# Mechanistic Classification of Neural Circuit Dysfunctions: Insights from Neuroeconomics Research in Animals

Steve W.C. Chang, David L. Barack, and Michael L. Platt

Many psychiatric conditions present complex behavioral symptoms, and the type and magnitude of underlying neural dysfunction may vary drastically. This review introduces a classification scheme for psychiatric symptoms, describing them in terms of the state of a dysfunctional neural circuit. We provide examples of two kinds of functional deficits: variance-shifted functionality, in which a damaged circuit continues to function albeit suboptimally, and state-shifted functionality, resulting in an absent or qualitatively different functional state. We discuss, from the perspective of neuroeconomics and related areas of behavioral investigation, three broad classes of commonly occurring symptoms in psychopathology based on selected studies of decision making in animals: temporal discounting, social preferences, and decision making under environmental volatility. We conclude that the proposed mechanistic categorization scheme offers promise for understanding neural circuit dysfunctions underlying psychopathology.

**Key Words:** Animals, decision, electronic circuit, neuroeconomics, psychopathology, reward, state-shifted, suboptimal, variance-shifted

Comprised of constellations of behavioral symptoms, psychiatric disorders frequently frustrate any simple attempt to translate observed phenotype into neurobiological mechanism. Even at the individual symptom level, such translation is challenging and not easily quantifiable. Behavioral symptoms are often compound and thus difficult to interpret. This presents a challenge for understanding their core neurobiological features, creating practical barriers to designing behavioral or diagnostic tests. This difficulty may be amplified when studying the illnesses manifested as a result of dysfunctions in the prefrontal, limbic, and paralimbic regions, which are less well understood, compared with, for example, the occipital cortex. A promising alternative to understanding the neurobiology of psychiatric disorders begins by classifying them according to the ways the underlying mechanisms may fail. In this issue exploring the benefits of a neuroeconomics approach for understanding psychopathology, we outline a mechanistic classification scheme grounded in the principles of neuroeconomic studies of cognition and behavior in animals.

## Variance-Shifted Versus State-Shifted Functionality: Insights from Electronics

Dysfunctional neural circuitry can be functionally classified into two different states based on the outputs of disrupted circuits. As an illustration, consider an electronic circuit designed to produce a specific output. A variance-shifted circuit operates with added noise and, therefore, generates a broadened output distribution, resulting in suboptimal performance. However, a suboptimal circuit may continue to process information (1). By contrast, a state-shifted circuit may generate a completely different functional output, either beyond the expectation of a downstream circuit or failing to generate any output at all, producing a qualitatively different or absent output and resulting in behavior drawn from a different distribution altogether (1).

As a simplified analogy, a simple band-pass filter illustrates the different classes of damage-induced functional states. A change in circuit resistance or capacitance will change the effective cutoff frequency, while a short in the system effectively halts filtering (1). Changes in a circuit's resistance will result in a noisier output, analogous to psychiatric conditions in which afflicted individuals show difficulty in evaluating changes in the environment. Such damage to the circuit reveals its critical role for producing adaptive, normal behavior. In contrast, the presence of a short in the system will prevent filtering of relevant information, analogous to situations where afflicted individuals completely lose sensitivity to changes in the environment. In this case, the state-shifted circuit reveals its necessary role in the production of a particular behavior.

The intricate balance between circuit components can result in functional changes that are either large and noticeable or small and subtle. Some neuropsychiatric symptoms only differ from others slightly, whereas others are so specific to a condition that they serve as a diagnostic hallmark. Furthermore, because of the complex and multilayered nature of neural circuits, initial perturbations may result at first in a state-shifted circuit that, due to neural plasticity, resolves back to a variance-shifted, or even fully restored, state. In summary, psychiatric symptoms may result from a relatively preserved neural circuit operating with added noise, producing deviant and suboptimal behavior (variance-shifted functionality). Alternatively, it may arise from a shorted circuit producing completely different or absent behaviors (state-shifted functionality).

The two damaged states can be described in terms of neural network models as well. In a trained neural network, the organizational principles involve individual computational units, or nodes, whose functionalities may be obscure and may encode information idiosyncratically (2,3). A variance-shifted functional state may result from damage to peripheral nodes, whereas a state-shifted state may be induced by damage to a central node in the network. The two functionalities can also be described based on the output statistics of an implicated circuit. A variance-shifted dysfunction in a neural circuit may produce circuit (or behavioral) outputs characterized by a broadened and/or attenuated distribution compared with optimal functionality (thus less specific or more noisy). In contrast, a state-shifted dysfunction in a circuit may produce an output

From the Department of Neurobiology (SWCC, MLP), Duke University School of Medicine; and Center for Cognitive Neuroscience (SWCC, DLB, MLP) and Departments of Philosophy (DLB), Evolutionary Anthropology (MLP), and Psychology and Neuroscience (MLP), Duke University, Durham, North Carolina.

Address correspondence to Michael L. Platt, Ph.D., Duke University, Department of Neurobiology, Center for Cognitive Neuroscience, LSRC Building Room B243F, Durham, NC 27710; E-mail: platt@neuro.duke.edu.

Received Aug 25, 2011; revised Feb 20, 2012; accepted Feb 21, 2012.

0006-3223/$36.00
doi:10.1016/j.biopsych.2012.02.017

drawn from a completely different distribution (thus qualitatively different) or may result in a complete failure to produce any output. It is worthwhile to note that a state shift could occur in the direction of extreme enhancement, resulting in exaggerated behavior such as positive symptoms in schizophrenia.

Our classification scheme, though neither exceptionless nor exhaustive, provides insight into the possible mechanisms underlying psychiatric symptoms. The two deficit types may occur simultaneously or sequentially (and the distinction sometimes can be ambiguous until a given circuit is fully understood) but may provide novel mechanistic insights into psychopathology and inform the relationship of pathology to health. This approach differs fundamentally from the *Diagnostic and Statistical Manual of Mental Disorders*, the *International Classification of Diseases*, and the like, which are designed to describe a disorder using a list of behavioral symptoms for diagnostic purposes. The present scheme is useful for directly comparing the functionality of neural mechanisms and their corresponding behaviors across normal and dysfunctional states of the brain. A successful distinction between variance- and state-shifted dysfunction is constrained by our understanding of a given circuit. For example, a variance-shifted dysfunction under one functional criterion could be seen as a state-shifted condition under a different framework. Such ambiguity, which is present in any classification scheme, can only be resolved through more comprehensive understanding of a circuit.

## Examples from Oculomotor System

Examples from oculomotor system help illustrate the two distinct dysfunctional states described above. The superior colliculus and frontal eye fields belong to a distributed oculomotor circuit spanning cortical and subcortical structures (4,5). Frontal eye field lesions increase variability in saccade trajectories and severely disrupt selection of targets in the contralesional hemifield (6). Frontal eye field lesioned animals, however, can still saccade (6). By contrast, superior colliculus lesions temporarily abolish contralesional saccades altogether (7). They also permanently increase saccade latencies and eliminate the animal's ability to make express saccades (saccades with reaction times less than 100 msec in monkeys) in a gap task (7), designed to bypass the time required to disengage from visual fixation by inserting a gap between the offset of a fixation stimulus and target onset (8). Therefore, for saccades, frontal eye field disruption results in noisy (i.e., variable) performance but preserves overall functionality, a variance-shifted dysfunction. Superior colliculus damage alone, by contrast, is sufficient to temporarily abolish saccades, which is consistent with a state-shifted dysfunction. These examples demonstrate that distinct mechanistic deficits can impair or abolish normal function.

## Neuroeconomics of Decision Making in Animals

Neuroeconomics, a discipline that marries the mathematical formalisms of classical economics, the psychophysical methods of behavioral economics, and contemporary neurosciences (9–11), provides an illuminating test of the functionality-based classification scheme for defining mechanistic pathologies in decision making (for a review regarding the benefits of animal models in neuroeconomics, see [12]). The approach applies mathematically tractable economic formalizations to the nervous system and focuses on basic economic concepts such as utility (9,13–15), risk (16,17), and temporal discounting (18,19), providing quantitative frameworks for examining the neural mechanisms underlying cognitive processes (12).

The neuroeconomic framework in animal models is advantageous for studying complex forms of decision making by tapping into their innate reward-seeking behaviors while maintaining ethological validity. Unlike in humans, animal models offer access to studying complex behaviors at the resolution of single neurons. Further, insights into different types of mechanistic deficits in neuropsychiatric symptoms can be obtained by studying decisions animals make following perturbation of neural circuits. Thus, animal models of decision making provide valuable insights into characterizing the biological mechanisms of behavior, detailing the formal operations the brain performs in realizing different cognitive capacities.

We discuss a selection of experiments, categorizing the observed deficits as the variance-shifted and state-shifted model of neural circuit dysfunctions. We organize this discussion around three examples of circuit dysfunction in light of neuroeconomics and other related disciplines: disorders of temporal discounting in addiction, social and other-regarding preferences (ORP), and decision making under environmental volatility. Our intention is not to establish necessary and sufficient conditions for connecting a specific dysfunction and a specific neural circuit. Doing so would not be practically possible. Instead, in this exercise, we attempt to label experimentally induced behavioral deficits observed in animals as dysfunctions arising from either a variance- or state-shifted functional state in the implicated circuit. Although this classification scheme can be just as easily applied to any perturbation results (e.g., microstimulation or drug infusion), we focus on lesion studies for their blunt effectiveness in perturbing circuit function.

## Addiction as a Disorder of Temporal Discounting

Single-unit recordings in animals, as well as neuroimaging in humans, have found that striatal dopaminergic signaling is critical for reward-related processing, including motivation and learning (20–22), and that dysfunctional dopaminergic signaling disrupts reward anticipation in drug addiction (for a review, see [23–25]). Firing rates of midbrain dopamine neurons compute economic decision parameters, such as reward probability, reward delay, and reward uncertainty (26–28). Dopaminergic signaling is also involved in evaluating the economic costs and benefits of upcoming rewards. For example, neurons in rodent nucleus accumbens (NAc) encode anticipated reward benefits, without encoding response costs to achieve the reward (28). Such economic computations by the mesolimbic dopamine system may contribute to addiction and other motivation-related disorders.

Temporal discounting describes a time-dependent devaluation of economic value (18). It is a phenomenon observed across multiple species including rodents, monkeys, and humans (18,29,30). When provided an option to choose an immediate but smaller reward over a larger reward with a longer delay, animals reliably prefer the immediate option (31). Addicted individuals discount more than nonaddicted individuals (24,32), as evidenced by behaviors manifested in addiction to cocaine, alcohol, opioid, nicotine, and gambling (for a review, see [32]). Therefore, a disruption in temporal discounting may be a common mechanistic deficit shared by many classes of addiction.

Single-unit recordings in monkeys demonstrate that neurons in the striatum mediate computations underlying temporal discounting (33). Rats with NAc lesions display severe difficulty in choosing a delayed reward option in an intertemporal choice task, suggesting a critical role of NAc in computing economic values of rewards in time (34). Further, NAc lesions do not abolish reward sensitivity altogether but impair the implementation of an optimal (reward-