Editorial TRENDS in Cognitive Sciences Vol.xxx No.x

ELSEVIER

Full text provided by www.sciencedirect.com ScienceDirect

Neuroeconomics

The first wave

P. Read Montague

Department of Neuroscience and Computational Psychiatry Unit, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA

"At the very least, brains are economic engines capable of valuing nearly everything that an organism must do in order to survive and propagate."

Anonymous, 2007

A new word: neuroeconomics

According to *Webster's New Millennium*™ *Dictionary of English*, neuroeconomics is 'the study of the brain in making economic decisions, esp. to build a biological model of decision-making in economic environments'. This definition strikes me as odd. Isn't 'decision-making in economic environments' a bit tautological? The concept of a decision always implies 'against some limited resource constraint', otherwise no decision would be necessary. And can anyone really describe for me a noneconomic environment? OK, I'm not a lexicographer, so I'll stop there, but I'm fairly sure that living creatures don't inhabit noneconomic environments even if they are abstractly possible to construct.

But what about neuroeconomics? Apart from its apparent and growing popularity [1–5], why should we consider neuroeconomics a real area with real problems not yet answered? The answer: because organisms die – they can be killed, they can run out of energy and, even in the best conditions, they will age and die. Consequently, choices must be made and some choices are better than others. If organisms lived forever, then there would be no need for the concept of a choice. If energy could be acquired without effort, then the nature of choice for biological creatures would be dramatically different [6]. But alas mobile organisms run on batteries and have always been forced to be rapid-fire economic decision-makers that know how to value their past, present and future. Bad valuations or even a small hiccup in ongoing valuation mechanisms can result in the ultimate loss – death. These constraints have existed ever since life first evolved. Conclusion: real-world economics has always been an implicit partner to the problems tackled by neuroscience.

Only recently has the knowledge and technology of neuroscience reached a point where basic economic questions can be probed in clearly defined experiments. From this perspective, neuroeconomics is a natural extension of the work that has taken place in neuroscience over the past 100 years. But this perspective is only one starting point. One might also begin a discussion of neuroeconomics by addressing the allocation of scarce resources in a more abstract context thought to model the kind of activity that characterizes human economic exchange (broadly construed). This maneuver would introduce different kinds of ideas about the pertinent variables, the dynamics that describe them, and the meaning of the entire framework for our understanding of human choice. This second trajectory might also be considered a reasonable introduction to the issues that interest neuroeconomics.

So which description and starting point is more representative? The best answer is neither. What is clear is that neuroscience and economics have always been natural cousin disciplines – whether or not this is yet broadly recognized – and there is a lot of fruitful work to be done at the interface of these two subjects without getting too hung up in definitions [1,3]. So let's not. Instead, let's try and focus our efforts on important components of decision-making that humans carry out. This special series on neuroeconomics includes five articles that are summarizing, yet forward-looking and synthetic. These articles all address issues fundamental to behavioral economics [7] and to neuroscience [1]. A common theme runs throughout all the articles – valuation. How does the nervous system differentially value time? How does the nervous system value a monetary gesture from another human? How does the nervous system value the fairness or unfairness of a partner in an exchange? How does the nervous system value unfairness that affects someone else? These are intrinsically interesting behavioral questions, and for a species as socially connectable as humans, these are important questions for understanding the human mind and its relationship to brain function.

The first two reviews in the series address issues that underwrite all the above questions: the valuation of time (Berns, Laibson and Lowenstein, *Trends.cognit.sci* [in press]) and the valuation and expression of fairness during exchanges with other humans (Camerer and Fehr, in this issue of *TICS*). Both problems are complex at many levels and so, as Camerer and Fehr say in their review, the 'first wave of studies' is only just now being logged into the literature. Below, I outline briefly some of the issues addressed by these two installments and give pointers to the future articles in the series.

Time is money

Time is a slippery idea. It's just like a cloud – seemingly solid and easy to spot at a distance, but hard to see and to grasp close up. Even in our physical theories, time still retains a certain slipperiness. Over the last several

Corresponding author: Montague, P.R. (read@bcm.tmc.edu).
Available online xxxxxx.

www.sciencedirect.com 1364-6613/$ – see front matter doi:10.1016/j.tics.2007.07.005

hundred years, physics has provided an escalating sophistication in what are called 'coordinate-free' descriptions of physical phenomena. The basic idea is that physical theories about the world should not change just because we change our point of view. A teacup with a handle on the right should be the 'same object' to someone sitting across the table from me who views the handle on the left. Therefore, physical theories about the teacup should take care of this by making themselves invariant to spatial rotations. That is, they should describe the same object despite any rotation in point of view. The same is true of translations in space – it should be the same teacup after being moved 10 cm in a straight line. The same is also true of translations in time – it is the same teacup even after letting it wait around for a bit (as long as nothing is added or taken away from the teacup).

Despite this progress in our physical theories, there is no complete physical account of the 'nature of time', yet we know that evolution has equipped our cognition with a range of temporal sensibilities. The intuitive concept of 'time' that we all pretty much share with one another is, like everything else in our cognitive toolbox, a construct – a kind of constructed fiction that at the least gives us ways to order (and therefore prioritize) our experience.

The idea that perceptual time is a flexible construct is well recognized by neuroscientists [8–10], but it is the ordering of perceptual events that is crucial here because it opens the door for differentially valuing one perceptual event over another perceptual event at a future time. And it is this ordering and prioritization of our experience that Berns, Laibson and Lowenstein address in their piece. How the 'psychological mechanics' governing the ordering and valuing of temporal experience will influence our understanding of our physical theories remains to be seen – but the issues addressed by these authors are important.

These authors start with the empirically demonstrable fact that, all things being equal, 'the perceptual present' is more valuable than 'the perceptual future'. Technically, the problem is called intertemporal discounting, and these authors address behavioral and neural findings that address three important features of intertemporal discounting: anticipation, self-control and representation. They review experiments suggesting that anticipation of a future rewarding event is itself reward and therefore value-laden. This fact plays into the second theme – self-control. Lastly, they take on the issue of how a choice problem is represented (i.e. framed) and how representations can dramatically change the character of the choices made.

But the issues surrounding time valuation run deep. Waiting for rewarding future outcomes [11–13] is different from the experience of or delay in aversive outcomes [14–16], and both types of experiences are subject to dramatic framing effects [17]. Moreover, timelines for perceptual experience are neither unitary nor unchanging. There is an extremely rich literature in rodents on how the brain represents time, the flexibility of these representations, and the neuroanatomical substrates necessary to support proper perceptual timing [9,10]. Modern neuroimaging work on temporal perception stands atop this rich foundation from the animal literature and has shown that multiple timelines are maintained during ongoing perceptual judgments of temporal order [18]. Much remains to be done, but the importance of understanding the neural underpinnings of time representations has ramifications for nearly every question in cognitive neuroscience.

Fairness

The second installment by Camerer and Fehr addresses another fundamental issue for human cognition – fairness. The central question here is: When should one cooperate with another agent? Fairness instincts have come under the scrutiny of neuroimaging using tasks lifted from the world of behavioral economics [19–22]. Camerer and Fehr give us an excellent synthesis of the first wave of neuroeconomics experiments. They also point out that to date most have been based on game-theoretic framing of the problem of cooperation.

As they say, a game theory model of choice 'is based on three concepts – the action (or opportunity) set, beliefs about chance events, and preferences over actions given beliefs.' By stating it this way, they implicitly make contact with another theoretical depiction of decision-making for individual agents that has been applied to a variety of neuroscience experiments – reinforcement learning. In the reinforcement learning model, an agent comes equipped with: (i) a representation of a problem (an underlying state space); (ii) a value function defined over this state space that assigns a 'long-term goodness' measure to each state; (iii) a guidance signal (error signal) generated by real or contemplated changes in state; and (iv) a decision function that takes the error signal or values (or both) as input and produces a choice as output. These two frameworks map naturally onto one another with the proviso that the reinforcement learning models have generally been applied to behaviors of individual decision-makers or their brain responses and the game theoretic descriptions are at least motivated more by how collections of agents will (or should) behave under prescribed constraints (e.g. see refs [23,24]).

One of the more difficult problems associated with fairness and cooperation is how and why such psychological mechanisms evolved, and beyond the early work of William Hamilton and Robert Trivers, modern efforts have provided some insight into the conditions under which cooperation and reciprocity should evolve (e.g. refs [25,26]).

Future installments

This series on neuroeconomics only scratches the surface – it is an area that is growing rapidly and is built upon areas of neuroscience and behavioral and experimental economics with a rich experimental history. Merely touching upon all the relevant areas in this short overview is impossible. However, for neuroeconomics to mature and contribute in an important way to our understanding of mind and brain it must reach up and reach down. It must reach up to provide new insights into how the human brain builds and sustains a representation of itself [27]. It must also reach down to inform and motivate new animal experiments where the full range of neurobiological techniques can be exploited. Such efforts are already underway [28–30], but these kinds of approaches must

Please cite this article in press as: Montague P. Read, The first wave, Trends Cogn. Sci. (2007), doi:10.1016/j.tics.2007.07.005

multiply to connect the underlying neurobiology to the important economic algorithms it supports.

Further installments are planned, covering the issue of preferences and utilities, the neural signals related to internal currencies used by the nervous system to rate outcomes, choices and contemplated behaviors, and finally ideas about how specific brain regions represent and simulate outcomes about the world and other humans.

References

1 Glimcher, P.W. (2003) *Decisions, Uncertainty, and the Brain. The Science of Neuroeconomics.* MIT Press
2 Glimcher, P.W. and Rustichini, A. (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306, 447–452
3 Camerer, C.F. and Fehr, E. (2006) When does "economic man" dominate social behavior? *Science* 311, 47–52
4 Ainslie, G. and Monterosso, J. (2004) A marketplace in the brain? *Science* 306, 421–423
5 Lehrer, J. (2006) Driven to market. *Nature* 443, 502–504
6 Montague, P.R. (2006) *Why Choose This Book?* Dutton Press
7 Camerer, C.F. (2003) *Behavioral Game Theory.* Princeton University Press
8 Mauk, M.D. and Buonomano, D.V. (2004) The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 304–340
9 Buhusi, C.V. and Meck, W.H. (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nature Neurosci. Rev.* 6, 755–765
10 Meck, W.H. (2006) Neuroanatomical localization of an internal clock: A functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. *Brain Res.* 1108, 93–107
11 Laibson, D.I. (1997) Golden eggs and hyperbolic discounting. *Q. J. Econ.* 62, 443–477
12 McClure, S.M. *et al.* (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507
13 McClure, S.M. *et al.* (2007) Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804
14 Wager, T.D. *et al.* (2004) Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* 303, 1162–1167
15 Berns, G.S. *et al.* (2006) Neurobiological substrates of dread. *Science* 312, 754–758
16 Tom, S.M. *et al.* (2007) The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518
17 De Martino, B. *et al.* (2006) Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687
18 Stetson, C. *et al.* (2006) Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 51, 651–659
19 McCabe, K. *et al.* (2001) A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11832–11835
20 Sanfey, A.G. *et al.* (2003) The neural basis of economic decision making in the ultimatum game. *Science* 300, 1755–1757
21 Delgado, M.R. *et al.* (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618
22 King-Casas, B. *et al.* (2005) Getting to know you: Reputation and trust in a two-person economic exchange. *Science* 308, 78–83
23 Daw, N.D. *et al.* (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879
24 Seymour, B. *et al.* (2004) Temporal difference models describe higher order learning in humans. *Nature* 429, 664–667
25 Axelrod, R. and Hamilton, W.D. (1981) The evolution of cooperation. *Science* 211, 1390–1396
26 Nowak, M.A. and Sigmund, K. (2005) Evolution of indirect reciprocity. *Nature* 437, 1291–1298
27 Damasio, A.R. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* Harcourt, Heinemann
28 Barraclough, D.J. *et al.* (2004) Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410
29 Sugrue, L.P. *et al.* (2004) Matching behavior and the representation of value in parietal cortex. *Science* 304, 1782–1787
30 Dorris, M.C. and Glimcher, P.W. (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378

Please cite this article in press as: Montague P. Read, The first wave, Trends Cogn. Sci. (2007), doi:10.1016/j.tics.2007.07.005