Published in final edited form as:
Neuron. 2014 November 5; 84(3): 638–654. doi:10.1016/j.neuron.2014.10.018.

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

[6]Clinical Neuroscience Division, VA National Center for PTSD, VA Connecticut Healthcare System, West Haven, CT 06516

Abstract

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, 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.

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

Neither Dr. Wang nor Dr. Krystal has any financial interests relevant to the advancement of computational psychiatry. to and engage with psychiatric research. Second, it will be important to develop a cadre of young psychiatrists who learn computational modeling, lest computational psychiatry.

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. 2013), 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, 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 feel 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 extremely superficial nature of our understanding of how circuits represent 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, MEG, PET), 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 for Psychiatric Disorders 5 (DSM-5). Although this new version of 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 inter-rater 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. Further, 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. 2013), 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 or 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. Further, even when aspects of molecular pathology are characterized, the impact on micro-and macro-circuit 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 signaling should be treated with drugs that increase the stimulation of the glycine co-agonist site of the NMDA receptor (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 prefrontal cortex (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, the present article covers two kinds. 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 biophysical properties of identified synaptic signaling mechanisms and other properties of microcircuits, has proven to be an effective approach to understanding the neurobiology underlying cortical functions and psychiatric disorders (Wang 2006, Anticevic et al. 2013).

Biologically-based neural circuit models

What is biologically-based neural circuit modeling? Simply put, it is a computational framework that is constrained by neurobiology and designed to achieve a cross-level understanding of brain functions in terms of neural dynamics, computation, and biological mechanisms (Figure 1). One may question whether such models are too complex to be useful in cognitive science or psychiatry (Carandini 2012, Montague et al. 2012). Three points are worth noting on this regard. First, biologically-based modeling is a broad term that embraces a diversity of models with varying degrees of complexity. A model does not necessarily improve when more biological details are included. There is always a tradeoff between incorporating important details in order for the model to be suitable (given a scientific question) on one hand and simplicity and generalizability on the other hand. It is also tremendously useful to be able to go back and forth between models differing in their levels of abstraction, for instance between a spiking network model and its reduced “mean-field” firing-rate model for population-level dynamics. Second, neuronal modeling is most appropriate for those functions for which we have some knowledge about the underlying neural processes, such as dopamine neural signaling of reward-prediction error, persistent activity subserving the internal representation of working memory and neural integrators in perceptual decision-making. By contrast, modeling at the neuronal level would seem premature for other behavioral phenomena such as hallucination or feeling of depression, in the absence of neurophysiological characterization. Finally, to the extent that biophysically-based neural circuit modeling begins by incorporating the simplest and most fundamental features of synaptic connectivity, it is arguably the simplest possible framework that permits us to elucidate the inter-relationship between biological mechanism, neural dynamics and computations, and circuit functional output (Figure 1A).

In a spiking network model, single neurons are often described by either the leaky integrate-and-fire model or the Hodgkin-Huxley model. These models are calibrated by physiological measurements, such as the membrane time constant and the input-output function (the spike firing rate as a function of the synaptic input current), which can be different for excitatory pyramidal cells and inhibitory interneurons. Furthermore, it is worth emphasizing that in biophysically-based models, synapses must be modeled accurately. Unlike connectionist models in which coupling between neurons is typically an instantaneous function of firing activity, synapses have their own rise-times and decay time constants, and they exhibit summation properties. Synaptic dynamics are crucial factors in determining the integration time of a neural circuit and the stability of a strongly recurrent network (Wang 1999). Finally, networks endowed with a biologically plausible architecture need to be constructed based on quantitative anatomy (Douglas and Martin 2004). For example, a commonly assumed circuit organization is local excitation between neurons of similar selectivity combined with a more global inhibition. Dynamic balance between synaptic excitation and inhibition is another feature of cortical microcircuits that has been increasingly recognized experimentally and incorporated in cortical network models (http://www.scholarpedia.org/article/Balance_of_excitation_and_inhibition).

Consider decision-making, the process of reaching a particular choice among several alternative options, such as rendering a judgment out of multiple possibilities given incomplete information or choosing one of actions expected to yield different outcomes (Glimcher 2003, Gold and Shadlen 2007, Wang 2008, Glimcher and Fehr 2013). Broadly speaking, there are two types of computational models of decision-making: behavioral models and neural circuit models. In Behavioral Psychology, decision-making is commonly modeled by the drift diffusion model (DDM) (Ratcliff 1978, Smith and Ratcliff 2004). In this model, an activity variable X represents the difference between the respective amounts of accumulated information about the two alternatives, say X_A and X_B, $X = X_A - X_B$. The dynamics of X is given by the drift diffusion equation, $\mathrm{d}X/\mathrm{dt}= \mu + \mathrm{w(t)}$, where μ is the drift rate, w(t) represents noise. The drift rate μ represents the bias (net difference in the evidence) in favor of one of the two choices (and is zero if there is no net bias). For instance, in a random-dot motion direction discrimination task, μ is proportional to the strength of motion signal. This system is a perfect integrator of the input. The integration process is terminated and the decision time is read out, whenever $X(t)$ reaches a positive threshold θ (choice A) or a negative threshold −θ (choice B). If the drift rate μ is positive, then choice A is correct, whereas choice B is an error. Therefore, this type of models is commonly referred to as ramping-to-threshold model, with the average ramping slope given by μ.

A biophysically-based neural circuit model has been proposed for decision-making (Wang 2002). This model reproduces not only behavioral observations but also single neural activity associated with decision-making observed in a monkey experiment (Roitman and Shadlen 2002). Moreover, it suggests a specific biological basis for temporal accumulation of evidence in decision-making. The drift diffusion model is an ideal perfect integrator (with an infinite time constant), whereas neurons and synapses are leaky with short time constants of tens of milliseconds. The neural circuit model suggests that a long integration time can be realized in a decision network through recurrent excitation. Reverberating excitation represents a salient characteristic of cortical local circuits (Douglas et al. 1995, Douglas and Martin 2004). When this positive feedback is sufficiently strong, recurrent excitation in interplay with synaptic inhibition can create multiple stable states ("attractors"). Such models have been initially proposed for working memory. The same models, provided that excitatory reverberation is slow (i.e. mediated by the NMDA receptors), has been shown to be capable of decision-making computations (Wang 2002, Machens et al. 2005, Miller and Wang 2006, Wong and Wang 2006, Soltani and Wang 2006, Deco et al. 2007, Wang 2008, Furman and Wang 2008, Deco et al. 2009, Engel and Wang 2011, Hunt et al. 2012). Interestingly, physiological studies in behaving non-human primates often reported neural activity correlated with decision making in cortical areas such as the prefrontal cortex or the parietal cortex, that also exhibit mnemonic persistent activity during working memory. Hence, this model and supporting experimental data suggest a common, "cognitive-type" circuit mechanism for decision-making and working memory in the brain (Wang 2013).

Behavioral modeling is often powerful in describing computations that solve a problem normatively or algorithmically. On the other hand, neural circuit models may be more suited for enabling us to investigate the underlying neural mechanisms and potentially pharmacologic or genetic manipulations of the circuits. Importantly, neural circuit models are not merely implementations of abstract mathematical models. For instance, the two types of models of perceptual decision-making have distinct predictions at the behavioral level (Wang 2008). These approaches are usually developed independently, but we are witnessing some convergence of the two in recent years. For example, spiking network models have been shown to have the capability of fitting quantitatively with behavioral performance (accuracy and reaction time) data (Lo et al. 2009), whereas such data fitting and model comparisons are commonly done with more abstract models due to their lower computational cost. Spiking network models can also be reduced to population rate models (Wong and Wang 2006), that have features of abstract connectionist models. On the other hand, connectionist neural network models have increasingly taken biological information (with identified brain structures, receptors, etc) into account (O'Reilly and Frank 2006). Thus, to bridge gaps in the current knowledge base and to facilitate research, there are advantages to move back and forth across several models that vary in their degree of abstraction, biological realism, their level of analysis (circuits, computational operations, behaviors).

Endophenotypes across brain disorder categories

Inasmuch as features of the pathophysiology of psychiatric disorders are shared across diagnostic boundaries (Krueger 1999), a promising research direction is to search for trans-diagnostic endophenotypes, i.e., quantitative heritable traits that are intermediate between risk genotypes and the psychiatric disorder syndrome itself (Figure 2A, Gottesman and Gould 2003). While it has yet to be demonstrated that endophenotypes have a more simple genetics than psychiatric diagnoses, there remains a hope that endophenotypes may be more precisely defined, measured, and related to the underlying biology and to animal models. For instance, impulsivity and compulsivity are behavioral endophenotypes that cut across a range of diagnostic categories including obsessive-compulsive disorders, substance dependence, attention-deficit/hyperactivity disorder. Neither impulsivity nor compulsivity may be unitary constructs, but they may derive from a set of psychological processes which themselves are candidate endophenotypes (Figure 2B, Robbins et al. 2012). Thus, one could show impulsive choice behavior because of an aversion to delayed gratification, or impulsive response due to motor disinihibiton or timing impairment. While this dimensional approach has not supplanted the prevailing psychiatric diagnostic schema, it has powerfully stimulated psychiatry research.

It is a major challenge to accurately and reliably identify endophenotypes. To make progress, it is beneficiary to complement consideration of symptoms (how people feel) with attention to what people do (choices and actions). By using behavioral paradigms that are designed to probe a specific cognitive function or functional domain, one can quantify the abnormalities of a particular function that are shared by multiple mental disorders. Those carefully designed tasks should be doable by both human subjects and nonhuman animals, thereby enabling more productive translational research (Carter et al. 2008, Wang 2013, Insel 2014). Theories can be developed and applied to both normal subjects and patients, providing insights into the core of a brain dysfunction.

Consider the case of disturbances in decision-making. Many people, who meet current diagnostic criteria for a number of neuropsychiatric disorders, repeatedly make bad choices in the social, vocational, and recreational domains that compromise the quality of their lives. There is increasing evidence that specific impairments in decision-making may represent cognitive endophenotypes across diagnostic boundaries (Robbins et al. 2012, Montague et al. 2012). A number of studies have dealt with the valuation process in reward-based decision-making. The computations that enable one to learn to evaluate alternative options through experience are fundamental for adaptive choice behavior, i.e., to make a choice, assess its outcome, and to use this experience to guide the next choice. Reinforcement learning (RL) theory (Rescorla and Wagner 1972, Sutton and Barto 1998, Rangel et al. 2008) offers a framework for this adaptive process and impairments associated with psychiatric conditions (Montague et al. 2012, Maia and Frank 2011, Lee 2013). This field, which lies at the interface behavior and neurobiological mechanisms, was galvanized by the discovery that phasic activity of dopamine neurons in the ventral tegmental area signals reward prediction error (RPE) (Montague et al. 1996, Schultz et al. 1997). Specifically, dopamine phasic firing has been shown to confirm with RPE according to temporal-difference RL (TDRL) (Sutton and Barto 1998, Dayan and Abbott 2001). TDRL computes the reward expectation in terms of all anticipated reward events in the future, and learns to predict reward by driving RPE to zero. For the sake of simplicity, here we describe a simplified notion of RPE, $\delta_t = r_t - V_t$, where r_t is the actual reward and V_t is the expected reward, at time t). The idea is that the mismatch between the actual reward and the expected reward generates an “error signal” that informs learning. RL is hypothesized to be driven by $\alpha\delta_t$, with the rate α controlling the speed of learning. Therefore, there is a solid foundation for bridging reward-related learning with a specific underlying brain circuit (the dopamine system). Empirical evidence for impaired RL has been documented for Parkinson’s disease, schizophrenia, Tourette’s syndrome, attention deficit disorder, drug addiction, depression (Maia and Frank 2011, Lee 2013, Huys et al. 2013), demonstrating powerfully the importance of function-based, transdiagnostic, approach in psychiatry.

For instance, addiction can be viewed as RL gone awry. Indeed, a pioneering application of RL to psychiatry (Redish 2004, Redish et al. 2007) was inspired by TDRL. It was proposed that addiction access the same RL system as in the normal brain, but drug-induced positive prediction errors could produce unbounded increases in the value of drug receipt. A merit of such quantitative models is that they are precise enough to be falsifiable by new experiments, a hallmark of scientific inquiry. Redish’s model predicts that a behavioral trait called blocking does not occur when drugs are used as unconditional reinforcers. Blocking refers to the observation that after a subject learns to associate a stimulus A with a reward, later pairing A with another stimulus B should not lead to learning to associate B with the reward. If, however, drugs (as stimuli A and B) lead to unlimited value increase, blocking should not be observed. Behavioral experiments using cocaine as unconditional stimulus showed that this is not the case, i.e. blocking does occur (Panlilio et al. 2007). One possible interpretation of this result is that blocking is not due to the specific form TDRL of RL. Indeed, blocking is accounted for in an alternative model of addiction that assumes the expected reward V_t to be computed by a weighted average over past reward events

(Dezfouli et al. 2009). Another possibility is that RL involves multiple competing systems (Redish et al. 2007).

The RL approach has also been applied to depression. Huys et al. (2013) set out to test the hypothesis that depression is associated with an altered sensitivity to reward; specifically, the RPE becomes $\delta_t = \rho r_t - V_t$, where the parameter ρ represents reward sensitivity. Meta-analysis of experiments with about 50 healthy subjects and 50 subjects with major depression disorder revealed has been carried out by fitting behavioral data with a RL model. It was found that compared to the control group, the patient group shows a significantly reduced reward sensitivity (a smaller value of ρ), but no change in the learning rate α, consistent with the anhedonia and lack of motivation found in patients with depression. Similar findings were also reported by Strauss et al. (2011). This work illustrates how computational modeling enables us to dissect distinct aspects (reward sensitivity but not learning rate) of a maladaptive behavior.

The RL theory is currently been extended beyond single-factor considerations. In particular, it has been recognized that RL involves two separate neural systems (Balleine and Dickinson 1998, Daw et al. 2005, 2011, Kahneman 2011, Dolan and Dayan 2013). One of these systems subserves habits and related behaviors. It is referred to as “model-free” because these behaviors are elicited in an automatized way by cues. The second, model-based, system is endowed with an internal representation of the causal structure of the environment and underlies goaloriented behaviors. The model-free and model-based systems must be balanced. A dual-system learning model (Daw et al. 2011) has been combined with human brain imaging to examine specific ways an imbalance of these two systems might lead to maladaptive choice behavior in mental illness. Using this framework, it was found that repeated exposure to addictive drugs shifts behavior from model-based to model-free emphasis (Kurth-Nelson and Redish 2011, Lucantonio et al. 2012). Likewise, data fitting by the dual-system model revealed that subjects diagnosed with obsessive-compulsive disorder display a bias towards model-free habit acquisition (Voon et al. 2014). The central control mechanisms governing the balance maintenance and shifts between model-based and model-free systems represent an area of intense ongoing research (Simon and Daw 2011).

Whereas the model-free system relies on RPE, the model-based system presumably depends on a more abstract “state prediction error” which might implicate lateral prefrontal cortex, giving rise to “dual system” RL models (Glascher et al. 2010). RL approaches have advanced translational neuroscience research on such phenomena as delusions that have been previously extremely challenging to study from this perspective. The focus on prediction error, a mismatch between expectation and experience, has inspired neurobiological studies of psychosis (Corlett et al. 2010). Delusions are false beliefs about the world that persist tenaciously despite repeated encounters with contradicting evidence. Corlett and his colleagues (Corlett et al. 2007) found that violations of causal associations activate the right lateral prefrontal cortex (rPFC) during fMRI, a putative prediction error signal. However, deficits in this fMRI prediction error signal among subjects with first episode psychosis strongly correlated with the severity of delusions across subjects (Corlett et al. 2007). Thus, false beliefs may be generated through compromised prediction error and sustained as aberrant learning transitions from being represented by model-based to model-free systems (Corlett et al. 2010).

RL has also been extended to hierarchically organized behaviors (Botvinick et al. 2009). These studies focused on RL illustrate well how theory and computational modeling, in conjunction with experimentation, can help dissect distinct component processes (such as reward sensitivity, learning rate, balance between model-free and model-based systems, et al.), each may be abnormal in multiple mental disorders but in different ways. This opens up the possibility that each cognitive endophenotype (such as impulsivity) could be defined in terms of a specific combination of quantitative impairments of these component processes. If so, future progress in this direction could yield a promising new framework to guide translational neuroscience studies of neuropsychiatric disorders.

Big data and model-aided diagnosis

Typically, the process of building from a behavioral experiment to a computational model follows several steps: (1) a cognitive task is strategically designed to probe a particular function (e.g. reward-related learning in decision-making), (2) an appropriate computational model (e.g. reinforcement learning) is chosen to simulate the process (e.g. valuation) under consideration, (3) model-fitting of data yields estimation of model parameters (e.g. reward sensitivity and learning rate). Many of these studies compare people deemed to be free of a psychiatric diagnosis to people who have been recruited specifically for the presence of a specific psychiatric diagnosis (e.g. according to DSM or international classification of diseases (ICD) criteria). Significant differences between the healthy group and patient group in some model parameters (e.g. reward sensitivity but not learning rate) provide the basis for characterizing the presumed “abnormality” in the patient group. However, computational psychiatry is not limited to existing diagnostic schema. Its focus on relating mechanisms to cognitive operations and behavioral processes promotes a transdiagnostic perspective. For instance, a similar bias toward model-free versus model-based learning has been found in disorders involving both natural (binge eating) and artificial (methamphetamine) rewards, as well as obsessive-compulsive disorder (Voon et al. 2014).

Recently, Frank and collaborators (Wiecki et al. 2014) proposed to extend this approach from subject groups to individuals. This requires a fourth step, i.e., to use sophisticated statistical analysis algorithms to investigate whether model parameter values extracted from individual subjects are clustered into distinct groups (Figure 3A). This step is crucial for this paradigm to potentially serve as a clinical tool, since diagnosis must obviously be done for single individuals. A similar approach has been advocated by Stephan and his colleagues (Figure 3B) (Brodersen et al. 2014). These authors proposed a cross-disciplinary approach that combines behavior, brain measures (fMRI) and computation (dynamical causal modeling, DCM (Friston et al. 2003, Stephan et al. 2007)). In a working memory study of schizophrenic patients they focused on DCM based estimates of effective connectivity between visual, parietal and prefrontal cortex, since these three cortical areas were critically involved in their visual working memory task. An unsupervised clustering procedure operating on the individual connectivity patterns yielded three distinct patient subgroups (Figure 3C): (a) those with greater fronto-parietal connectivity, (b) those with weaker fronto-parietal connectivity, and (c) those with greater visuo-frontal connectivity. The authors further pushed the approach by including two more steps (Figure 3B): (5) assessment of whether clusters of subjects obtained by model-fitting are correlated with different severity of behavioral impairment (indeed they found that subjects in the three clusters display a different degree of negative symptom severity (Figure 2D)), and (6) interpretation of the results from step (5) that attributes the behavioral deficit (negative symptom) to a possible underlying brain substrate (visual-parietal-prefrontal circuitry connectivity), generating new hypotheses to be tested in future research.

This line of work raises the question of whether it might be possible to use brain imaging data (or models of such data) rather than symptoms as the substrate for diagnostic classification schema. A related line of thinking is to view psychiatric illness from the perspective of brain connectome (Rubinov and Bullmore 2013), according to which the analysis of functional connectivity patterns inferred from brain imaging offers a window to pathoconnectomics associated with mental disorders. It would be interesting to know the impact of attempting to, on a very large scale, develop model parameters that cluster patients in new ways. Would this approach yield a classification schema different from DSM? Would this classification schema be replicable and generalizable? Would it suggest new directions for research and treatment? This type of strategy might address a conundrum in psychiatry, i.e., the absence of biomarkers. It may be impossible to develop meaningful illness biomarkers within a diagnostic framework that is not based in biology. However, if the diagnostic framework were, itself, built around an imaging biomarker, then it would seem highly likely that this biomarker would have predictive power with regards to diagnosis and treatment.

A number of factors will determine the success of this framework: very large samples of subjects, efficient and statistically reliable analysis methods, and judicious choices of computational models. With the advance of big data science, and computational modeling, a radical modern paradigm shift may be on the horizon.

Biophysically-based neural circuit modeling: understanding across levels

In contrast to more abstract models, biophysically realistic neural circuit modeling has the potential to be rigorously calibrated by quantitative neurophysiology and anatomy. Ultimately, this is necessary to elucidate deficits at the molecular, cellular and circuit levels that underlie cognitive and behavioral disorders in mental illness.

Among hierarchically inter-related cognitive dysfunctions associated with schizophrenia (Millan et al. 2012), perhaps the best studied is working memory (Park and Holzman 1992, Lee and Park 2005, Lewis and Gonzalez-Burgos 2006, Barch and Ceaser 2012). Working memory, the brain’s ability to encode and sustain the neural representation of information in the absence of direct sensory stimulation and to manipulate this information in the service of future action, is a core cognitive function that depends on the PFC (Fuster 2008, Goldman-Rakic 1995, D’Esposito 2007, Baddeley 2012). Fortunately, working memory has been particularly amenable to biophysically-based neural circuit modeling, because of the richness of experimental data at multiple levels of study.

A well known working memory paradigm is the delayed oculomotor response task, in which a subject is required to remember a visual cue (a directional angle) across a delay period in order to perform a memory guided saccadic eye movement (Funahashi et al. 1989, Constantinidis and Wang 2004). A biologically-based network model of spiking neurons has been developed for this spatial working memory experiment (Fig. 4A) (Compte et al. 2000, Renart et al. 2003, Wang et al. 2004, Carter and Wang 2007, Wei et al. 2012, Kilpatrick et al. 2013, Hansel and Mato 2013, Pereira and Wang 2014). Fig. 4B shows a model simulation of the delayed oculomotor task. Initially, the network is in a resting state in which all cells fire spontaneously at low rates. A transient input drives a subpopulation of cells to fire at high rates. As a result they send recruited excitation to each other via horizontal connections. This internal excitation is large enough to sustain elevated activity, so that the firing pattern persists after the stimulus is withdrawn. Synaptic inhibition ensures that the activity does not spread to the rest of the network, and persistent activity has a localized, bell shape ("bump attractor"). At the end of a mnemonic delay period the cue information can be retrieved by reading out the peak location of the persistent activity pattern; and the network is reset back to the resting state. This type of spatial working memory network is endowed with a continuous family of bump attractors, each encoding a specific potential location.

In this model, a mnemonic persistent activity pattern is sustained internally by strong recurrent excitation, which the model predicts to be slow and dependent on the NMDA receptor mediated synaptic transmission at local synapses (Wang 1999, 2001) (Figure 4C). In a recent experiment with monkeys performing a working memory task (Wang et al. 2013), iontophoresis of drugs that blocked the NMDA receptors suppressed delay-period persistent activity of PFC (Figure 4D), in support of an important role of the NMDA receptors in PFC processes. Another monkey experiment showed that ketamine (an NMDA receptor antagonist) reduces task selectivity of PFC neurons in parallel with behavioral impairment (Skoblenick and Everling 2012). These findings are directly relevant to psychiatry. Indeed, it has been hypothesized that NMDA hypofunction underlies working memory deficits in schizophrenia (Coyle et al. 2003, Moghaddam and Krystal 2012), and sub-anesthetic dose of ketamine produces working memory impairment in healthy human subjects, similar to that seen in schizophrenia (Krystal et al. 1994). The finding that NMDA receptors are critical for mnemonic persistent activity and its selectivity offers a possible mechanistic explanation as to why NMDA signaling pathway is essential for working memory function.

Like Yin and Yang in Ancient Chinese Philosophy, the dynamic balance between synaptic excitation and inhibition within local and distributed networks is a fundamental property of cortical function. This balance is important for normal functions within a biophysically-based PFC neural circuit model because it defines many emergent properties of the network including: dynamic network stability (because if unchecked by inhibition, strong recurrent excitation would lead to runaway positive feedback), fast coherent oscillations (generated by the interplay between fast AMPA receptor mediated excitation and slower GABA$_A$ receptor mediated inhibition), stimulus-selectivity (synaptic inhibition is critical for neural tuning), and resistance to distractors (reduced responsiveness to distracting stimuli by neurons not involved in memory storage) (Compte et al. 2000, Brunel and Wang 2001, Wang 2013).

These results have functional implications for the observed pathology of inhibitory circuits associated with schizophrenia (Lewis et al. 2005, Lewis et al. 2012). In particular, enhanced distractibility represents a common behavioral deficit in schizophrenic patients (Goldman-Rakic 1987, Mesulam 2000, Luck and Gold 2008). A recent computational study examined how a reduced inhibition might lead to PFC's deficient ability to filter out distracting stimuli during working memory (Murray et al. 2014). Disinhibition induced a broadening of the neural representation for the memorandum maintained in working memory through persistent activity (Figure 5A). Importantly, this feature of the circuit was a function of the overall balance between excitation and inhibition (Figure 5B). Neural broadening, in turn, induced specific behavioral deficits, making working memory more vulnerable to intervening distractors. In the model, distractibility depends on the similarity between the distractor and the mnemonic representation, and therefore broadening the mnemonic representations increases the range of distractors that can disrupt behavior. The authors tested this model prediction by analyzing behavior from healthy humans administered ketamine, a pharmacological model of schizophrenia, during a spatial delayed match-to-sample task. Matching the model prediction, ketamine increased the rate of errors specifically for distractors that would overlap with a broadened mnemonic representation (Figure 5C). Just as the biophysical basis of the model allows instantiation of potential pathologies, it can also readily incorporate pharmacological treatments to compensate for these deficits. In particular, in this model it was demonstrated as proof-of-principle that glutamatergic or GABAergic manipulations could restore excitation-inhibition balance, reversing the broadened mnemonic representations and corresponding distractibility induced by disinhibition (Figure 5D). An open question is concerned with the brain mechanisms for deciding which information should be considered task-relevant versus distracting and how this may or may not be related to reward value processing of potentially relevant or distracting stimuli. Impairments of this decision process could be relatively independent from those of working memory circuit's ability to resist distractors as described above, which would suggest an orthogonality between these deficits. Future research is needed to assess whether this is indeed the case.

In the model, the network's ability to filter out distractors is impaired by a reduced excitation in inhibitory neurons. The main insight is that predominant behavorial disturbance due to modest disinhibition may not be so much the inability of memory storage per se as the difficulty of ignoring behaviorally irrelevant inputs during memory maintenance. The observation that ketamine in human subjects leads to impaired resistance against near distractors, as predicted by the model, suggests that disinhibition involves NMDARs. Intuitively, this could be caused by a reduced NMDAR mediated excitation in inhibitory neurons. In support of this view, there is evidence that, in rodents, acute ketamine administration led to a decreased activity of putative fast-spiking (FS) interneurons, and increased activity of putative pyramidal cells (Homayoun and Moghaddam 2007). Moreover, since FS inhibitory neurons are critically involved in the generation of fast γ oscillations (Buzsaki and Wang 2012, Wang 2010), a reduced excitation of those neurons could explain abnormal γ synchrony observed in schizophrenic patients (Spencer et al. 2004, Lisman et al. 2008).

However, in fast-spiking interneurons of the mice frontal cortex, NMDAR mediated excitation is small and insensitive to NMDAR blocker AP5 (Rotaru et al. 2011). In adult rats, the majority of fast-spiking interneurons are devoid of NMDA receptors, whereas NMDAR dependent synaptic excitation is more significant in other subclasses of regular-spiking and low-threshold spiking inhibitory cells (Wang and Gao 2009). The latter mediate dendritic inhibition, thereby gating synaptic inputs onto pyramidal cells. Further, the dendrite-targeting interneurons function in an input-specific manner, enabling pyramidal neurons to be selectively activated by task-relevant inputs. This has been incorporated in an extended working memory microcircuit model endowed with three subtypes of inhibitory neurons: (a) PV-expressing soma-targeting interneurons that control pyramidal firing output, (b) interneurons that express calbindin or somatostatin and gate dendritic inputs to pyramidal cells, (c) interneurons that express calrintinin or VIP and preferentially target dendrite-targeting interneurons (thereby providing a new disinhibition mechanism) (Wang et al. 2004, Wang 2013). It was found that dendritic inhibition controls the network's ability to resist irrelevant distractors more effectively than peri-somatic inhibition that controls the spiking output of pyramidal neurons. Taken together, one plausible scenario consistent with currently available evidence is that disinhibition induced by ketamine results from a reduction of NMDAR dependent excitation of dendrite-targeting interneurons. This prediction can be tested using cell-type specific genetic tools (Kepecs and Fishell 2014, Higley 2014) in future animal experiments.

What happens when the excitation-inhibition balance is tilted in a way that synaptic excitation becomes excessively strong? Model simulations showed that one consequence of such an imbalance could lead to behavioral inflexibility: attractor states encoding memory items become so robust that it becomes difficult to switch off from one memory attractor state either to rest (memory erasure) or another memory state (Rolls et al. 2008, Durstewitz and Seamans 2008, Gruber et al. 2010). This idea is interesting especially in the light of the fact that working memory is not limited to sensory stimuli but also more abstract information such as behavioral task sets or rules (Miller and Cohen 2001, Wallis et al. 2001, Sakai 2008, Buckley et al. 2009, Lapich et al. 2010, Sigala et al. 2010), and attractor network models have been extended to internal representation of behavioral rule or context in flexible behavior (Rigotti et al. 2010, 2013). Thus, behavioral inflexibility may be reflected in the difficulty to make a transition from a behavioral context to another one, which is a hallmark of abnormal cognition in schizophrenia.

This framework is also useful for analyzing abnormal neuromodulation in mental illness. The dopamine system represents an example *par excellence*. It is well known that working memory performance exhibits an inverted U-shaped dependence on dopamine modulation: too little dopamine, you loss working memory; too much dopamine, you are inflexible with switching on and off in a working memory system. Dopamine modulation acts on targets such as NMDA receptor mediated excitatory synaptic excitation and GABA mediated inhibitory synaptic inhibition (Brunel and Wang 2001, Seamans et al. 2001, Durstewitz et al. 2000), or the gain of single-neuron input-output relationship (Cohen and Servant-Schreiber 1992). Computational modeling showed that an inverted-U shape of dopamine modulation can be readily explained if dopamine modulation has a differential sensitivity to the NMDA conductance and GABA conductance (Brunel and Wang 2001). Furthermore, interestingly, the network’s ability to ignore distractors is sensitive to modulation by dopamine of recurrent excitation and inhibition. Therefore, even a mild impairment of dopaminergic signaling in the prefrontal cortex could be very detrimental to robust working memory maintenance in spite of ongoing sensory flow.

These studies on working memory demonstrate how biophysically-based modeling in interplay with experimentation can play a powerful role in making discoveries and producing new hypotheses about the brain mechanisms of core cognitive processes implicated in psychiatric disorders.

Looking forward: building a new cross-disciplinary field

The economic cost of mental illness represents an enormous burden on the society Wittchen et al. (2011, Olesen et al. (2012, Vos et al. (2012). The critical nature of our knowledge gap for the clinical neuroscience fields, including neurology, neurosurgery, psychiatry, and psychology, is well known. In the United States, National Institute of Health initiatives including the Human Connectome Project (http://www.humanconnectomeproject.org) and the BRAIN Initiative (http://www.nih.gov/science/brain/) are designed to advance current approaches and to develop new technologies to characterize brain circuit function. Parallel initiatives are underway in Europe and Asia.

In this *Perspective*, we marshaled findings from recent work on reinforcement learning and working memory to argue for a Computational Psychiatry approach to brain disorders. This perspective emphasizes an integration of experimentation, data analysis and theory in concerted efforts to understand neural circuits involved in mental illness. Although we have focused on local circuit mechanisms, computational psychiatry must also be developed for large brain systems. A notable line of research in this regard is concerned with the interplay between cortex and basal ganglia, which is important for both working memory and decision-making (O’Reilly and Frank 2006, Lo and Wang 2006, Ding and Gold 2013). In fact, behavioral evidence from a cleverly designed experiment suggests that impaired RL in schizophrenia is attributable, largely, to working memory deficits rather than valuation process (Michael Frank, personal communication). Another interplay involves cortex and thamalus (Vukadinovic 2011, Anticevic et al. 2013). More broadly, new approaches applied to the study of the connectivity properties of large-scale brain systems are exciting developments (Sporns 2009, Bullmore and Sporns 2009, Markov et al. 2013) with important implications for psychiatric disorders (Anticevic et al. 2013, Rubinov and Bullmore 2013, Yang et al. 2014).

Unprecedented ongoing progress in neuroscience offers extraordinary opportunities as well as challenges. First, progress in genomics, massive neuroimaging and other advances are creating enormous datasets that, in turn, require new mathematical/statistical tools. Second, there is an increasing recognition that, so far, mechanistic preclinical studies have been almost exclusively focused on local circuits, but we need to develop large-scale brain circuit models in order to investigate how the PFC controls and interacts with many other brain regions in a highly interconnected large system. Third, major mental disorders like schizophrenia, autism qnd ADHD are neurodevelopmental diseases (Moore et al. 2006, Belujon and Grace 2008, Insel 2010, Fair et al 2012). Thus it is critical that computational models incorporate developmental changes in synaptic and circuit function in disease-related models. For instance, the human neural representation of working memory assessed with fMRI changes during adolescence (Satterthwaite et al. 2013). Similarly, synaptic mechanisms evolve during adolescence. In rodents, for example, NMDA receptors are abundant on PV interneurons early in life, but they are present more sparsely in adults (Belforte et al. 2010). In these circuits, reducing NMDA receptor expression early in life, but not in adulthood impairs cognitive function in adulthood. There is a dearth of computational modeling dedicated to understanding critical periods in neurodevelopment and the impact of even "transient" developmental disruption on circuit development and cognitive function in adulthood. Progress along these lines will require sophisticated neural circuit modeling in conjunction with genetic, physiological and imaging experimentation. Fourth and finally, can one quantitatively capture specific features of the normal and dysfunctional flow of thought associated with mental illness? A recent work took the view that language could be used "as a privileged measuring lens into thought", and showed that quantitative analysis of speech could yield accurate sorting of schizophrenia versus mania with high sensitivity and specificity (Mota et al. 2012). Language is a human cognitive ability implicated in mental disorders, thus elucidation of brain's language circuit represents another neuroscientific theme relevant to Psychiatry.

It is our belief that these challenges cannot be overcome without theory and computational modeling. To advance the field, we need new infrastructure, resources and training of cross-disciplinary young talents who are well versed both in mathematical modeling and experimentation. In particular, it would be important to develop training programs whereby graduate students and postdoctoral fellows trained in the physical and mathematical sciences could more easily be introduced psychiatry develop without the input of physician-scientists. Third, government funding agencies and non-profit organizations and foundations should offer new programs to promote highly cross-disciplinary education and research in computational psychiatry. Through these concerted efforts, we are optimistic that computational psychiatry could play an indispensable role in addressing the great challenges of mental health in the twenty-first century.

Acknowledgments

We thank John Murray, Alan Anticevic, and Philip Corlett for discussions related to computational psychiatry and John Murry for making Figure 5; Michael Frank, Nathaniel Daw, Ray Dolan and Klaas Enno Stephan for their comments on an earlier version of the manuscript. This work was supported by NIH grant R01MH062349 and ONR grant N00014-13-1-0297 (X.-J.W.); National Institute on Alcohol Abuse and Alcoholism grant (2P50AA012870-11 to PI: J.H.K.), National Center for Advancing Translational Science grants (UL1RR024139, PI: R. Sherwin, UH2TR000960, PI: JHK), and the U.S. Department of Veterans Affairs through their support of the VA National Center for PTSD and the Consortium to Alleviate PTSD (J.H.K).

References

Abbott LF. Theoretical neuroscience rising. Neuron. 2008; 60:489–495. [PubMed: 18995824]

Anticevic A, Cole MW, Repovs G, Murray JD, Brumbaugh MS, Winkler AM, Savic A, Krystal JH, Pearlson GD, Glahn DC. Characterizing Thalamo-Cortical Disturbances in Schizophrenia and Bipolar Illness. Cereb. Cortex. 2013

Anticevic A, Cole MW, Repovs G, Savic A, Driesen NR, Yang G, Cho YT, Murray JD, Glahn DC, Wang XJ, Krystal JH. Connectivity, pharmacology, and computation: toward a mechanistic understanding of neural system dysfunction in schizophrenia. Front Psychiatry. 2013; 4:169. [PubMed: 24399974]

Arnsten AF, Paspalas CD, Gamo NJ, Yang Y, Wang M. Dynamic Network Connectivity: A new form of neuroplasticity. Trends Cogn. Sci. (Regul. Ed.). 2010; 14:365–375. [PubMed: 20554470]

Baddeley A. Working memory: theories, models, and controversies. Annu Rev Psychol. 2012; 63:1–29. [PubMed: 21961947]

Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology. 1998; 37:407–419. [PubMed: 9704982]

Barch DM, Ceaser A. Cognition in schizophrenia: core psychological and neural mechanisms. Trends Cogn. Sci. (Regul. Ed.). 2012; 16:27–34. [PubMed: 22169777]

Belforte JE, Zsiros V, Sklar ER, Jiang Z, Yu G, Li Y, Quinlan EM, Nakazawa K. Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. Nat. Neurosci. 2010; 13:76–83. [PubMed: 19915563]

Belujon P, Grace AA. Critical role of the prefrontal cortex in the regulation of hippocampus-accumbens information flow. J. Neurosci. 2008; 28:9797–9805. [PubMed: 18815264]

Botvinick MM, Niv Y, Barto AC. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. Cognition. 2009; 113:262–280. [PubMed: 18926527]

Brandon NJ, Millar JK, Korth C, Sive H, Singh KK, Sawa A. Understanding the role of DISC1 in psychiatric disease and during normal development. J. Neurosci. 2009; 29:12768–12775. [PubMed: 19828788]

Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, Stephan KE. Dissecting psychiatric spectrum disorders by generative embedding. Neuroimage Clin. 2014; 4:98–111. [PubMed: 24363992]

Brunel N, Wang X-J. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. J Comput Neurosci. 2001; 11:63–85. [PubMed: 11524578]

Buchanan RW, Keefe RS, Lieberman JA, Barch DM, Csernansky JG, Goff DC, Gold JM, Green MF, Jarskog LF, Javitt DC, et al. A randomized clinical trial of MK-0777 for the treatment of cognitive impairments in people with schizophrenia. Biol Psychiatry. 2011; 69:442–449. [PubMed: 21145041]

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, Tanaka K. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science. 2009; 325:52–58. [PubMed: 19574382]

Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 2009; 10:186–198. [PubMed: 19190637]

Buzsaki G, Wang X-J. Mechanisms of gamma oscillations. Annu. Rev. Neurosci. 2012; 35:203–225. [PubMed: 22443509]

Carandini M. From circuits to behavior: a bridge too far? Nat. Neurosci. 2012; 15:507–509. [PubMed: 22449960]

Carter E, Wang X-J. Cannabinoid-mediated disinhibition and working memory: dynamical interplay of multiple feedback mechanisms in a continuous attractor model of prefrontal cortex. Cereb. Cortex. 2007; 17(Suppl 1):16–26.

Carter CS, Barch DM, Buchanan RW, Bullmore E, Krystal JH, Cohen J, Geyer M, Green M, Nuechterlein KH, Robbins T, Silverstein S, Smith EE, Strauss M, Wykes T, Heinssen R. Identifying cognitive mechanisms targeted for treatment development in schizophrenia: an overview of the first meeting of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Initiative. Biol. Psychiatry. 2008; 64:4–10. [PubMed: 18466880]

Compte A, Brunel N, Goldman-Rakic PS, Wang X-J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb. Cortex. 2000; 10:910–923. [PubMed: 10982751]

Constantinidis C, Wang X-J. A neural circuit basis for spatial working memory. Neuroscientist. 2004; 10:553–565. [PubMed: 15534040]

Corlett PR, Murray GK, Honey GD, Aitken MR, Shanks DR, Robbins TW, Bullmore ET, Dickinson A, Fletcher PC. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain. 2007; 130:2387–2400. [PubMed: 17690132]

Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH. Toward a neurobiology of delusions. Prog. Neurobiol. 2010; 92:345–369. [PubMed: 20558235]

Courchesne E, Mouton PR, Calhoun ME, Semendeferi K, Ahrens-Barbeau C, Hallet MJ, Barnes CC, Pierce K. Neuron number and size in prefrontal cortex of children with autism. JAMA. 2011; 306:2001–2010. [PubMed: 22068992]

Coyle JT. Glutamate and schizophrenia: beyond the dopamine hypothesis. Cell. Mol. Neurobiol. 2006; 26:365–384. [PubMed: 16773445]

Coyle JT, Tsai G, Goff D. Converging evidence of NMDA receptor hypofunction in the pathophysiology of schizophrenia. Ann N Y Acad Sci. 2003; 1003:318–327. [PubMed: 14684455]

Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 2005; 8:1704–1711. [PubMed: 16286932]

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans’ choices and striatal prediction errors. Neuron. 2011; 69(6):1204–1215. [PubMed: 21435563]

Dayan, P.; Abbott, LF. Theoretical Neuroscience. Cambridge MA: MIT Press; 2001.

Deco G, Rolls ET, Romo R. Stochastic dynamics as a principle of brain function. Prog. Neurobiol. 2009; 88:1–16. [PubMed: 19428958]

Deco G, Scarano L, Soto-Faraco S. Weber's law in decision making: integrating behavioral data in humans with a neurophysiological model. J. Neurosci. 2007; 27:11192–11200. [PubMed: 17942714]

D’Esposito M. From cognitive to neural models of working memory. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 2007; 362:761–772. [PubMed: 17400538]

Dezfouli A, Piray P, Keramati MM, Ekhtiari H, Lucas C, Mokri A. A neurocomputational model for cocaine addiction. Neural Comput. 2009; 21:2869–2893. [PubMed: 19635010]

Ding L, Gold JI. The basal ganglia's contributions to perceptual decision making. Neuron. 2013; 79:640–649. [PubMed: 23972593]

Dolan RJ, Dayan P. Goals and habits in the brain. Neuron. 2013; 80:312–325. [PubMed: 24139036]

Douglas RJ, Martin KAC. Neuronal circuits of the neocortex. Annu Rev Neurosci. 2004; 27:419–451. [PubMed: 15217339]

Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH. Recurrent excitation in neocortical circuits. Science. 1995; 269:981–985. [PubMed: 7638624]

Durstewitz D, Seamans JK. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol. Psychiatry. 2008; 64:739–749. [PubMed: 18620336]

Durstewitz D, Seamans JK, Sejnowski TJ. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. J. Neurophysiol. 2000; 83:1733–1750. [PubMed: 10712493]

Engel TA, Wang X-J. Same or different? A neural circuit mechanism of similarity-based pattern match decision making. J. Neurosci. 2011; 31:6982–6996. [PubMed: 21562260]

Fair DA, Bathula D, Nikolas MA, Nigg JT. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. Proc. Natl. Acad. Sci. U.S.A. 2012; 109:6769–6774. [PubMed: 22474392]

Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, Gabbard GO, Gau SS, Javitt DC, Oquendo MA, Shrout PE, Vieta E, Yager J. The initial field trials of DSM-5: new blooms and old thorns. Am J Psychiatry. 2013; 170:1–5. [PubMed: 23288382]

Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003; 19:1273–1302. [PubMed: 12948688]

Friston KJ, Stephan KE, Montague R, Dolan RJ. Computational psychiatry: the brain as a phantastic organ. Lancet Psychiatry. 2014; 1:148–158.

Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey’s dorsolateral prefrontal cortex. J. Neurophysiol. 1989; 61:331–349. [PubMed: 2918358]

Furman M, Wang X-J. Similarity effect and optimal control of multiple-choice decision making. Neuron. 2008; 60:1153–1168. [PubMed: 19109918]

Fuster, JM. The Prefrontal Cortex. fourth ed.. New York: Academic Press; 2008.

Glascher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 2010; 66:585–595. [PubMed: 20510862]

Glimcher, PW. Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics. Cambridge MA: MIT Press; 2003.

Glimcher, PW.; Fehr, CF. Neuroeconomics: Decision Making and the Brain. 2nd ed.. London: Academic Press; 2013.

Goff DC. Bitopertin: the good news and bad news. JAMA Psychiatry. 2014; 71:621–622. [PubMed: 24696065]

Gold JI, Shadlen MN. The neural basis of decision making. Annu. Rev. Neurosci. 2007; 30:535–574. [PubMed: 17600525]

Goldman-Rakic, PS. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: Plum, F.; Mountcastle, V., editors. Handbook of Physiology – The nervous system V. Vol. Chapter 9. Bethesda, Maryland: American Physiological Society; 1987. p. 373-417.

Goldman-Rakic PS. Working memory dysfunction in schizophrenia. J Neuropsychiatry Clin. Neurosci. 1994; 6:348–357. [PubMed: 7841806]

Goldman-Rakic PS. Cellular basis of working memory. Neuron. 1995; 14:477–485. [PubMed: 7695894]

Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry. 2003; 160:636–645. [PubMed: 12668349]

Gruber AJ, Calhoon GG, Shusterman I, Schoenbaum G, Roesch MR, O'Donnell P. More is less: a disinhibited prefrontal cortex impairs cognitive flexibility. J. Neurosci. 2010; 30:17102–17110. [PubMed: 21159980]

Hansel D, Mato G. Short-term plasticity explains irregular persistent activity in working memory tasks. J. Neurosci. 2013; 33:133–149. [PubMed: 23283328]

Higley MJ. Localized GABAergic inhibition of dendritic Ca(2+) signalling. Nat. Rev. Neurosci. 2014; 15:567–572. [PubMed: 25116141]

Homayoun H, Moghaddam B. NMDA receptor hypofunction produces opposite effects on prefrontal cortex interneurons and pyramidal neurons. J. Neurosci. 2007; 27:11496–11500. [PubMed: 17959792]

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 2012; 15:470–476. [PubMed: 22231429]

Huys QJ, Pizzagalli DA, Bogdan R, Dayan P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. Biol Mood Anxiety Disord. 2013; 3:12. [PubMed: 23782813]

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010; 167:748–751. [PubMed: 20595427]

Insel TR. Rethinking schizophrenia. Nature. 2010; 468:187–193. [PubMed: 21068826]

Insel TR. The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. Am. J. Psychiatry. 2014; 171:395–397. [PubMed: 24687194]

Johansen JP, Cain CK, Ostroff LE, LeDoux JE. Molecular mechanisms of fear learning and memory. Cell. 2011; 147:509–524. [PubMed: 22036561]

Josselyn SA. Continuing the search for the engram: examining the mechanism of fear memories. J. Psychiatry Neurosci. 2010; 35:221–228. [PubMed: 20569648]

Kahneman, D. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux; 2011.

Kepecs A, Fishell G. Interneuron cell types are fit to function. Nature. 2014; 505:318–326. [PubMed: 24429630]

Kilpatrick ZP, Ermentrout B, Doiron B. Optimizing working memory with heterogeneity of recurrent cortical excitation. J. Neurosci. 2013; 33:18999–19011. [PubMed: 24285904]

Krueger RF. The structure of common mental disorders. Arch. Gen. Psychiatry. 1999; 56:921–926. [PubMed: 10530634]

Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, Heninger GR, Bowers MB, Charney DS. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch. Gen. Psychiatry. 1994; 51:199–214. [PubMed: 8122957]

Krystal JH, State MW. Psychiatric disorders: diagnosis to therapy. Cell. 2014; 157:201–214. [PubMed: 24679536]

Kurth-Nelson, Z.; Redish, AD. Modeling decision-making systems in addiction. In: Gutkin, B.; Ahmed, SH., editors. Computational Neuroscience of Drug Addiction. Springer Publishing; 2011. p. 163-188.

Lapish CC, Durstewitz D, Chandler LJ, Seamans JK. Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:11963–11968. [PubMed: 18708525]

Lee D. Decision making: from neuroscience to psychiatry. Neuron. 2013; 78:233–248. [PubMed: 23622061]

Lee J, Park S. Working memory impairments in schizophrenia: a meta-analysis. J Abnorm Psychol. 2005; 114:599–611. [PubMed: 16351383]

Lewis D, Hashimoto T, Volk D. Cortical inhibitory neurons and schizophrenia. Nat. Rev. Neurosci. 2005; 6:312–324. [PubMed: 15803162]

Lewis DA, Curley AA, Glausier JR, Volk DW. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. Trends Neurosci. 2012; 35:57–67. [PubMed: 22154068]

Lewis DA, Gonzalez-Burgos G. Pathophysiologically based treatment interventions in schizophrenia. Nat. Med. 2006; 12:1016–1022. [PubMed: 16960576]

Lisman J, Coyle J, Green R, Javitt D, Benes F, Heckers S, Grace A. Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. Trends in Neurosci. 2008; 31:234–242.

Lo CC, Wang X-J. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat. Neurosci. 2006; 9:956–963. [PubMed: 16767089]

Lo CC, Boucher L, Paré M, Schall JD, Wang X-J. Proactive inhibitory control and attractor dynamics in countermanding action: a spiking neural circuit model. J. Neurosci. 2009; 29:9059–9071. [PubMed: 19605643]

Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G. The impact of orbitofrontal dysfunction on cocaine addiction. Nat. Neurosci. 2012; 15:358–366. [PubMed: 22267164]

Luck SJ, Gold JM. The construct of attention in schizophrenia. Biol. Psychiatry. 2008; 64:34–39. [PubMed: 18374901]

Machens CK, Romo R, Brody CD. Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science. 2005; 18:1121–1124. [PubMed: 15718474]

Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. 2011; 14:154–162. [PubMed: 21270784]

Markov NT, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H. Cortical high-density counterstream architectures. Science. 2013; 342:1238406. [PubMed: 24179228]

Mesulam, M-M. Principles of Behavioral and Cognitive Neurology. 2nd ed.. New York: Oxford University Press; 2000.

Millan MJ, Agid Y, Brune M, Bullmore ET, Carter CS, Clayton NS, Connor R, Davis S, Deakin B, DeRubeis, B RJ, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat Rev Drug Discov. 2012; 11:141–168. [PubMed: 22293568]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annu Rev Neurosci. 2001; 24:167–202. [PubMed: 11283309]

Miller P, Wang X-J. Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. Proc Natl Acad Sci U S A. 2006; 103:201–206. [PubMed: 16371469]

Moghaddam B, Krystal JH. Capturing the angel in ”angel dust”: twenty years of translational neuroscience studies of NMDA receptor antagonists in animals and humans. Schizophr Bull. 2012; 38:942–949. [PubMed: 22899397]

Montague P, Dayan P, Sejnowski T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. 1996; 16:1936–1947. [PubMed: 8774460]

Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. Trends Cogn. Sci. (Regul. Ed.). 2012; 16:72–80. [PubMed: 22177032]

Moore H, Jentsch JD, Ghajarnia M, Geyer MA, Grace AA. A neurobehavioral systems analysis of adult rats exposed to methylazoxymethanol acetate on E17: implications for the neuropathology of schizophrenia. Biol. Psychiatry. 2006; 60:253–264. [PubMed: 16581031]

Mota NB, Vasconcelos NA, Lemos N, Pieretti AC, Kinouchi O, Cecchi GA, Copelli M, Ribeiro S. Speech graphs provide a quantitative measure of thought disorder in psychosis. PLoS ONE. 2012; 7(4):e34928. [PubMed: 22506057]

Murray JD, Anticevic A, Gancsos M, Ichinose M, Corlett PR, Krystal JH, Wang XJ. Linking Microcircuit Dysfunction to Cognitive Impairment: Effects of Disinhibition Associated with Schizophrenia in a Cortical Working Memory Model. Cereb. Cortex. 2014; 24:859–872. [PubMed: 23203979]

100. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jonsson B, Jordanova A, Musayev A, Gustavsson A, Gabilondo A, Maercker, B A, et al. The economic cost of brain disorders in Europe. Eur. J. Neurol. 2012; 19:155–162. [PubMed: 22175760]

101. Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychol. Rev. 1992; 99:45–77. [PubMed: 1546118]

102. O’Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 2006; 18:283–328. [PubMed: 16378516]

103. Panlilio LV, Thorndike EB, Schindler CW. Blocking of conditioning to a cocaine-paired stimulus: testing the hypothesis that cocaine perpetually produces a signal of larger-than-expected reward. Pharmacol. Biochem. Behav. 2007; 86:774–777. [PubMed: 17445874]

104. Park S, Holzman PS. Schizophrenics show spatial working memory deficits. Arch. Gen. Psychiatry. 1992; 49:975–982. [PubMed: 1449384]

105. Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. 1998; 21:227–277. [PubMed: 9530497]

106. Pereira J, Wang X-J. A trade-off between accuracy and flexibility in a working memory circuit endowed with slow feedback mechanisms. Cerebral Cortex. 2014 in press.

107. Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 2008; 9:545–556. [PubMed: 18545266]

108. Ratcliff R. A theory of memory retrieval. Psychol. Rev. 1978; 85:59–108.

109. Redish AD. Addiction as a computational process gone awry. Science. 2004; 306:1944–1947. [PubMed: 15591205]

110. Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. Psychol Rev. 2007; 114:784–805. [PubMed: 17638506]

111. Renart A, Song P, Wang X-J. Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. Neuron. 2003; 38:473–485. [PubMed: 12741993]

112. Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Black, AH.; Prokasy, WF., editors. Classical Conditioning II. New York: Appleton-Century-Crofts; 1972. p. 64-69.

113. Rigotti M, Ben Dayan Rubin DD, Wang X-J, Fusi S. Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. Front. Comput. Neurosci. 2010; 4:24. [PubMed: 21048899]

114. Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S. The importance of mixed selectivity in complex cognitive tasks. Nature. 2013; 497:585–590. [PubMed: 23685452]

115. Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. Trends Cogn. Sci. (Regul. Ed.). 2012; 16:81–91. [PubMed: 22155014]

116. Roitman JD, Shadlen MN. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci. 2002; 22:9475–9489. [PubMed: 12417672]

117. Rolls E, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nat. Rev. Neurosci. 2008; 9:696–709. [PubMed: 18714326]

118. Rotaru DC, Yoshino H, Lewis DA, Ermentrout GB, Gonzalez-Burgos G. Glutamate receptor subtypes mediating synaptic activation of prefrontal cortex neurons: relevance for schizophrenia. J. Neurosci. 2011; 31:142–156. [PubMed: 21209199]

119. Rubinov M, Bullmore E. Fledgling pathoconnectomics of psychiatric disorders. Trends Cogn. Sci. (Regul. Ed.). 2013; 17:641–647. [PubMed: 24238779]

120. Sakai K. Task set and prefrontal cortex. Annu. Rev. Neurosci. 2008; 31:219–245. [PubMed: 18558854]

121. Satterthwaite TD, Wolf DH, Erus G, Ruparel K, Elliott MA, Gennatas ED, Hopson R, Jackson C, Prabhakaran K, Bilker WB, et al. Functional maturation of the executive system during adolescence. J. Neurosci. 2013; 33(41):16249–16261. [PubMed: 24107956]

122. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]

123. Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ. Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. Proc. Natl. Acad. Sci. U.S.A. 2001; 98:301–306. [PubMed: 11134516]

124. Sejnowski TJ, Koch C, Churchland PS. Computational neuroscience. Science. 1988; 241:1299–1306. [PubMed: 3045969]

125. Sigala N, Kusunoki M, Nimmo-Smith I, Gaffan D, Duncan J. Hierarchical coding for sequential task events in the monkey prefrontal cortex. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:11969–119674. [PubMed: 18689686]

126. Simon, DA.; Daw, ND. dual-system learning models and drugs of abuse. In: Gutkin, B.; Ahmed, SH., editors. Computational Neuroscience of Drug Addiction. Springer Publishing; 2011. p. 145-161.

127. Skoblenick K, Everling S. NMDA antagonist ketamine reduces task selectivity in macaque dorsolateral prefrontal neurons and impairs performance of randomly interleaved prosaccades and antisaccades. J. Neurosci. 2012; 32:12018–12027. [PubMed: 22933786]

128. Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. Trends Neurosci. 2004; 27:161–168. [PubMed: 15036882]

129. Soltani A, Wang X-J. A biophysically based neural model of matching law behavior: melioration by stochastic synapses. J. Neurosci. 2006; 26:3731–3744. [PubMed: 16597727]

130. Spencer K, Nestor P, Perlmutter R, Niznikiewicz M, Klump M, Frumin M, Shenton M, McCarley R. Neural synchrony indexes disordered perception and cognition in schizophrenia. Proc. Natl. Acad. Sci. U.S.A. 2004; 101:17288–17293. [PubMed: 15546988]

131. Sporns, O. Networks of the Brain. Cambridge, MA: MIT Press; 2009.

132. Stephan KE, Harrison LM, Kiebel SJ, David O, Penny WD, Friston KJ. Dynamic causal models of neural system dynamics:current state and future extensions. J. Biosci. 2007; 32:129–144. [PubMed: 17426386]

133. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM. Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. Biol. Psychiatry. 2011; 69:424–431. [PubMed: 21168124]

134. Sutton, RS.; Barto, AG. Reinforcement Learning : an Introduction. Cambridge, MA: MIT Press; 1998.

135. Szczepanski SM, Knight RT. Insights into human behavior from lesions to the prefrontal cortex. Neuron. 2014; 83:1002–1018. [PubMed: 25175878]

136. Voon V, Derbyshire K, Ruck C, Irvine MA, Worbe Y, Enander J, Schreiber LR, Gillan C, Fineberg NA, Sahakian BJ, Robbins TW, Harrison NA, Wood J, Daw ND, Dayan P, Grant JE,

Bullmore ET. Disorders of compulsivity: a common bias towards learning habits. Mol. Psychiatry. 2014 in press.

137. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans, J V, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380:2163–2196. [PubMed: 23245607]

138. Vukadinovic Z. Sleep abnormalities in schizophrenia may suggest impaired trans-thalamic cortico-cortical communication: towards a dynamic model of the illness. Eur. J. Neurosci. 2011; 34:1031–1039. [PubMed: 21895800]

139. Wallis J, Anderson K, Miller E. Single neurons in prefrontal cortex encode abstract rules. Nature. 2001; 411:953–956. [PubMed: 11418860]

140. Wang H, Stradtman GG, Wang X-J, Gao WJ. A specialized NMDA receptor function in layer 5 recurrent microcircuitry of the adult rat prefrontal cortex. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:16791–16796. [PubMed: 18922773]

141. Wang HX, Gao WJ. Cell type-specific development of NMDA receptors in the interneurons of rat prefrontal cortex. Neuropsychopharmacology. 2009; 34:2028–2040. [PubMed: 19242405]

142. Wang M, Yang Y, Wang CJ, Gamo NJ, Jin LE, Mazer JA, Morrison JH, Wang X-J, Arnsten AF. NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. Neuron. 2013; 77:736–749. [PubMed: 23439125]

143. Wang X-J. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. J. Neurosci. 1999; 19:9587–9603. [PubMed: 10531461]

144. Wang X-J. Synaptic reverberation underlying mnemonic persistent activity. Trends in Neurosci. 2001; 24:455–463.

145. Wang X-J. Probabilistic decision making by slow reverberation in cortical circuits. Neuron. 2002; 36:955–968. [PubMed: 12467598]

146. Wang X-J. Decision making in recurrent neuronal circuits. Neuron. 2008; 60:215–234. [PubMed: 18957215]

147. Wang X-J. Neurophysiological and computational principles of cortical rhythms in cognition. Physiol. Rev. 2010; 90:1195–1268. [PubMed: 20664082]

148. Wang, X-J. The prefrontal cortex as a quintessential ‘cognitive-type’ neural circuit: Working memory and decision making. In: Stuss, DT.; Knight, RT., editors. Principles of Frontal Lobe Function. Second ed.. New York: Cambridge University Press; 2013. p. 226-248.

149. Wang X-J, Tegnér J, Constantinidis C, Goldman-Rakic PS. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. Proc Natl Acad Sci U S A. 2004; 101:1368–1373. [PubMed: 14742867]

150. Wei Z, Wang X-J, Wang DH. From distributed resources to limited slots in multiple-item working memory: a spiking network model with normalization. J. Neurosci. 2012; 32:11228–11240. [PubMed: 22895707]

151. Wiecki TV, Poland JS, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. Clinical Psychological Science. 2014 in press.

152. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jonsson B, Olesen J, Allgulander C, Alonso J, Faravelli C, Fratiglioni L, Jennum P, Lieb R, Maercker A, van Os J, Preisig M, Salvador-Carulla L, Simon R, Steinhausen HC. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011; 21:655–679. [PubMed: 21896369]

153. Wong KF, Wang X-J. A recurrent network mechanism of time integration in perceptual decisions. J. Neurosci. 2006; 26:1314–1328. [PubMed: 16436619]

154. Yang GJ, Murray JD, Repovs G, Cole MW, Savic A, Glasser MF, Pittenger C, Krystal JH, Wang XJ, Pearlson GD, Glahn DC, Anticevic A. Altered global brain signal in schizophrenia. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:7438–7443. [PubMed: 24799682]

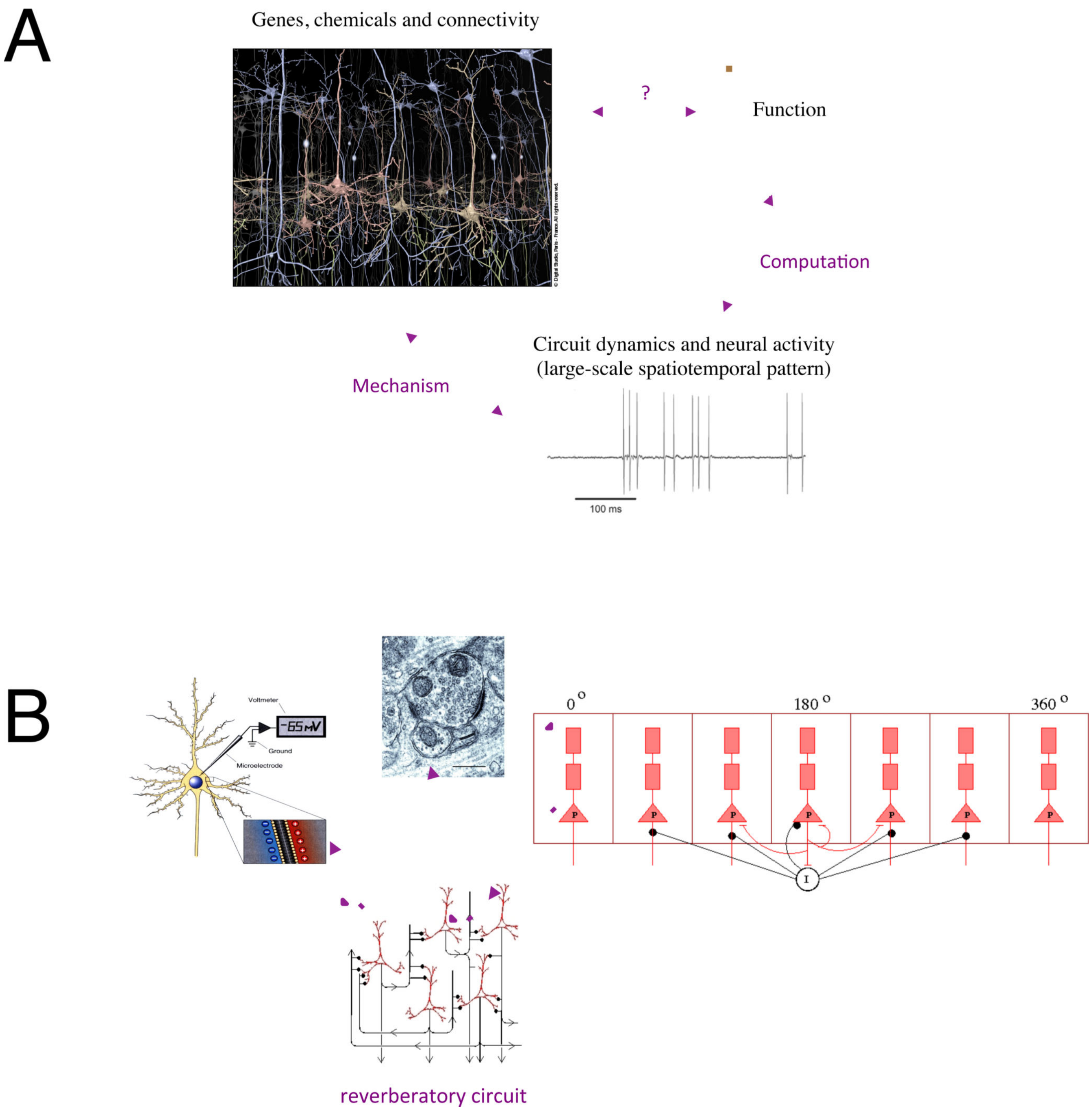
A
Genes, chemicals and connectivity
© Digital Studio, Paris - France. All rights reserved.
?
Function
Computation
Mechanism
Circuit dynamics and neural activity
(large-scale spatiotemporal pattern)
100 ms
B
Voltmeter
-65mV
Ground
Microelectrode
0 o
180 o
360 o
P
I
reverberatory circuit

Figure 1.

(A) Mechanistic understanding of brain functions must relate structure (molecules, cells and network connectivity) and dynamics with behavior. Brain measures probe spatiotemporal neural activity patterns that are correlated with specific aspects of behavior. Theory and modeling provide a powerful tool to elucidate how such a pattern is produced by its biological substrate, on one hand, and give rise to computations necessary to account for brain function, on the other hand. (B) Biologically-based neural circuit modeling is calibrated by physiology of single neurons and synapses, and constrained by quantitative network connectivity data. This approach is arguably necessary for the 3-way understanding between function, neural dynamics and computation, biological mechanism.

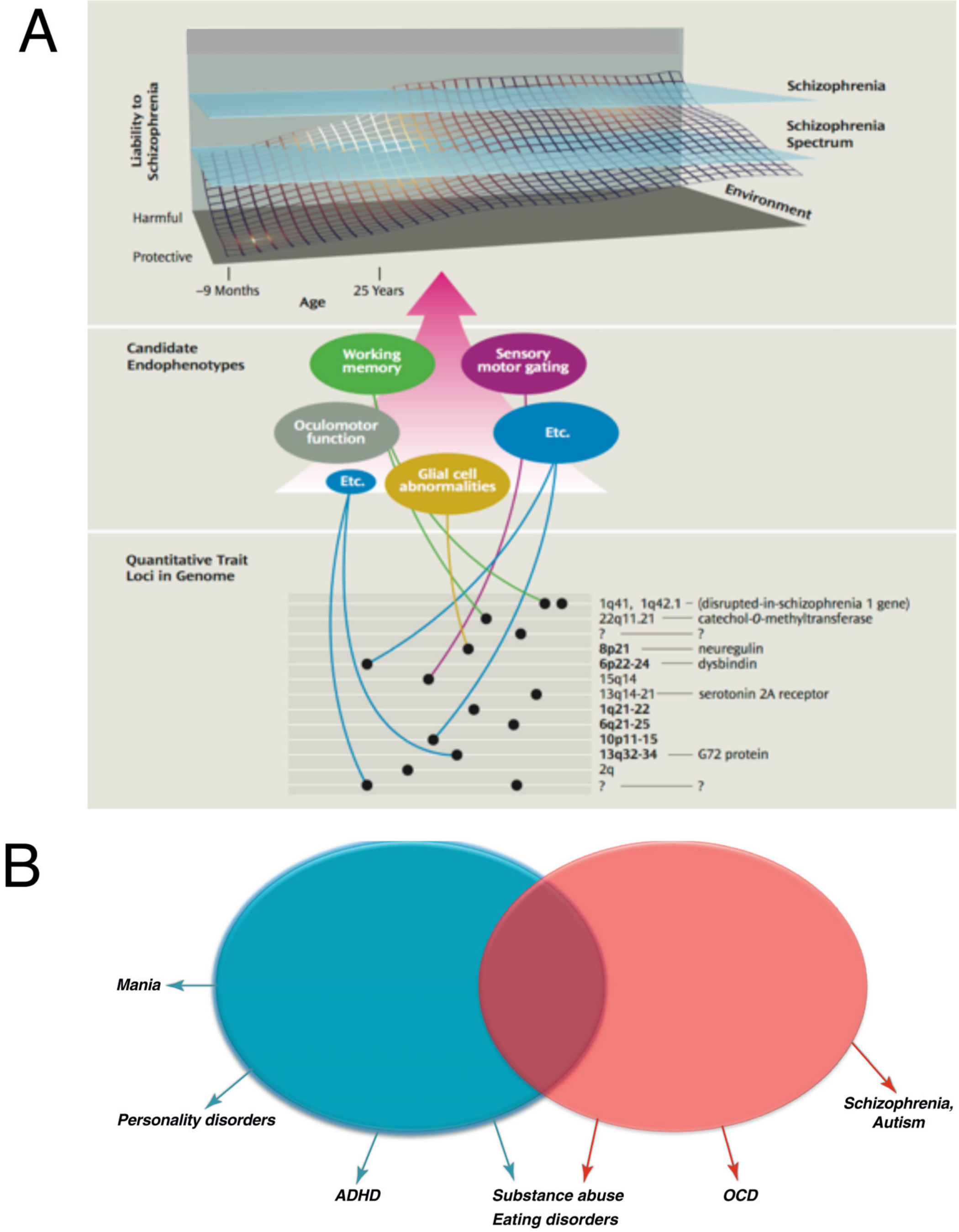

Figure 2.

(A) Gene regions, genes, and putative endophenotypes implicated in a biological systems approach to schizophrenia research. The dynamic developmental interplay among genetic, environmental, and epigenetic factors that produce cumulative liability to developing schizophrenia. Endophenotypes as schizophrenia discriminators involve sensory motor gating, oculomotor function, working memory, and glial cell abnormalities. Many more gene loci, genes, and candidate endophenotypes remain to be discovered (represented by question marks). The figure is not to scale. (B) The impulsivity and compulsivity constructs.

The diagram describes possible psychological component mechanisms underlying the two constructs. It would appear that these different measures likely do not inter-correlate well, which would argue against a unitary construct for either impulsivity or compulsivity, but this issue is still actively being researched. Both impulsivity and compulsivity involve motor/response disinhibition, but at different stages of the response process. (A) was reproduced from Gottesman and Gould (2003), (B) from Robbins et al. (2012), with permission.

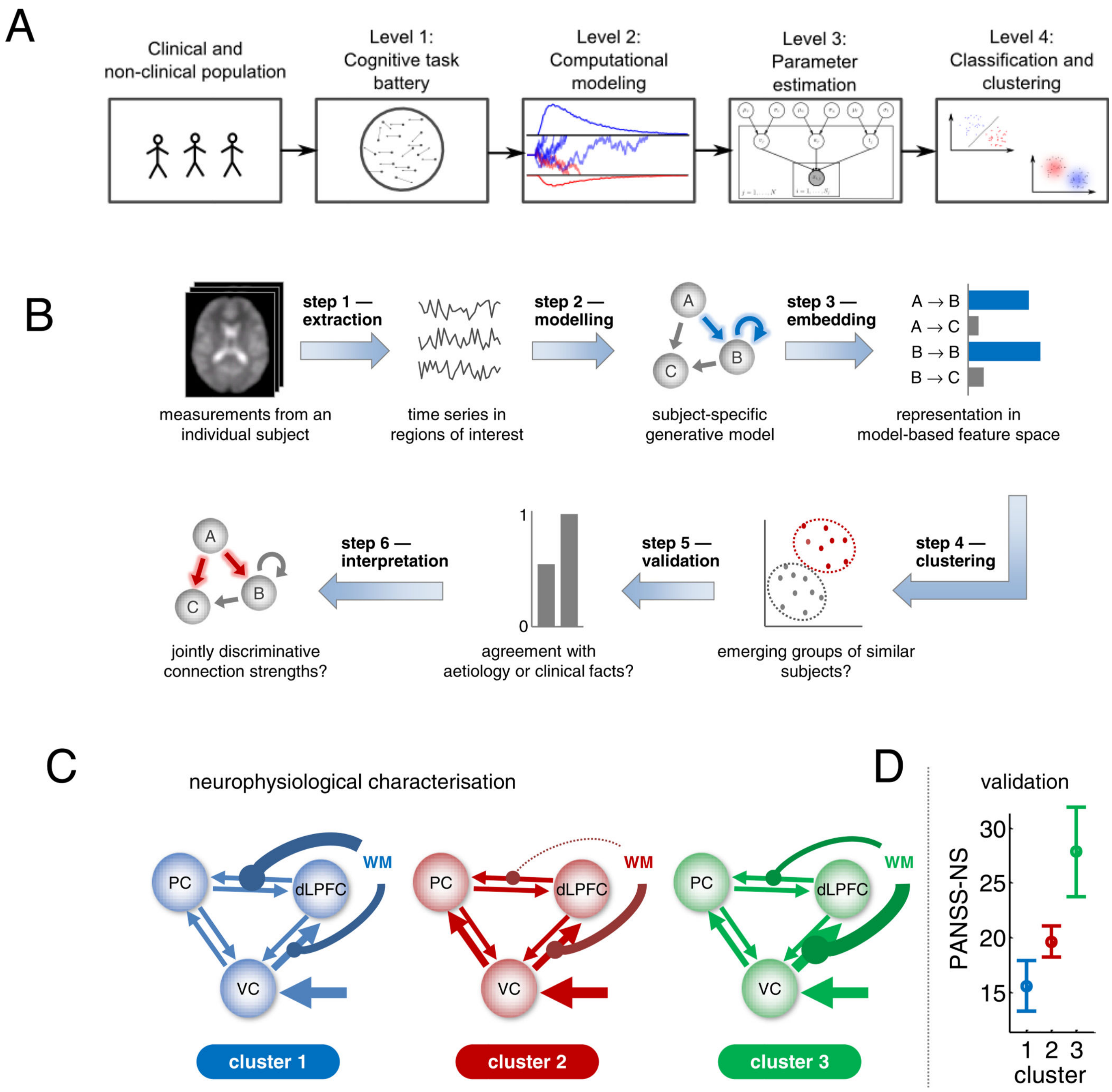
A
Clinical and non-clinical population
Level 1: Cognitive task battery
Level 2: Computational modeling
Level 3: Parameter estimation
Level 4: Classification and clustering
B
step 1 — extraction
step 2 — modelling
step 3 — embedding
measurements from an individual subject
time series in regions of interest
subject-specific generative model
representation in model-based feature space
A → B
A → C
B → B
B → C
step 4 — clustering
step 5 — validation
step 6 — interpretation
emerging groups of similar subjects?
agreement with aetiology or clinical facts?
jointly discriminative connection strengths?
1
0
C
neurophysiological characterisation
PC
dLPFC
VC
WM
cluster 1
cluster 2
cluster 3
D
validation
PANSS-NS
15
20
25
30
1 2 3
cluster

Figure 3.

(A) Illustration of the 4 levels of computational psychiatry. Clinical and nonclinical populations are tested on a battery of cognitive tasks. Computational models can relate raw task performance (e.g. RT and accuracy) to psychological and/or neurocognitive processes. These models can be estimated via various methods. Finally, based on resulting computational multidimensional profile, training using learning algorithms can either uncover groups and subgroups in clinical and healthy populations, or relate model parameters to clinical symptom severity. (B) Conceptual overview of model-aided clustering of fMRI data. First, separately for each subject, BOLD time series are extracted from a number of regions of interest. Second, subject-specific time series are used to estimate the parameters of a model. Third, subjects are embedded in a score space in which each dimension represents a specific model parameter. This space implies a similarity metric under which any two subjects can be compared. Fourth, a clustering algorithm is used to identify salient substructures in the data. Fifth, the resulting clusters are validated against known external (clinical) variables. Once validated, a clustering solution can, sixth, be interpreted mechanistically in the context of the underlying model. (C–D) Model-based clustering of fMRI data from schizophrenic patients in a working memory task. (C) An unsupervised clustering analysis of the patient group only, using Gaussian mixture models operating on dynamical causal model (DCM) parameter estimates, yield the average posterior parameter estimates (in terms of maximum a posteriori estimates) for each coupling and input parameter in the model. This is displayed graphically by the thickness of the respective arrows. (D) The three subgroups, which are defined on the basis of connection strengths, also differ in terms of negative clinical symptoms as operationalized by the negative symptoms (NS) subscale of the PANSS score. (A) was reproduced from Wiecki et al. (2014), (B–D) from (Brodersen et al. 2014), with permission.

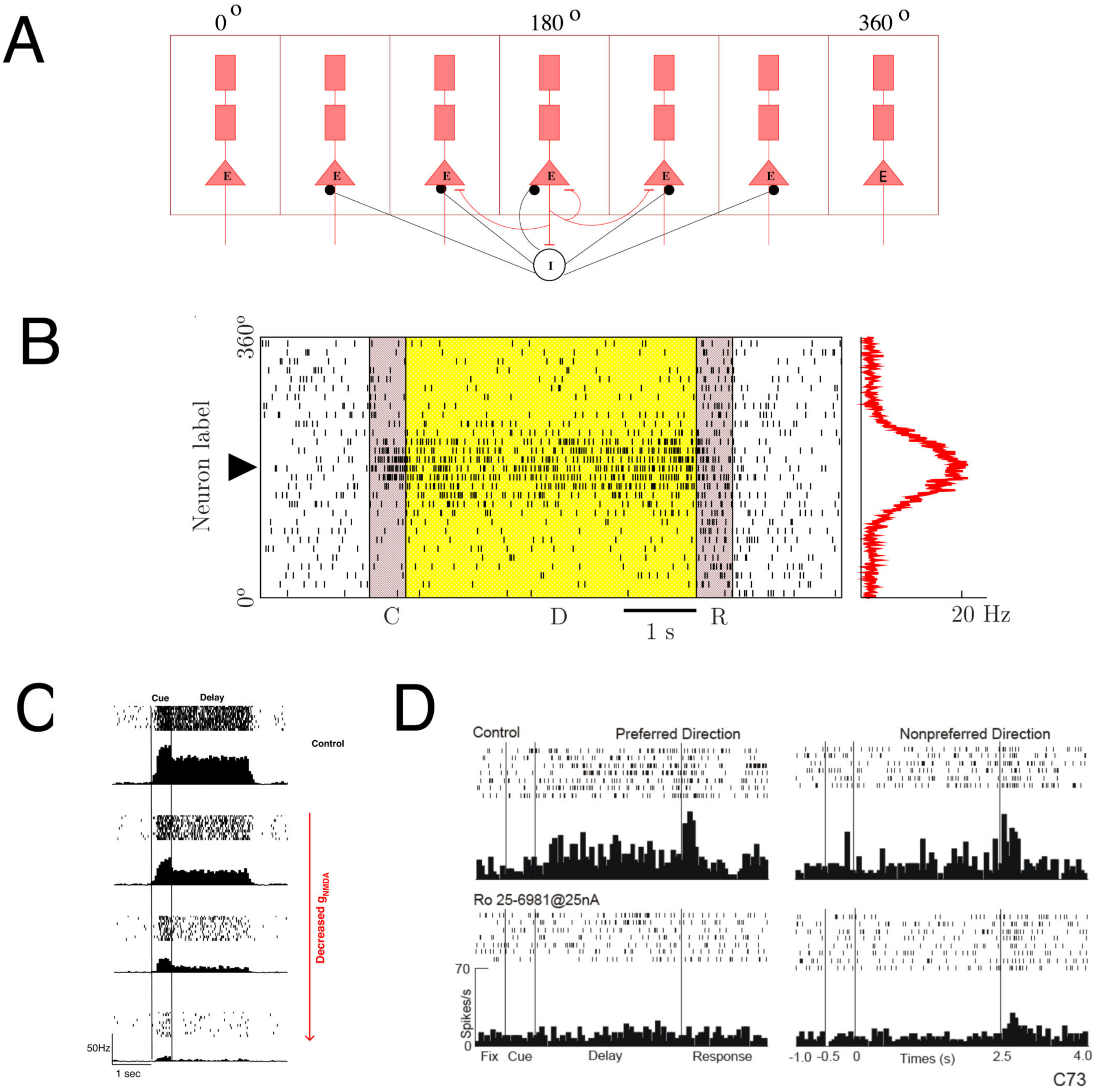

Figure 4.

(A–B) Spiking network model of working memory. (A) Model architecture. Excitatory pyramidal cells are labeled by their preferred cues (0° to 360°). Pyramidal cells of similar preferred cues are connected through local excitatory-to-excitatory connections. Inhibitory interneurons receive inputs from excitatory cells and send feedback inhibition by broad projections. (B) A stimulus is encoded and actively maintained by a self-sustained network persistent activity pattern (a "bump attractor") in a simulation of the delayed oculomotor experiment. C: cue period D: delay period, R: response period. Pyramidal neurons are labeled along the y-axis according to their preferred cues. The x axis represents time. A dot in the rastergram indicates a spike of a neuron whose preferred location is at y, at time x. An elevated and localized neural activity is triggered by a transient cue stimulus and persists during the delay period. (C) The effects of iontophoretic NMDA blockade on working memory activity in a computational model of working memory. Under control conditions, a stimulus cue selectively activates a group of neurons, leading to persistent activity sustained by NMDAR-dependent recurrent excitation. NMDA conductance is reduced from control to 90%, 80%, and 70% (to bottom) of a reference level in a few pyramidal neurons in the network model. Stimulus-selective persistent activity gradually decreases with more NMDAR blockade and eventually disappears in these affected cells. (D) An example of an individual dorsolateral PFC cell recorded from behaving monkey in a delayed oculomotor response task. Upper panels: control condition, lower panels: after iontophoresis of Ro 25-6981 (25 nA), a blocker of NR2B-containing NMDA receptors. The rasters and histograms show firing patterns for the neuron’s preferred direction and the nonpreferred direction (opposite to the preferred direction). Iontophoresis of Ro 25-6981 markedly reduced mnemonic delay period firing to baseline. (B) was adapted from Compte et al. (2000), (C–D) from Wang et al. (2013), with permission.

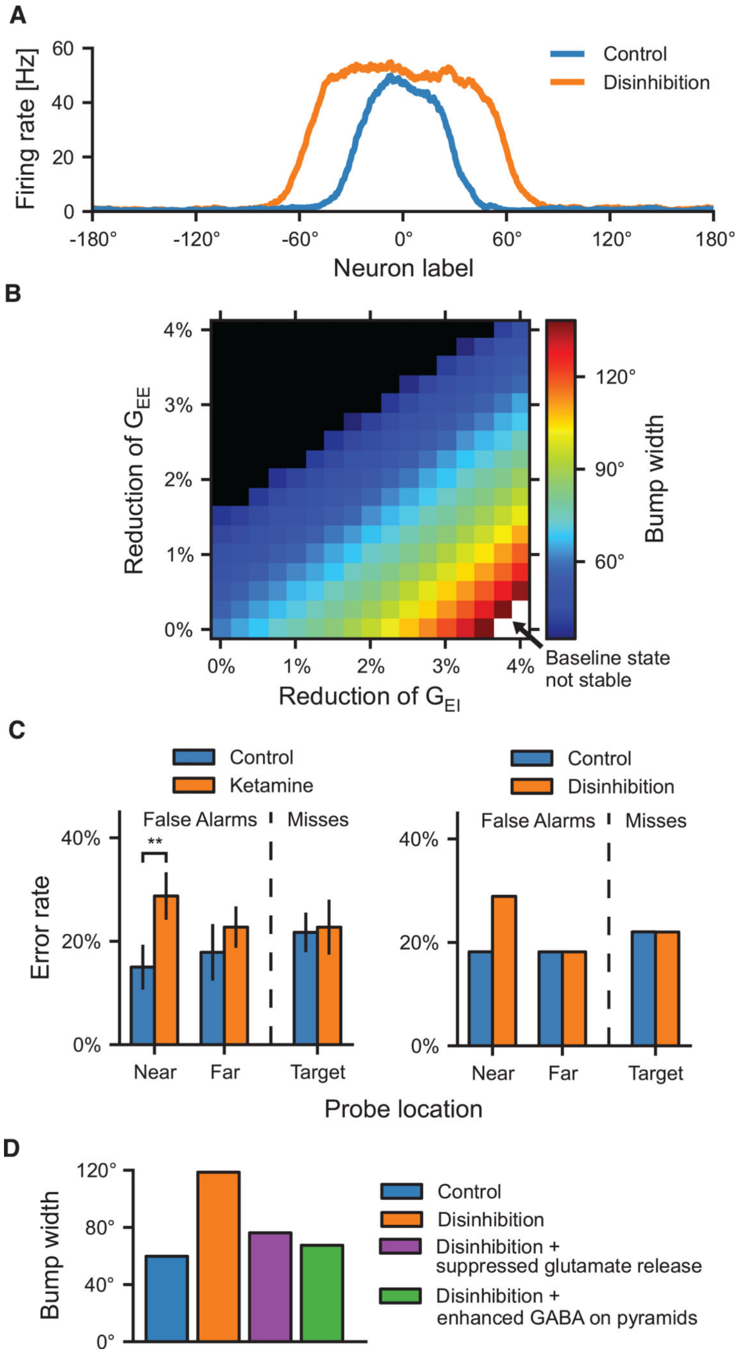

Figure 5.

Computational modeling of excitation-inhibition (E/I) balance in working memory circuits. (A) A spatial working-memory model can generate a bump-shaped stimulus-selective persistent activity pattern following stimulus withdrawal. Disinhibition, mediated by NMDAR hypofunction on interneurons, broadens working-memory representations at the neural level. (B) The parameter space of NMDAR hypofunction highlights the importance of E/I balance for working memory function. If the E/I ratio is elevated as in disinhibition, the width of the representation increases. In contrast, if the E/I ratio is reduced too much through weakened recurrent excitation between pyramidal cells, the circuit cannot support memory-related persistent activity (upper left corner). (C) Broadening of working-memory representations was tested using behavioral data from human subjects performing a spatial working-memory task combined with ketamine infusion, a pharmacological model of schizophrenia. Consistent with broadening, ketamine induced errors specifically for near distractor probes (left), as predicted by the model (right). (D) Compensations can restore E/I balance and ameliorate behavioral deficits in the model. We paired the disinhibition mechanism with either reduced excitation (purple) or increased inhibition (green), following proposed pharmacological treatments. Adapted with permission from (Murray et al. 2014).