robinaugh D, Fried E. On the control of psychological networks. 2020. Preprint at <https://psyarxiv.com/7vpz2/>.
141. Chang J-Y, Pigorini A, Massimini M, Tononi G, Nobili L, Van Veen BD. Multivariate autoregressive models with exogenous inputs for intracerebral responses to direct electrical stimulation of the human brain. *Front Hum Neurosci.* 2012;6:317.
142. Yang Y, Sani OG, Chang EF, Shanechi MM. Dynamic network modeling and dimensionality reduction for human ecog activity. *J Neural Eng.* 2019;16:056014.
143. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczurkin AN, et al. Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat Commun.* 2018;9:3003.
144. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TE, Glasser MF, et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci.* 2015;18:1565–7.
145. Bianconi G. *Multilayer networks: structure and function.* Oxford University Press; 2018.
146. Menichetti G, Dall’Asta L, Bianconi G. Control of multilayer networks. *Sci Rep.* 2016;6:20706.
147. Gagnier JJ, Moher D, Boon H, Beyene J, Bombardier C. Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. *BMC Med Res Methodol* 2012;12:111.
148. Vu M-AT, Adali T, Ba D, Buzsaki G, Carlson D, Heller K, et al. A shared vision for machine learning in neuroscience. *J Neurosci.* 2018;38:1601–7.
149. Ertekin S, Rudin C. A Bayesian approach to learning scoring systems. *Big Data.* 2015;3:267–76.
150. Platt JR. Strong inference. *Science.* 1964;146:347–53.
151. Bertolero MA, Bassett DS. On the nature of explanations offered by network science: a perspective from and for practicing neuroscientists. *Top Cogn Sci.* 2020. Advanced publication. <https://doi.org/10.1111/tops.12504>.
152. Ross LN. Causal selection and the pathway concept. *Philos Sci.* 2018;85:551–72.
153. Ross LN. Causal concepts in biology: how pathways differ from mechanisms and why it matters. *Br J Philos Sci.* 2018;0:1–30.
154. Dworkin JD, Linn KA, Teich EG, Zurn P, Shinohara RT, Bassett DS. The extent and drivers of gender imbalance in neuroscience reference lists. *Nat Neurosci.* 2020;23:918–26.
155. Maliniak D, Powers R, Walter BF. The gender citation gap in international relations. *Int Organ.* 2013;67:889–922.
156. Caplar N, Tacchella S, Birrer S. Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nat Astron.* 2017;1:0141.
157. Chakravarty P, Kuo R, Grubbs V, McIlwain. C. #CommunicationSoWhite. *J Commun.* 2018;68:254–66.
158. Thiem Y, Sealey KF, Ferrer AE, Trott AM, Kennison R. Just Ideas? The status and future of publication ethics in philosophy: a white paper. Technical report. 2018. <https://publication-ethics.org/white-paper/>.
159. Dion ML, Sumner JL, Mitchell SM. Gendered citation patterns across political science and social science methodology fields. *Political Anal.* 2018; 26:312–27.
160. Zhou D, Cornblath EJ, Stiso J, Teich EG, Dworkin JD, Blevins AS, et al. Gender diversity statement and code notebook v1.0. 2020. <https://github.com/dalejn/cleanBib>.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020