Economics and Philosophy, **24** (2008) 407–418
doi:10.1017/S0266267108002046

WELL-BEING AND NEUROECONOMICS

Julian C. Jamison*
University of Southern California

Neuroscience can contribute to economics by inspiring new models, helping to distinguish models that have similar implications for readily available data, and guiding interpretations of decision-making processes by policy-makers. However, there is an additional less straightforward role for it to play: augmenting, along with survey data and other non-revealed-preference sources, assessments of well-being. The need for such augmentation lies in the slightly bizarre stance taken by modern economic theory, namely that economics is concerned only with choices and not with welfare per se. It is shown that this is neither historical nor at all necessary, even within the standard paradigm. Although neuroscience is by no means a panacea for determining true utility, which ultimately remains a subjective concept, it provides a uniquely useful complementary dataset.

What men are poets who can speak of Jupiter if he were like a man, but if he is an immense spinning sphere of methane and ammonia must be silent?

- Richard Feynman

The goal of this note is primarily twofold: to take a critical look at the notion of utility within modern economics; and to ask how neuroeconomics can contribute to that notion, if at all. For most academic economists, the concept of utility is inviolable and sacrosanct, to be questioned neither in terms of its antecedents nor its constituents. Agents have preferences, given as primitives once and for all (at birth? age 18? this is not discussed), against which all else is defined. Those preferences are revealed to the

* I would like to thank many people for conversations and comments on these ideas over the last few years, including (but not limited to) Colin Camerer, Leemore Dafny, Jodi Halpern, Glenn Harrison, Johannes Hörner, George Loewenstein, George Mailath, Tymofiy Mylovanov, Tanya Rosenblat, Tymon Tatur, Jon Wegener and David Wolpert. However, not only are they not responsible for anything I say here, some of them disagree with various of my conclusions.

outside world through choices, and these choices can then be used to perform welfare analysis. Although the terms involved ("preference", "utility" and "welfare") appear to invoke some notion of actual well-being, in the standard framework they are nothing more than placeholders that refer to choices and decidedly not to any real notion of happiness or life satisfaction.[1]

It is naturally impossible to argue with a definition, and any readers who feel that economics as a profession has nothing to do with true well-being – who have never used the idea of a sunk cost to try to make someone feel better; who have never wished that a policymaker would face up to the fact that agricultural subsidies hurt 99% of the population; who have never been frustrated because their significant other didn't quite grasp the implications of marginal analysis for daily life – should perhaps stop here.

This is not to say that all of economics is about well-being, much less hedonic states such as joy or fear. For instance, the design of trading institutions to maximize information disclosure, under the assumption of profit-maximization by firms, has very little to do with philosophical notions of happiness. But as soon as the discussion turns to, say, "consumer surplus" in the efficiency of competitive markets, we need to ask whether the parties involved truly think that they are discussing surplus as defined via choices only, or whether they think that they are, at least in part, addressing something more fundamental. I believe that in practice it is the latter, and indeed that it ought to be the latter – economists have worthwhile contributions to make in this regard.

And while I also believe that revealed preference truly does tell us something valid and valuable about well-being, I don't believe that it is the final word. This is where neurophysiological data (and 'mundane' survey data) can add colour to the picture. Unfortunately, the latter is not a silver bullet either. In fact, there is no final word: as much as we might like a definitive answer (whether from observed choices or neuronal firing patterns), we must face the fact that well-being will remain subjective. This does not imply, however, that all types of data do not have a major role to play in assessing utility for the purpose of economic analysis.

The paper proceeds in three parts. Section 1 takes a very brief look at some of the current discussions and controversies surrounding neuroeconomics, and in particular raises some recurring themes. Section 2 deals with the rationales for defining utility within economics in terms of

[1] For instance, Gul and Pesendorfer (2008) state that, "Discussions of hedonic experiences play no role in standard economic analysis. . ." As an aside, it is worth pointing out that early economists most certainly did not share this view, so it would not be supportable as a historical statement about the content of the discipline. To take but one example, Jevons (1888) begins his chapter 3 ("Theory of Utility") with a section entitled "Definition of Terms". The first sentence reads, "Pleasure and pain are undoubtedly the ultimate objects of the Calculus of Economics."

revealed preference and discusses possible alternatives. Section 3 suggests some possibilities for using neuroimaging and other non-choice data to augment current notions of utility.

1. PRELIMINARIES

As an approximate definition, neuroeconomics consists of experimental, empirical, or theoretical analyses of decision-making that take the physical embodiment of the decision-maker into account. There are some fairly uncontroversial ways in which it is clear that neuroscientific studies can help to advance economics. For instance, brain images can suggest whether models of risk and ambiguity are likely to be more descriptively accurate if they focus on a single joint mechanism or not (see Hsu *et al.*, 2005). In much the same way that some experimental economics papers inspire new theories (rather than test existing ones), neural results can do the same; see McCabe (2008) for an elaboration of this and related ideas. Whether this is considered to be doing neuroscience or doing economics is beside the point.

As a similar example, consider the paper by Singer *et al.* (2006), which claims to show (very roughly speaking) that men and women punish unfairness in similar amounts, but that men enjoy punishing more, as evidenced by known reward areas in the brain. Although this difference in emotion is perhaps not directly relevant to outcomes, it obviously has implications for differentially targeted interventions. On the other hand, as has been pointed out many times before, if there are in fact different behavioural predictions under some circumstances, then we don't need neuroscience to find this out. But even with that limited goal, it is a purely practical question as to which method is more efficient (or feasible) in getting us there. More generally, economics as a field is becoming more and more interested in studying heterogeneous behaviours at the individual level (as opposed to universal behaviours such as the law of demand), and this is a topic that neuroscientists and geneticists know much about.

A somewhat more speculative potential application of neuroimaging to economic policy is of the type found in Wegener *et al.* (2007). We find that the brain systems associated with imagining oneself in the future are precisely those that are associated with thinking about the intentions and behaviours of other agents (sometimes known as "theory of mind"). This can be interpreted as biological evidence of the multi-self model, and in particular it suggests an argument against the default libertarian doctrine that individuals will necessarily make decisions in their own best interest – people have no more inherent predilection to optimize their future utility than they do to optimize an acquaintance's utility. This is not, and never can be, a proof, but it is a suggestive and useful piece of evidence when designing institutions.

Meanwhile, it is clear that economists have a lot to offer to neuroscience, primarily in the design of tasks which can be used to elucidate specific brain mechanisms. Consider the controversy surrounding the question of whether subjects are aware that they are making the 'correct' choice in the Iowa Gambling Task, which involves choosing from four decks of cards that can yield either gains or losses (both trial-by-trial and in overall expected terms). This debate has garnered quite a bit of attention in the literature, as it touches on issues of consciousness and on the prerequisites for rational decision-making.

In brief, Bechara *et al.* (1997) found that at the moment when subjects started to consistently choose a positive-EV deck, they were unable to articulate the reasons for their choice (in response to open-ended questions regarding their knowledge of the game). Maia and McClelland (2004) replicated the task, but asked subjects more directed queries regarding their estimated value of each deck. They found that in this framework the participants correctly gave a higher subjective rating to positive decks as soon as they began to choose them regularly, and they interpreted this as awareness. Finally, Persaud *et al.* (2007) took a slightly more incentive-compatible approach by requiring subjects to place a high or low wager on the next card in sequence, after choosing a deck on each trial. They found support for the original conclusion that participants were making good decisions before being consciously aware that they were doing so, at least to the extent that they didn't start placing high wagers on the good decks until well after they started choosing them.

One immediate problem with all of these studies is that they use hypothetical payments only ("facsimile money" in the words of Persaud *et al.*), they have relatively small sample sizes, and they are not careful to actually randomize the order of the cards in the decks – or to give any common priors to the subjects. However, a much more serious concern with the final study, itself the most sophisticated of the three in terms of the task, is that it completely disregards the possibility of risk aversion. Given that subjects are uncertain throughout the session, it makes perfect sense to provisionally settle on a certain deck but continue to make low wagers at first, while continually updating one's position. This hardly constitutes active evidence against the results or conclusions, which may well be valid despite these shortcomings, but it does display a critical lack of attention to key elements of experimental design.

Another potential contribution of economics to neuroscience is via formal modelling (and thus improved understanding) of competing brain systems, e.g. using the entire toolkit of principal-agent theory. For instance, Brocas and Carrillo (2007) model competing brain systems as interacting agents with asymmetric information and are able to thereby endogenously produce a form of discounting. Again, it may not be immediately clear whether this constitutes doing economics *per se*, nor does it seem to matter.

But it would be odd to believe that there was no role for economists or economic techniques to play in developing such models.

Turning to some of the controversies that have arisen, the paper by Gul and Pesendorfer (2008) has generated a lot of attention, including detailed responses by Camerer (2008) and others in Caplin and Schotter (2008). The grossly oversimplified synopsis of their original paper is as follows:

- Neuroeconomics was not possible until recently;
- Hence, standard economics makes no reference to brains;
- Thus, there is no way for neuroeconomics to contribute to economics.

In short, they start by defining economics (including welfare economics) in terms of [potentially] observed choices only, from which they are swiftly able to conclude that, indeed, choice data is entirely sufficient in order to do economics. In addition to the practical drawbacks of this approach (see discussion in Harrison 2008, and in section 3 below), the authors exemplify the relatively narrow view of a subset of theorists that economists as a whole do not and should not care about any notion of welfare other than one based on revealed preference.

Gul and Pesendorfer also take great pains to attack the methods, perceived sloppiness, and overall lack of results shown by (in their view) most of the current neuroeconomics – and in fact behavioural economics more broadly – literature. Whether or not their criticism is well-founded, it says nothing about the viability and usefulness of the field of neuroeconomics to economists. This brings to mind a statement from Sobel (2005): "Rejecting a model is easy, rejecting an approach is nearly impossible."

2. THEORY

Microeconomics begins with a preference relation, which can be thought of as a selection function between any two alternatives A and B. This function selects A if it is preferred to B; it selects B if the reverse is true; and it selects both if the decision-maker (DM) is indifferent between them. From the preference relation, a utility function can be constructed under suitable conditions; and from there, all sorts of beautiful theorems can be proven. Note, however, that there is no reason that the original selection function has to correspond to revealed preference in the sense of observed choices. Here is a list of a few possible selection criteria:[2]

(a) Would the DM choose A or B?
(b) Before deciding, does the DM claim that she will choose A or B?

[2] These are reminiscent of Kahneman's (1994) four types of utility, e.g. "experienced" or "remembered".

https://doi.org/10.1017/S0266267108002046 Published online by Cambridge University Press

(c) At the moment of consumption, would the DM claim to enjoy A or B more?[3]
(d) After consumption, would the DM claim to have regretted A or B less?
(e) Does a panel of psychologists believe that A or B would maximize the DM's well-being? (taking her individual characteristics into account, of course)
(f) Does a panel of doctors believe that A or B would maximize the DM's lifespan?
(g) Does a panel of biologists believe that A or B is more likely to maximize the number of surviving grandchildren of the DM?
(h) Does a panel of philosophers believe that A or B is morally superior?
(i) Does A or B induce higher levels of activation in the ventral striatum of the DM's brain at the time of choice? What about at the moment of consumption?
(j) Does A or B maximize the DM's true well-being?

Admittedly, the last one is not even conceivably observable, but that simply serves to illustrate the point that the normative theory of utility is completely independent of the interpretation given to the underlying selection function.

As an example, consider the idea of a Perfect Bayesian Equilibrium in a game with incomplete information. This is a normative theory (and a beautiful and powerful one) of how players ought to behave, given particular utility functions – themselves defined (if at all) as "that thing which the players are trying to maximize". It quite purposefully does not say anything at all about where those utility functions come from or how they relate to the primitives of the situation (such as monetary payoffs). And there is an excellent reason for this, namely that it has absolutely no need to.

Meanwhile, Gul and Pesendorfer (2007) provide a useful separation of welfare analysis into three categories: the first studies the possibility of

[3] Some might argue that in many circumstances it is impossible to think of consuming both A and B, even sequentially, which hardly seems like a fair argument in a theoretical exercise. But in any case, the pure version of the standard framework (a) involves the DM making but one complete contingent plan of action at the beginning of time, leaving very few actual choices to observe. This can be circumvented with appropriate separability and IIA assumptions, but versions of those can be applied just as well to (c) – meaning the difference is a matter of degree only. Put another way: if the choice is between apples and oranges, then there are plenty of data before, during, and after consumption of each. If the choice is whether to spend the rest of one's life in Florida or Norway, no sensible observer should claim that the DM has enough information for her decision to constitute a definitive welfare statement one way or the other; in expected terms, perhaps – but the whole point is to ask, why stop there?

Pareto efficiency across environments (and hence the persistence of various institutions); the second evaluates the behaviours induced by different policies (whether governmental, corporate, or individual) against a stated objective function; and the third is purely normative, asking what policies one ought to pursue in the first place.[4] The lacuna in their partition is that many exogenously given objective functions in the real world do not concern themselves with observable outcomes but instead with the maximization of (some weighted sum of) well-being. Even if the economist *qua* economist has no role getting his feet wet in the moral philosophical domain of which objective function to choose, he can be very helpful to a policymaker who desires to maximize some measure of welfare – assuming they can first agree on the appropriate definition of welfare. The point is that the skills of the economist can be brought to bear regardless of the definition they ultimately settle upon.

Despite this inherent flexibility, it has been commonplace among economists (or at least theorists) to interpret the utility function as deriving from selection criterion (a) above.[5] However, it seems to me (and, I would be willing to wager, the majority of actual policymakers) that for many purposes (c), (e), (i), and especially (j) are at least as compelling. Indeed, the introduction of the canonical graduate textbook (Mas-Colell *et al*. 1995) refers to "... the decision maker's tastes, as summarized in her preference relation." And later: "Welfare analysis concerns itself with the evaluation of the effects of changes in the consumer's environment on her well-being."[6] Since neither "tastes" nor "well-being" are ever formally defined in the book, one can only presume that they are being used in their lay sense.

As far as I can tell, there are two main reasons espoused for using criterion (a) above, apart from tradition. The first is that economists are often trying to predict behaviour, and for that purpose it certainly makes sense to use behaviour itself as a guide (although again it would be silly to ignore other sources of data even for this purpose; see section 3). This is the positive branch of economics, and for a firm trying to maximize profits this is clearly the appropriate notion. But for a government (or firm, or individual in daily life) trying to maximize welfare, it is not. Of course, such a government may still want to focus on choices insofar as they are trying to influence actual behaviour in order to conform to their goals. However, the appropriate notion of welfare to use when elucidating those

[4] They dismiss the role of economics in this third type of welfare analysis, in large part because they claim that very few examples of it appear in economics journals. It strikes me that such an argument could have been made against every new application of economics, and indeed I am not convinced that economics has no productive role to play in this regard, but that is not the focus of the current paper.

[5] The concept of revealed preference was introduced by Samuelson (1938).

[6] Compare with Gul and Pesendorfer (2008), who believe that "... standard economics has no therapeutic ambition; it does not try to improve the [happiness of the] decision maker".

goals in the first place is entirely separate, and there is no reason to believe that actual (or hypothetical) choices should play a special role in the latter process.

The second reason is the greater availability of data on choices. Economics as a field was for many years associated almost exclusively with markets and prices, and this was in large part due to the fact that that was where the data was. Although we now have much broader access to data sources, it is true that many of them still involve choice behaviour. However, this is a purely empirical question and hardly a theoretical rationale for a particular definition of utility. Furthermore, even if one could observe all previous choices for an 80-year-old individual, this would be nowhere near enough information to deduce a complete utility function (under any sort of realistic assumptions).

There is potentially a third reason, though it does not seem to be mentioned as often. This is that evolution 'created' the sensation of well-being exactly in order to affect behaviour, so there is good reason to believe that the two are closely intertwined. That is, even if we are interested in (j) at the end of the day, perhaps (a) is a sufficient statistic. One immediate concern here is that this argument does not apply to anyone above the age of about 45, since from an evolutionary standpoint such individuals no longer had much impact on the survival of their genes (usually because they were dead by then).

But more interestingly, it does not specify exactly how the 'desired' outcome was to be achieved: If a mother gives up food to support her children, is this because she receives pleasure directly from their welfare? Or because she receives pleasure from thinking that she is doing the [morally] right thing? Or because she receives pleasure from dinner-table conversation with her spouse, allowing her to substitute intellectual consumption for caloric consumption? To a greater or lesser degree, these all serve the purpose of evolution, but obviously they have quite different implications for interventions meant to increase well-being.

I have tried in this section to pursue a distinction between choices as the defining characteristic of a utility function (about which I am sceptical) and choices made in the pursuit of some given utility function (about which economic theory has a lot more to say, almost all of it entirely reasonable). In fact, this seems to open up a fruitful line of inquiry for theorists: how shall we model, in a well-defined way, possible relationships between the selection criteria listed above? Why not apply our methods to utility itself?

In this vein, numerous lines of attack suggest themselves. One approach would be to apply evolutionary arguments (involving efficiency and optimality), such as those sketched above, to derive conclusions about when revealed preference, experienced utility, and biological fitness are most or least likely to coincide. Another would be to view choices (effected

by a brain with limited resources) as strategic commitment devices that may or may not replicate actual well-being; since commitment will be more useful in some domains than others, this would have predictive power for when and how choices should deviate furthest from well-being.[7] As a last example, consider the idea of self-serving forgetfulness: the body (including the brain) may recall the full pain of a marathon (or of childbirth), while the conscious decision-making self recalls only the outcome (in part as a positive adaptation to hardship) – and once again it is clear that theories can help distinguish the relevant situations.

3. PRACTICE

In this final section, I would like to present some practical considerations concerning the use of revealed-preference data along with neurophysiological data and survey data. The previous section argued that there was no a priori reason to use choices as the sole building block of utility for either individual-level or policy-level decisions. Nevertheless, there are good reasons (such as the evolutionary one discussed above) to think that, even if we are seeking true well-being, revealed preference is a splendid place to begin. It's just that, no matter how convenient or 'clean' it may appear, there is no justification for relying solely on it.

For instance, suppose you wanted to know if I would get more [true] utility from listening to A: "I'm a Believer" by the Monkees; or B: "True Companion" by Marc Cohn. You could look at all my past musical choices (and possibly those of others), posit a theoretical relationship between musical styles, demographics, circumstances, and choices, and then run a regression. This strains credulity on realistic data collection, but even if feasible it might be more efficient as well as more accurate to simply ask me which I'd prefer.[8] On the other hand, if I weren't sure which I would like better, or you thought I might make an error (according to my own actual preferences, as always) in forecasting my corresponding utility levels, then it could be helpful to (e.g.) measure my galvanic skin response – known to be associated with intensity of emotions – while I pondered each of the options in turn (or even listened to a brief clip of the respective background melodies). Utilizing more data sources can't help being an improvement, whether the goal is to predict my future behaviour or to attempt to assess my actual well-being.

As a more in-depth and germane example, consider a patient who must decide between a risky surgery (with a possibility of full recovery

[7] A readable introduction to the experimental psychology literature on prediction errors involving well-being may be found in Gilbert (2006). Jamison and Wolpert (2008) provide a more technical game-theoretic treatment for why agents might strategically choose to direct their own actions via payoff functions that differ from their true utility.

[8] B

but a complementary possibility of death) versus a lifetime of symptom management. This patient has never faced such a decision before and never will again. For obvious reasons, it is impossible to run properly incentivized laboratory experiments regarding such decisions. Depending on the procedure in question, there may be a limited amount of data on similar decisions by similar patients, and this will surely be useful to the patient under consideration.[9] But such data is suspect, in part because those other patients had even less solid evidence to go on, and in part because of selection biases that are even more difficult to avoid in such a life-or-death situation than they are in usual.

Survey data can be useful here, and a whole literature on contingent valuation (see Carson and Hanemann 2005 for an overview) has been developed to deal with these types of situations, arising especially often in regard to health and the environment. This is also suspect, for the usual reasons, since not only does the DM still have limited information about what the possible outcomes will be like, he (consciously or otherwise) may not be sincerely expressing whatever best estimate he has about his utility in those future states. Nevertheless, it is a useful addition to the arsenal.

Neuroimaging techniques add a third potential source of data. One possible method is to begin with a series of hypothetical decisions made by this patient while being scanned, along with a series of real (incentivized) questions about similar topics. These could be in a non-medical context, but the goal would be to observe brain activation patterns that correspond with 'authentic' as opposed to 'inauthentic' hypothetical decisions, calibrated against the known real choices. With such data, it may well be possible to better parse the validity of a planned decision (also observed in the fMRI magnet) regarding the risky surgery. Or it may not; or it may prove to be exorbitantly expensive (although structural medical scans are currently performed all the time); in any case, it should be considered and explored further.

Alternately, suppose that one could obtain data for this same patient about scans taken during imagined future (at the time) states, along with data believed to be correlated with the actual level of well-being in those states (e.g. from revealed preferences or calibrated surveys when the states are later experienced). Then it is possible, or at the very least conceivable, to compute a relationship between neuronal activation in certain regions at the moment of anticipation and the associated true utility. Hence, inferring the same relationship between scans taken while focusing on the possible future medical states and what the ultimate well-being would

[9] Although note that it is difficult to ask the subset of surgery patients who were unlucky and died whether they think they made the right choice. Then again, to a purist there's no reason to actually ask anybody anything.

be in each of them, decision making is improved – not by blindly following colourful pictures of the brain, but by adding a novel and relevant datum that is outside the convex hull of what was otherwise available (so to speak).

To be crystal clear, I am not suggesting that economics suddenly drop the idea of defining welfare only in terms of revealed preference. I am suggesting that it has already done so (or rather, that this was never the case) and that the way forward is to acknowledge this and fully incorporate it into our theories. No one believes that the revealed preference notion of utility applies to young children[10] or to the severely mentally ill – but this is already a subjective conclusion, made by external observers. It is not that we, the observers, wish to impose our own preferences onto others, but that – even with respect to their own preferences – we believe that their choices are not a sufficient guide. How do we decide for which individuals this is the case? Do their choices remain valid in some contexts, and if so which ones? If we believe that we can sometimes identify "mistakes" (choices that conflict with actual well-being), then shouldn't we apply such a test to all individual decisions? Instead of ad hoc answers to these questions, which are rarely even raised explicitly, it is preferable to develop better measures of utility, both theoretically and empirically.

Does this mean that neuroscience will provide the answer to long-standing questions in welfare economics? No. There is no ultimate 'utility node' in the brain, whose level of activation would yield a measure of true well-being.[11] At the end of the day, the judgment as to what constitutes well-being will remain a subjective one. It will be made by individuals and by societies, and it will be based on choice data; on survey data; on neurophysiological data; on personal experience; on introspection and reasoned analysis; and on other sources, existing or yet to be developed. None of this obviates the need for careful theories (as discussed briefly at the end of section 2) to link these various sources, asking which are more applicable in various circumstances (and why) and delving ever deeper into the constituents of well-being.

[10] As an aside, children are woefully understudied in economics and often provide a useful test case for arguments about libertarianism and benevolent paternalism. On a more practical level, models of preference development (and indeed belief formation) are conspicuously rare, although they are starting to make an appearance.

[11] In fact, the most common neuroimaging marker – blood oxygen levels – can only be used to look at changes (not absolute levels). But that is a technological impediment, and my claim is a more fundamental one: by its very nature, well-being depends on processes in addition to instantaneous values, and hence can never be subject to a single complete measurement. See Halpern and Jamison (2008) for a more in-depth discussion of accurately assessing subjective well-being.

REFERENCES

Bechara, A., H. Damasio, D. Tranel and A. R. Damasio. 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275: 1293–5.

Brocas, I. and J. Carrillo. 2008. The brain as a hierarchical organization. *American Economics Review* forthcoming.

Camerer, C. 2008. The Case for Mindful Economics. In *Foundations of positive and normative economics*, ed. A. Caplin and A. Schotter, 43–69. New York: Oxford University Press.

Caplin, A. and A. Schotter, eds. 2008. *Foundations of positive and normative economics*. New York: Oxford University Press.

Carson, R. T. and W. M. Hanemann. 2005. Contingent valuation. In *Handbook of environmental economics*, Vol. 2, ed. K.-G. Mäler and J. R. Vincent, 821–936. Amsterdam: Elsevier.

Gilbert, D. 2006. *Stumbling on happiness*. New York: Knopf.

Gul, F. and W. Pesendorfer. 2007. Welfare without happiness. *AEA Papers and Proceedings* 97: 471–6.

Gul, F. and W. Pesendorfer. 2008. The case for mindless economics. In *Foundations of positive and normative economics*, ed. A. Caplin and A. Schotter, 3–42. New York: Oxford University Press.

Halpern, J. and J. Jamison. 2008. On the assessment and prediction of subjective well-being. UC Berkeley working paper.

Harrison, G. 2008. Neuroeconomics: a critical reconsideration. *Economics and Philosophy* 24.

Hsu, M., M. Bhatt, R. Adolphs, D. Tranel and C. Camerer. 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310: 1680–3.

Jamison, J. and D. H. Wolpert. 2008. Persona Games. USC working paper.

Jevons, W. S. 1888. *The Theory of Political Economy*, 3rd edn. London: MacMillan and Co.

Kahneman, D. 1994. New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics* 150: 18–36.

Maia, T. V. and J. L. McClelland. 2004. A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Natural Sciences, USA* 101: 16075–80.

Mas-Colell, A., M. D. Whinston and J. R. Green. 1995. *Microeconomic theory*. New York: Oxford University Press.

McCabe, K. A. 2008. Neuroeconomics and the economic sciences. *Economics and Philosophy* 24.

Persaud, N., P. McLeod and A. Cowey. 2007. Post-decision wagering objectively measures awareness. *Nature Neuroscience* 10: 257–61.

Samuelson, P. A. 1938. A note on the pure theory of consumer's behaviour. *Economica* 5: 61–71.

Singer, T., B. Seymour, J. P. O'Doherty, K. E. Stephan, R. J. Dolan and C. D. Frith. 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439: 466–9.

Sobel, J. 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* XLIII: 392–436.

Wegener, J., K. Madsen, M. Christensen and J. Jamison. 2007. Parieto-temporal activation modulates pure time preference. University of Copenhagen working paper.