The Scylla and Charybdis of Neuroeconomic Approaches to Psychopathology

P. Read Montague

As neuroeconomics enters its second decade, it is fair to ask whether it represents a useful confluence of two disciplines or simply a compound word with aspirations to be much more. Neuroeconomics rests on two important assumptions: the first is the ideal rational agent model, which guides the definition of optimal decision making and the valuations that should underlie it. The second is that neuroimaging is capable of providing insights into the neurobiology of the decision-making process. This special issue attempts to stretch the methods and metaphors of neuroeconomics toward the characterization and understanding of psychiatric disorders.

Neuroeconomics gains extra strength from three features that may help it reach beyond its beginnings: 1) its capacity to connect decision-making variables to details of neural circuits including details extracted from model animal systems; 2) its natural connection to computational theory with the best example being reinforcement learning (1,2); and 3) its connection to game theory (3).

All three areas provide support for the current effort to use neuroeconomics to forge a new understanding of certain psychopathologies (e.g., [4]). One hope is that computational modeling and game theoretic structuring of behavioral probes will help leverage this effort. To be successful, these early efforts must meet at least five important criteria:

1. Does the approach produce pointers to consistent neural circuits underlying functions perturbed by mental disease?
2. Do ideas about mental disease gain extra richness from mappings on to computational models, neural structures, or functional activations in brain imaging experiments?
3. Will useful and new endophenotypes emerge from this work?
4. Is the metaphor of the rational agent helpful or limiting?
5. Is the neural data in humans still to crude to reach up to disease or down to basic neurobiologic mechanisms?

The contributions to this special issue touch on all these issues, but these are still early days and the jury is out. However, some of the most fruitful first steps appear to be in the area of neural circuits and their connection to clearly identified derangements in mental function.

Connecting to Neural Circuits

Hartley and Phelps (5) in this issue provide an excellent example of how the detailed exploration of neural systems for fear conditioning can guide the understanding of how anxiety disorders could be acquired and sustained in a human and how they influence decision making. Reaction to and learning about fearful stimuli are part of the normal behavioral repertoire of all mobile creatures; hence, neurobiologists have long been interested in fear conditioning as a model system for relating brain and behavior. Their review details the neural circuitry involved in many aspects of anxiety's influence on decision making. This is one area in which the neural structures and the logic of behavioral acquisition and extinction give much more mechanistic insight than simply studying anxiety in humans in either isolation or brain imaging devices. Delgado and Dickerson (6) make a similar compelling case for considering multiple memory systems in the kinds of complex, real-world choices that humans routinely navigate. Similar to Hartley and Phelps (5), their contribution shows how detailed consideration of animal experiments has immediate impact on how to understand the neural basis of economic choice in humans.

Chang *et al.* (7) escort the neural circuit analysis perspective straight into a proposal for how to use circuit-level descriptions to classify symptoms expressed in numerous psychopathologies. They focused on decision making, social preference, and the valuation of time (temporal discounting) as examples. Modern genetic tools allow for the production of model organisms with specific molecular changes targeted precisely to known neural systems. One can imagine a future here in which a cluster of important symptoms for some psychopathology is isolated, classified according to the information-based scheme outlined by Chang *et al.* (7), and then used to guide the production of a model organism exhibiting deficits in a contributing neural system. These possibilities are quite exciting not only because of the possible insights into basic mechanisms but also because of the potentially fruitful interplay with clinical applications. But herein lies a gap.

What kinds of training programs can prepare a student or fellow adequately for such a vast array of theoretical and empirical approaches to problems that themselves may also require serious clinical understanding? Training programs for cognitive neuroscience, computational neuroscience, psychiatry, and others will have to evolve to meet the needs latent in such profoundly cross-disciplinary endeavors. Computational approaches to psychiatry represent one response to such challenges, but one suspects that new computational approaches alone will not be enough. Increasing the reach and resolution of noninvasive imaging of human brain function will be required to connect the cognitive part of neuroimaging to detailed neurobiologic mechanisms.

Active Social Exchange as a Research Tool

Connecting features of mental dysfunction to animal models has been fruitful at the mechanistic level; however, some behavioral and perceptual events find such a uniquely human expression that it appears unlikely that we can simply guess which behavioral features in a model organism correspond to analogs in the human.

One approach that has arisen in recent years is to probe human mental function in active interpersonal exchange and often using probes designed around ideas from game theory. It is here that computational models have helped guide the design and interpretation of experiments (8–10) (also see [11] for review). King-Casas and Chiu (12) in this issue highlight the important role of social impairments in a range of mental disorders. Their piece conveys a palpable tension between the excitement of gaining new mechanistic understanding in a range of important clinical maladies and

From the Virginia Tech Carilion Research Institute, Department of Physics, Virginia Tech, Roanoke, Virginia; and the Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom.

Address correspondence to P. Read Montague, Ph.D., Virginia Tech Carilion Research Institute, Department of Physics, Virginia Tech, 2 Riverside Circle, Roanoke, VA 24016; E-mail: read@vt.edu.

Received May 9, 2012; revised May 20, 2012; accepted May 21, 2012.

0006-3223/$36.00
http://dx.doi.org/10.1016/j.biopsych.2012.05.010

the vastness of the problem ahead. One challenge for social exchange probes is that we simply do not have good models of social behavior in humans, but we need them desperately. Kishida and Montague (13) make many of the same points but point ahead to experimental settings in which a subject's behavior is studied in group settings, which could serve to reveal new classes of control signals and contributing neural substrates (4,8,14). They make an argument for the pressing need to develop endophenotypes around mental function and the need to extend physical measurements in healthy and diseased human brains. As always, the presence and use of reinforcement learning models (1,2,14,15) continue to be central to such endeavors.

Clinical End Points as Starting Points

This special issue also presents a more traditional starting point of beginning with a known disorder (like addiction or depression) but then using a different perspective (neuroeconomics) to redefine elements of these problems or to provide new hypotheses about their neural substrates.

Monterosso *et al.* (16) make a case for neuroeconomics approaches providing new breakthroughs in the area of addiction. They focus on the way that humans value the near-term and medium-term future (delay discounting) and on prediction-error signaling, an area defined exclusively by computational models and their connection to neural substrates.

Ernst (17) reviewed how goal-directed and reward-related processing can map many of the symptoms in depression to hypotheses about malfunctioning neural substrates. The details in this piece reveal the dependence of the various proposals either on direct elements of computational models or on qualitatively features associated with the models (1,2). Sonuga-Barke and Fairchild (18) make a similarly strong case for understanding important features of attention-deficit/hyperactivity disorder (ADHD) by reconceptualing them as issues related to personal agency and the parts of this agency that malfunction in ADHD. Their piece is provocative in suggesting that the current medical model in this domain is fundamentally misleading and that an economic agency model can go some distance to a better understanding of how to characterize ADHD.

Challenges to a Way Forward

The importance of computational modeling is one theme apparent throughout all the contributions to this special issue; however, these models are still quite simple and in many cases require a significant number of ad hoc choices to make them work. We must remain vigilant and skeptical about facile interpretations of such efforts because of the complex environments in which human minds now operate.

The second challenge is development and its interaction with contributing genetic substrates. Developmental insults combined with genetically prescribed sensitivities provide one important source of malfunctioning mentation later in life.

The third challenge is scope. Decision making and the valuations that underlie it do not exhaust the range of functions required to have healthy cognition. A broadening of issues addressed and approaches used is warranted.

The fourth challenge is the crudeness in space and time of noninvasive neuroimaging and modeling. The problems with high-density electroencephalogram are well known. The two main limitations with functional magnetic resonance imaging are its crude spatial and temporal scales and a remaining gap in our understanding of all the detailed signal sources that contribute to the measurement. Our understanding of this signal space will need to improve vastly before neural activations can be considered mechanisms in the same way neurobiology at the cellular level identifies and tracks mechanisms.

The author reports honoraria from all universities at which he gives talks and honoraria from the MacArthur Foundation for work on the Neuroscience and Law Network.

This work was supported by National Institutes of Health Grants R01DA011723, R01MH085496, R01DA030241, and RC4AG039067 and a Wellcome Trust Principal Research Fellowship.

1. Dayan P, Daw ND (2008): Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* 8:429–453.
2. Rangel A, Camerer CF, Montague PR (2008): A framework for studying the neurobiology of value-based decision-making. *Nat Rev Neuroscience* 9:545–556.
3. Camerer CF, Fehr E (2006): When does "economic man" dominate social behavior? *Science* 311:47–52.
4. Kishida KT, King-Casas B, Montague PR (2010): Neuroeconomic approaches to mental disorders. *Neuron* 67:543–554.
5. Hartley CA, Phelps EA (2012): Anxiety and decision-making. *Biol Psychiatry* 72:113–118.
6. Delgado MR, Dickerson KC (2012): Reward-related learning via multiple memory systems. *Biol Psychiatry* 72:134–141.
7. Chang SWC, Barack DL, Platt ML (2012): Mechanistic classification of neural circuit dysfunctions: insights from neuroeconomics research in animals. *Biol Psychiatry* 72:101–106.
8. Ray D, King-Casas B, Montague PR, Dayan P (2008): Bayesian model of behaviour in economic games. *Adv Neural Inf Process Syst* 21:1345–1353.
9. Hampton AN, Bossaerts P, O'Doherty JP (2008): Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci USA* 105:6741–6746.
10. Yoshida W, Dolan RJ, Friston KJ (2008): Game theory of mind. *PLoS Comput Biol* 4:e1000254.
11. Behrens TEJ, Hunt LT, Rushworth MFS (2009): The computation of social behaviour. *Science* 29:1160–1164.
12. King-Casas B, Chiu PH (2012): Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biol Psychiatry* 72:119–125.
13. Kishida KT, Montague PR (2012): Imaging models of valuation during social interaction in humans. *Biol Psychiatry* 72:93–100.
14. Maia T, Frank MJ (2011): From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14:154–162.
15. Montague PR, Dolan RJ, Friston K, Dayan P (2012): Computational psychiatry. *Trends Cogn Sci* 16:72–80.
16. Monterosso J, Piray P, Luo S (2012): Neuroeconomics and the study of addiction. *Biol Psychiatry* 72:107–112.
17. Ernst M (2012): The usefulness of neuroeconomics for the study of depression across adolescence into adulthood. *Biol Psychiatry* 72:84–86.
18. Sonuga-Barke EJS, Fairchild G (2012): Neuroeconomics of attention deficit/hyperactivity disorder: Differential influences of medial, dorsal and ventral prefrontal brain networks on suboptimal decision-making? *Biol Psychiatry* 72:126–133.