CellPress

Neuron
Perspective

Computational Psychiatry

Xiao-Jing Wang[1,2,3,*] and John H. Krystal[3,4,5,6]
[1]NYU-ECNU Institute of Brain and Cognitive Science, NYU-Shanghai, Shanghai, China
[2]Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA
[3]Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA
[4]Department of Psychiatry, Yale University School of Medicine, 300 George Street, Suite #901, New Haven, CT 06520, USA
[5]Psychiatry Service, Yale-New Haven Hospital, New Haven, CT 06510, USA
[6]Clinical Neuroscience Division, VA National Center for PTSD, VA Connecticut Healthcare System, West Haven, CT 06516, USA
*Correspondence: xjwang@nyu.edu
http://dx.doi.org/10.1016/j.neuron.2014.10.018

Psychiatric disorders such as autism and schizophrenia, arise from abnormalities in brain systems that underlie cognitive, emotional, and social functions. The brain is enormously complex and its abundant feedback loops on multiple scales preclude intuitive explication of circuit functions. In close interplay with experiments, theory and computational modeling are essential for understanding how, precisely, neural circuits generate flexible behaviors and their impairments give rise to psychiatric symptoms. This Perspective highlights recent progress in applying computational neuroscience to the study of mental disorders. We outline basic approaches, including identification of core deficits that cut across disease categories, biologically realistic modeling bridging cellular and synaptic mechanisms with behavior, and model-aided diagnosis. The need for new research strategies in psychiatry is urgent. Computational psychiatry potentially provides powerful tools for elucidating pathophysiology that may inform both diagnosis and treatment. To achieve this promise will require investment in cross-disciplinary training and research in this nascent field.

Introduction

In 1988, a computational neuroscience "manifesto" (Sejnowski et al., 1988) mentioned three reasons for the emergence of this new research field: advances in neuroscience had generated a large body of neurophysiologic data, new computers possessed sufficient power to conduct neural model simulations, and simplified brain models were introduced that provided insights into complex neural circuit functions. Since then, dramatic advances made on all three fronts fundamentally changed the computational neuroscience landscape (Abbott, 2008). Notably, computational neuroscience initially focused on the early stages of sensory processing (Sejnowski et al., 1988), because studies of the neural bases of higher cognitive functions were beyond empirical neuroscience of that era. Indeed, only in recent years has the confluence of single-unit physiology, human functional brain imaging, and advances in computational modeling made significant strides in tackling executive functions (such as working memory and decision making) that underlie cognitively controlled flexible behavior. These higher functions critically depend on the prefrontal cortex (PFC) (Fuster, 2008; Miller and Cohen, 2001; Wang, 2013; Szczepanski and Knight, 2014). Because impairments of the PFC and related circuits are implicated in major psychiatric disorders, such as schizophrenia and autism (Goldman-Rakic, 1994; Insel, 2010; Courchesne et al., 2011; Anticevic et al., 2013a), the newly acquired insights and computational models offer an opportunity to elucidate how cellular and circuit level pathologies give rise to cognitive deficits observed in mental illness, advances in this direction could inform studies of psychiatric diagnosis, pathophysiology and treatment.

Therefore, the time is ripe for computational psychiatry to emerge as a field at the interface between basic and clinical neuroscience (Montague et al., 2012; Friston et al., 2014). In this Perspective, we review recent work demonstrating that computational psychiatry introduces novel approaches and tools to investigate neural circuit mechanisms underlying the cognitive and behavioral features of neuropsychiatric disorders. First, we will spell out the rationale of a computational approach to psychiatry, i.e., "why computational psychiatry? What theories and models are relevant to this field?" Second, we will discuss how theories and models have been applied to the investigation of behavioral impairments in terms of transdiagnostic endophenotypes. Third, we will summarize recent work that advocates for a model-aided framework of diagnosis and treatment. The fourth part will be devoted to biophysically based neural circuit modeling that we argue represents the optimal approach for cross-level understanding from cellular processes to collective and emergent circuit dynamics and ultimately to behavior. Fifth and finally, we will end with practical recommendations related to the training and funding needed to foster this nascent field.

Why Computational Psychiatry?

It is widely acknowledged that current psychiatric diagnostic schema and the treatments for psychiatric disorders lack a firm biological foundation. The complexity of the brain presents unique challenges to the development of highly specific mechanistic hypotheses to guide research in psychiatry. Advances in genetics, and molecular and cellular neurosciences are providing, at long last, clues to the etiology of human cognitive, emotional, and behavioral problems. For example, candidate-gene studies have revealed gene variations (such as DISC1; Brandon et al., 2009) associated with psychiatric disorders. However, many in the field think that attempts to seek single genes underlying complex psychiatric phenotypes have been largely

disappointing, and that efforts to link genes to more basic cognitive and behavioral functions and functional impairments could be more promising. The progress in these areas has yet to provide a firm basis for a diagnostic system or a single pharmacotherapy for common psychiatric disorders (Krystal and State, 2014).

A major hindrance in our capacity to develop novel pharmacotherapies for psychiatric disorders is the still superficial nature of our understanding of how circuits produce behavior. In this regard, synaptic and systems physiology are producing remarkable advances in our specific understanding of the functional properties of microcircuits and the beginnings of connecting these insights into behavioral processes including basic visual perception (Parker and Newsome, 1998), fear conditioning and extinction (Johansen et al., 2011), and mental representations in working memory (Arnsten et al., 2010). There are even examples where aspects of the neural representation of distinct fear memories can be ascribed to the functional integrity of a few distinct sets of cells in the amygdala (Josselyn, 2010). Yet, perhaps as a consequence of the limitations of our animal models combined with the limited spatial and temporal resolution of current neuroimaging technologies (MRI, magnetoencephalograpy, positron emission tomography), there is not a single symptom of a single psychiatric disorder for which we fully understand its physiologic basis at a molecular, cellular, and microcircuit level. In other words, we have only a somewhat vague idea of how the brain generates the cognitive, emotional, and behavioral problems that lead people to seek treatment by psychiatrists and other mental health clinicians.

As a consequence of our limited understanding of how circuits represent information, there are a plethora of attempts to explain circuit dysfunction in psychiatric disorders in superficial ways, giving rise to an equally large number of relatively risky potential pharmacologic strategies to address the unmet need for more effective treatments. The implications of this knowledge gap are profound for the field of psychiatry and for society. For example, psychiatric diagnoses have categorical qualities as exemplified by the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5). Although this new version of the DSM takes into consideration the recent explosions in the genetics of disorders, such as autism and schizophrenia (Krystal and State, 2014), it is widely criticized for lack of a solid biological foundation based on either etiology or pathophysiology. Categorizing patients by symptom checklists results in enormous clinical heterogeneity within diagnostic categories, surprisingly poor interrater reliability for many common psychiatric diagnoses (Freedman et al., 2013), and very likely, poorer clinical outcomes.

An alternative schema has emerged from the recognition that behavioral impairments are traits that may be shared across psychiatric disorders (Krueger, 1999). The shift from a categorical diagnostic focus to a dimensional transdiagnostic approach emerged in the form of the Research Domain Criteria (RDoC, http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml) (Insel et al., 2010; Insel 2014). The RDoC program aims at identifying core cognitive, emotional, and social dysfunctions, then elucidating their brain mechanisms bridging different levels (from molecules, cells, circuits to functions). Yet, the next step in this process is to determine whether the circuits are dysfunctional in the same way across disorders or whether, when characterized in increasingly accurate molecular and physiological ways, categorical features of psychiatric diagnoses reemerge. Furthermore, diagnoses may have both categorical and dimensional features. For example, schizophrenia appears to be a more severe form of circuit dysfunction than bipolar disorder with respect to the thalamo-cortical functional connectivity (Anticevic et al., 2013b), but a completely distinct type of disorder than bipolar disorder with respect to the variance or "noise" level of cortical activity (Yang et al., 2014). Neither DSM nor RDoC in its current form provides guidance as to how to integrate the dimensional and categorical features of psychiatric pathophysiology. A second consequence is the lack of precision with which one can predict whether a particular treatment mechanism will work for psychiatric disorders. It is not just that biomarkers of illness are lacking, but rather the biomarkers that we have are not sufficiently mechanistically precise as to specify a particular treatment. In addition, even when aspects of molecular pathology are characterized, the impact on micro-and macrocircuit functions and the paths to correct that circuit dysfunction are not clear. As a result, in the case of schizophrenia, it is not clear that GABA signaling deficits (Lewis et al., 2005, Lewis and Gonzalez-Burgos, 2006) should be treated by $GABA_A$ receptor agonists nor deficits in NMDA receptor (NMDAR) signaling should be treated with drugs that increase the stimulation of the glycine coagonist site of the NMDAR (Buchanan et al., 2011; Goff, 2014).

The gap between genetic, molecular, and cellular studies, on the one hand, and systems and behavioral neuroscience studies, on the other, currently cannot be bridged purely through experimentation. Take, again, the example of the PFC. Its crucial role in a wide range of executive functions (Fuster, 2008; Miller and Cohen, 2001; Wang, 2013) begs the question: what are the key properties that enable the PFC to subserve cognitive processes, in contrast to primary sensory or motor systems? This question is difficult to address by laboratory experiments alone, partly because PFC circuitry is endowed with powerful positive and negative feedback loops and the behavior of any such dynamical system is not predictable by intuition alone. While physiological studies in animals and humans yield data on the correlation of particular measurements to specific cognitive operations, theory and modeling are usually needed, together with experimentation, to investigate the "follow-up" questions: what circuit mechanisms give rise to the observed neuronal and other brain signals? What are the computational algorithms and generalizable principles that are reflected in the observed biological signals and sufficient to explain behavior?

Computational modeling offers a suitable approach to quantitatively explore the properties of complex systems across levels of investigation. Therefore, by incorporating computational neuroscience modeling within translational neuroscience research programs, it may be possible to develop more specific hypotheses related to circuit dysfunction in model systems and psychiatric disorders. There are many forms of computational models; we will present two types. Models of Mathematical Psychology or algorithmic models from Computer Science are enormously useful for quantifying behavioral data and relating their fitted parameters to neural computations (Maia and Frank, 2011; Montague et al., 2012). On the other hand, biophysically informed computational modeling, that are constrained by the