Journal of Economic Psychology 31 (2010) 151–154

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Book Review

Neuroeconomics: Decision Making and the Brain. Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, Russell A. Poldrack (Eds.). Academic Press, London (2009). +xviii 538 pp., Hardcover, ISBN: 978-0-12-374176-9

Neuroeconomics is a large, beautifully produced and ambitious book that aims to be to for this emerging field what Gazzaniga's *The Cognitive Neurosciences* is for the emerged one. It offers essays on the history of the discipline; reflections on whether (and what, and why) it offers anything of unique value to science; and reports of cutting-edge research in a number of areas including: the neural mechanisms of choice, valuation and social decision-making; the evolutionary origins of economic behavior; and the various ways of using neuroscience data to adjudicate between competing theories in behavioral economics. In general, I would say the book succeeds in its ambition. Its publication will undoubtedly be welcomed by everyone in neuroeconomics, as well as by cognitive scientists and experimental economists more broadly, and it can certainly serve as a high-quality textbook for courses on the subject at the advanced undergraduate and graduate levels. That this book *is* likely to be so successful raises something of a puzzle, the same puzzle raised by the field itself: how did it come so far, so fast?

It is worth reflecting on the challenges facing any new subject of inquiry in the academy, especially a cross-disciplinary one. In general, and despite abundant praise for the *idea* of interdisciplinary work, the academy remains a place of conservative habits and entrenched departmental structures. Financial and other support for emerging research is almost always channeled through existing departments, and existing departments consistently (although not, of course, universally) hire scientists addressing the customary questions with well-understood methods over those crossing disciplinary boundaries. Theoretical novelty is widely embraced, but methodological, and what one might call *interrogative* novelty are viewed with much more suspicion. This is hardly surprising. Interdisciplinary work is often hard to understand and hard to evaluate. I am sure every reader of this review has watched a job candidate or visiting lecturer serially lose alternating halves of her audience as she switches between biology and psychology, or physics and neuroscience, or anthropology and economics, in an effort to explain and situate her research and its implications. Moreover, cross-disciplinary research is an implicit challenge to each of the established disciplines involved, suggesting as it does than none of these perspectives is adequate by itself. Bewildering and vaguely (if unintentionally) insulting one's colleagues is hardly a recipe for generating warm feelings, and is a fairly poor foundation for a successful research career.

The fact that the field of neuroeconomics is flourishing despite such challenges suggests the strength of the perceived need for a supplement to business-as-usual in both cognitive neuroscience and experimental economics. What is it that each of these disciplines offers the other, or needs from the other, that has made the match so productive? The evidence of this volume suggests that most of the practitioners of this new field are asking themselves the same question, and although there is no consensus on the *right* answer, there are several recurring themes that run through both the explicit justifications for the field and the implicit assumptions that lie behind the research agendas.

Consider first that cognitive neuroscience is itself a relatively young field, that is only now figuring out how to properly deploy its remarkable technological resources – functional magnetic resonance imaging (fMRI) chief among them – to answer scientific questions. For far too long it its early history, cognitive neuroscience was dominated by a "neural correlates" approach to understanding the brain, in which scientists ask (too) simply: what parts of the brain are active when people do 'X'? But discovering neural correlates sheds zero light on the function of the brain or the nature of cognition. That it sometimes seems to provide insight is an illusion created by the myth of selectivity and abetted by the practice of reverse inference. Selectivity is the idea that individual regions of the brain are active under only a very narrow range of task conditions (think of the popular vision of brain function, where vision happens in the back of the brain, language in the left temporal lobe, and executive control near the front). Reverse inference is the common practice of inferring the engagement of a particular cognitive process from the observation of activity in a brain region (Poldrack, 2006). Less formally, it means taking brain activity to have the same cognitive significance, regardless of the task conditions under which it was observed; for instance, assuming that emotion is involved whenever the amygdala is active. Readers will doubtless have seen many examples of this same inferential move.

What's wrong with this practice? Two related things. First, focusing on individual regions assumes that most cognitive action is local, and emerging research suggests to the contrary that much of cognition may rest on irreducibly relational

doi:10.1016/j.joep.2009.11.008

features of brain activity – things like relative spike timing, large-scale coherence, and long range functional connectivity (Gross et al., 2004; Honey et al., 2007; Sporns, Chialvo, Kaiser, & Hilgetag, 2004; Varela, Lachaux, Rodriguez, & Martinerie, 2001). To understand brain function, we should look less at the *action* of individual regions, and more at their *interaction*.

Second, the evidence increasingly points away from selectivity as a prominent feature of brain function. To take just one example, discussed in chapter 16 of this volume, amygdala activity has been associated with the perception of biological motion, the detection of oddball tones, the perception of sharp contours, and framing effects in economic decisions, in addition to various other perception- memory- and emotion-related functions (see Anderson, 2007, 2008, for evidence that activation in a broad range of cognitive tasks is typical across the brain). Given that diversity, the simple observation of amygdala activation in a given task condition sheds little light on the nature of the task, or the function of the neural structure, for one does not know if it is doing the same thing as it was under other conditions, or if one is witnessing some new operation. Even if amygdala *is* doing one thing, it is not something that can be expressed in the typical vocabulary of cognitive functions; amygdala is not an emotion area, or a perception area, or a memory area. What then is it, and how do we find out? We will return to this point, below.

Unfortunately, there is plenty of neural correlate work (and reverse inference) in this volume. I would guess the continuing tolerance for this approach is a result of the fact that even when leading researchers explicitly criticize it – Glimcher, Kahneman, Phelps, Poldrack, and others all do so here – they soften the criticism with the suggestion that the technique can be useful for hypothesis generation, or to give some sense of the functional "personality" of an area, or for some other purpose, and generally conclude that the approach should be "used with caution". I think this is a mistake, and the fastest way for this field to fail would be to get caught up with finding the neural correlates for various economic processes and variables. The solution to the problems exemplified by neural correlate work and reverse inference is not to use the techniques with caution, but to *stop doing it*. Of course, the scientists who desist will then need something else to do, and the main strength of this volume is in showcasing the very many other options that are on firmer inferential and methodological ground.

At the center of those techniques that seem both more scientifically and logically compelling is a re-evaluation of the best way to take scientific advantage of fMRI experiments. Contrary to the impression given in the media, and perpetuated in a few of the essays here, fMRI is *not* a window into the mind, able to discover and localize the subjective experience of experimental subjects. It is not a reliable localizer of feelings, emotions, decisions, cognitive functions, nor even of representations. Rather, it is a (good but not perfect) localizer of (some but not all) *brain activity*; every other claim is necessarily based on theory, argument, and model-based inference. This is to say, fMRI provides us with a very convenient and useful dependent variable, but one that is no different in its scientific role from any other, be it response times, error rates, choice patterns or what-have-you. One implication of this stance that is vital to appreciate is that finding the *same* brain activity under two task conditions should be treated like any other null result. The most reliable experimental designs, and the most inferentially sound outcomes, are those based on finding measurable *differences* in the dependent variable of choice.

One clever application of this perspective involves using the brain as a difference detector. Suppose, for instance, that a given theory predicts that two economic situations should be treated as equivalent, while another predicts they will be distinguished; similarly, on some economic models concepts like utility have a unified role, but serve diverse purposes in others. Chapters 10, 11, 20, 24 and 25 all suggest that finding differences in brain activity in relevantly different task conditions can be an important part of the argument in favor of those theories predicting distinctions, even (or especially) when attempts to find behavioral differences have so far failed.

Another area in which the essays in this volume help point the way is in using fMRI and other brain activity measuring technologies (such as direct neural recording) for straightforward model testing. Naturally, this means having models with testable implications for brain activity, and this ideal is a long way off for most economic models. The one place where the ideal seems closest to fruition is in the areas of conditioning and learning (chapters 12, 22, 24). Here the models are well-specified and quantitatively sophisticated, and researchers have started mapping model-specific concepts onto neuroscientifically relevant quantities, to allow brain activity to be used as a viable dependent variable. The best of these projects depend not on simple predictions of activation in some area 'Y' (for this to be useful often requires knowing what 'Y'-activation signifies), but rather on the specific *modulation* of activity as model variables are manipulated. Progress in this area should be monitored closely, for it will be very instructive regarding the most reliable approaches to modifying existing behavioral models to best take advantage of this new dependent variable.

The reverse is also true. As the introduction to this volume observes, cognitive neuroscience offers a treasure-trove of data, but a relative dearth of unifying theories. One thing that behavioral economics can offer, and that might be taken up into cognitive neuroscience more broadly, is a particular expertise in theory-driven modeling. The more precise and quantitative cognitive models can become – along the lines of those in conditioning and learning or behavioral economics (e.g. chapters 3, 4, 8, 11) – the faster the field can move entirely beyond reverse inference and begin to approach function-to-structure mapping with cross-domain modeling (Anderson, 2008). In light of the apparent lack of selectivity in individual neural structures, the only hope of assigning function to local structure is to look beyond the single-domain study and consider *all* of the various tasks that appear to recruit a given piece of neural real estate. To attribute a function to amygdala, for instance, it will be necessary to consider not just the cognitive model of one or a few emotion-related functions, and the various options for mapping the processes and variables in that model to specific brain areas, but also the cognitive models of all the *other* functions recruiting the area, such that the elements of each model attribute the same operation to the brain areas where they overlap. Finding the function of the amygdala will be something like finding the right letter to go into a box on a (multidimensional) crossword puzzle, determined not just by the answer to a single clue, but by all the clues whose answers cross that box. This makes the task both harder, because it is multiply constrained, but also easier, because it offers the possibility of leveraging information from several sources to make the attribution. Being both a young and an inherently cross-disciplinary field, neuroeconomics is an ideal place for this practice to begin to take root.

Of course, the claim that the fields of behavioral economics and cognitive neuroscience can help each other in these ways is far from uncontroversial. None of the authors here expresses more than cautious optimism, and a few – notably B. Douglas Bernheim (chapter 9) and Charles Gallistel (chapter 27) – express grave doubts. Each author suggests for different reasons that these two fields are unlikely to find any useful common ground, and that no hybrid entity can do what the individual fields cannot. Although there is reason for caution in expressing the possible benefits of any interdisciplinary partnership, these claims strike me as somewhat overstated. Gallistel is quite right to think that the prospects for a reduction of economic concepts to neuroscientific ones are dim – and we should be thankful for that – but reduction is almost never how sciences at different explanatory levels are unified. Thus, a failure of reduction need not mean a failure of useful integration (Bechtel & Hamilton, 2007; Darden, 1986). Similarly, I think Bernheim over-states the case when he claims, first, that economic variables like utility and preference have no meaning outside of the equations in which they appear, and second, that the only way for neuroeconomics to prove its usefulness is to generate a novel economic model-based entirely on neuroeconomic data that outperforms existing behavioral models.

Although it is of course true that the strict scientific meaning of any experimental variable is its operationalized version – carefully circumscribed and defined so as to be directly or indirectly measurable – the idea that such operationalized variables thereby lose their connection with their psychological origins misunderstands the continuing role of this connection in influencing the fertility of the model. Suppose it in fact turns out to be the case that there is a psychological distinction between wanting and liking (chapter 11) such that animals are not motivated by utility maximization, or that there is no single psychological utility function (chapter 24), or that psychologically speaking, gains are not the opposite of losses (chapter 25). Such findings may not necessitate the immediate change of any axiom in any behavioral model so long as it remained highly predictive. But it would certainly change how the results of that model were interpreted – it would change what they *meant* for the nature of the decision-making subjects on which economic theories are ultimately based. This, in turn, would make a difference to how the model was modified in light of anomalies, and how (and whether) it might be extended to cover other phenomena. Shaping the trajectory of theoretical development strikes me as a quintessential sort of scientific contribution, and so points us in the direction of the ways neuroeconomics might prove valuable even if it never generates an entirely novel economic model. Indeed, even if neuroeconomics does nothing more than offer an expanded set of scientific methods or dependent variables to help get at some questions of already acknowledged interest, this alone could be a contribution to science, insofar as having a diversity of approaches in an inherent good. Moreover, the methodological approaches might have extra-scientific advantages. For instance, they might be capable of drawing more public attention and thereby more funding to a scientific question; or they might be capable of attracting scientific talent that would otherwise do something else; or they might be capable of answering *some* questions more quickly or accurately. Any of these benefits might be sufficient to justify the existence of the field.

But whether it has been justified or not, neuroeconomics is here, and seems likely to remain with us for some time. If it does, I suspect that the interesting, thought-provoking, occasionally exasperating articles collected in *Neuroeconomics* will have played an important role in maintaining that longevity. As a whole, they have soberly laid out the promise and prospects of this young field – what it might offer, and how it might fail – in a way that is sure to become a guide *to* research in the field for those outside of it, and a guide *for* research for those within. One can ask little more of any text.

References

Anderson, M. L. (2007). Evolution of cognitive function via redeployment of brain areas. *The Neuroscientist, 13*(1), 13–21.

Anderson, M. L. (2008). Circuit sharing and the implementation of intelligent systems. *Connection Science, 20*(4), 239–251.

Bechtel, W., & Hamilton, A. (2007). Reductionism, integration, and the unity of the sciences. In T. Kuipers (Ed.), *Philosophy of science: Focal issues: Vol. 1. Handbook of the philosophy of science*. New York: Elsevier.

Darden, L. (1986). Relations amongst fields in the evolutionary synthesis. In W. Bechtel (Ed.), *Integrating scientific disciplines* (pp. 113–123). Dordrecht: Martinus Nijhoff.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., et al (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences, USA, 101*(35), 13050–13055.

Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences, USA, 104*, 10240–10245.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*(2), 59–63.

Sporns, O., Chialvo, D., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8*, 418–425.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*(4), 229–239.

Michael L. Anderson
Department of Psychology,
Franklin and Marshall College,
P.O. Box 3003,
Lancaster,
PA 17604,
USA

Institute for Advanced Computer Studies,
University of Maryland,
College Park,
MD 20742,
USA
Tel.: +1 717 291 3826
E-mail address: michael.anderson@fandm.edu

Available online 22 December 2009