MPRA

Munich Personal RePEc Archive

Decision-Making: A Neuroeconomic Perspective

Hardy-Vallee, Benoit

University of Toronto

July 2007

Decision-Making: A Neuroeconomic Perspective

Benoit Hardy-Vallée
Department of Philosophy
University of Toronto
215 Huron Street
Toronto, ON, Canada
M5S 1A2
Tel: (416) 978-3311
Fax: (416) 978-8703
ben.hardy.vallee@utoronto.ca
http://decisis.net

Abstract. This article introduces and discusses from a philosophical point of view the nascent field of neuroeconomics, which is the study of neural mechanisms involved in decision-making and their economic significance. Following a survey of the ways in which decision-making is usually construed in philosophy, economics and psychology, I review many important findings in neuroeconomics to show that they suggest a revised picture of decision-making and ourselves as choosing agents. Finally, I outline a neuroeconomic account of irrationality.

1. The Rational Animal In Philosophy, Economics and Psychology

Human life is one long decision tree. (Sterelny)

Rational agents display their rationality mainly by making decisions. Some decisions are basic (turn left or turn right), other ones concern more crucial issues ("to be or not to be"). Even abstinence is decision, as thinkers like William James or Jean-Paul Sartre once pointed out. Since choice is central to life, it is not surprising that many disciplines attempt to properly characterize decision-making. Philosophy, psychology and economics, among others, all have different and sometimes conflicting views about

the nature of decision-making and the conditions that make it rational. Reviewing different construal of decision will therefore illuminate the importance of neuroeconomics at the theoretical level.

Philosophers since Aristotle have reflected on the normative features of decisions—what makes a decision rational or not. In philosophy of mind, the standard conception of decision-making equates *deciding* and *forming an intention before an action* (Davidson *Essays on Actions and Events;* Searle; Audi). According to many analyses, this intention can be equivalent to, inferred from or accompanied by, desires and beliefs. As Davidson explains,

> *If someone acts with an intention then he must have attitudes and beliefs from*
> *which, had he been aware of them and had he the time, he could have reasoned*
> *that his act was desirable. (Davidson, Essays on Actions and Events 85)*

The decisions rational agents make are thus motivated by *reasons*. If Paul sends a manuscript to a publisher, his decision is explained by a desire to be read and a belief that sending the manuscript can lead to publication. Rational actions are *explained* by these reasons, the purported causes of actions. Beliefs and desire are also constitutive of rationality because they *justify* rational action: there is a logical coherence between beliefs, desires and actions.

Rational decision-making is also a core concept in economics. According to a standard definition, economics is the "science which studies human behavior as a relationship between ends and scarce means which have alternative uses" (Robbins 15). This definition shows the centrality of decision-making in economic science: since means are scarce, we should use them efficiently to accomplish our goals. The two branches of rational-choice theory, decision theory and game theory, specify the formal

constraints on optimal decision-making in individual and interactive contexts. Rational-choice theory can be construed as logic of action. It specifies which logical conditions the preferences of an agent should meet in order to make her decision rational (see Baron, for an introduction). For instance, preferences must be transitive: if one prefers A to B and B to C, one must also prefer A to C, but should not prefer C to A. An individual agent facing a choice between two actions can make a rational decision if she takes into account two parameters: the probability and the utility of each action's consequence. Multiplying the subjective probability by the subjective utility of an action's outcomes allows her to find which action has the higher subjective expected utility. The philosopher's belief-desire model is thus reflected in the economist's probability-utility model: probabilities represent beliefs while utilities represent desires, at least according to the standard interpretation of decision theory.

Game theory considers agents making decisions in strategic contexts, situations where the preferences of at least another agent must be taken into account. Decision-making is represented as the selection of a strategy in a *game*, a set of rules that dictate the range of possible actions as well as the payoffs of conjoining actions. For example, in the prisoner's dilemma, police officers hold in separate cells two individuals accused of robbing a bank. The suspects, Bob and Alice, cannot communicate with each other. The police officers offer them the following options: confess or remain silent. If one confesses—implicating his or her partner—and the other one remains silent, the first is freed and the other one receives a ten-year sentence. If they both confess, they will each serve a five-year sentence. If they both remain silent, they will serve only a two-year sentence. This situation can be represented as the following payoff matrix (Fig.1), where Bob's payoff is in bold characters and Alice's in italics.

		Alice	
		Confess	Remain silent
Bob	Confess	(-**5**, -*5*)	(-**10**, *0*)
	Remain silent	(-**10**, *0*)	(-**2**, -*2*)

Fig.1. Payoff matrix in the prisoner's dilemma

Assuming that Bob and Alice have common knowledge—everybody knows that everybody knows that everybody knows... *ad infinitum*—of both each other's rationality and of the rules of the game, they should confess. Confessing gives you either freedom or a five-year sentence, whereas remaining silent brings either a two-year or a ten-year sentence. They will expect each other to make the best move, confessing, since the best reply to this move is also confessing. Even if they would be better off by remaining silent, this choice is suboptimal: should the other one confess, they risk a ten-year sentence. In other words, they should not choose the cooperative move. Although the prisoner's dilemma is a bare-bones picture of strategic decision-making, its simplicity allowed generations of scholars to investigate the nature of strategy in economic, social, psychological, political, or even military contexts (see Poundstone for a historical review).

Finally, decision-making is also a topic for experimental economics, behavioral economics, cognitive science and psychology (I will refer to these empirical approaches of rationality as 'psychology'). Whereas philosophy and rational-choice theory focuses

more on normative or prescriptive issues, psychology is generally interested in the descriptive aspects of decision-making. Psychologists study how subjects make decisions, and on which mechanisms do they rely for making them. Their patterns of inference and behavior can then provide an empirical counterpart to rational-choice theories. In numerous studies, Amos Tversky and Daniel Kahneman showed that decision-makers' judgments deviate markedly from normative theories (Kahneman; Tversky; Kahneman, Slovic and Tversky). Subjects tend to make decisions according to their 'framing' of a situation (the way they represent the situation, e.g. as a gain or as a loss), and exhibit loss-, risk- and ambiguity-aversion. For instance, they prefer a sure gain of $10 instead of 10% chance of winning $110, even if the second option's subjective expected utility is higher.

Psychological studies also showed that we are not as selfish and greedy as the common interpretation of rational-choice theory suggests. Experimental game theory indicates that subjects cooperate massively in prisoner's dilemma (Ledyard; Sally) and in other similar games, such as the ultimatum game. In this one-shot bargaining situation, a 'proposer' (Alice) makes an offer to a 'responder' (Bob) who can either accept or refuse the offer (Güth, Schmittberger and Schwarze). Alice proposes to split an amount of money between them. If Bob accepts her offer, he keeps the amount he is offered and Alice keeps the rest. If Bob rejects it, both players receive nothing. According to game theory, rational agents must behave as follows: Alice should offer the smallest possible amount, in order to keep as much money as possible, and Bob should accept any amount, because a small amount is better than nothing. If there is $10 to split, Alice should offer $1 and keep $9, while Bob should accept the split.

The ultimatum game has been studied in many contexts by varying parameters such as culture, age, sex, the amount of money, the degree of anonymity, the length of the game, and so on (Oosterbeek, S. and van de Kuilen; Samuelson). Overall, results show a robust tendency: the game-theoretic strategy is rarely played, because people tend to make 'fair' offers. Proposers tend to offer about 50% of the amount, and responders tend to accept these offers, rejecting most of the 'unfair' offers (less than 20%). Since the rules of the game are quite simple, subjects do not deviate from the optimal strategy because they do not understand it. Normal subjects seem instead to have a tendency to cooperate and to value fairness. A similar pattern of behavior is also apparent in the trust game. In this game, Alice has an initial amount of money she could either keep or transfer to Bob. If she transfers it to Bob, the amount is tripled. Bob could keep this amount, or transfer it (partially or totally) to Alice. Game theory predicts that either Alice should keep everything, or if she transfers any amount, Bob should keep all of it. Experimental studies have shown that players in Alice's position invest about 50% of their money and get more or less what they invest (Camerer).

Experimental approaches to rationality can thus be informative for theory, as they highlight two features of our practical rationality. First, our practical reasoning does not honor the axioms of game and decision theory. Second, morality might have been left out of the picture of strategic rationality so far. However, the study of decision-making is recently benefiting from another source of empirical inputs: neuroscience. When we make decisions, complex mechanisms inside our skulls process information and control our bodies. It was therefore predictable that, after psychology, neuroscience would contribute to the study of decision-making. In the next two sections, I briefly review

some studies in the field of decision neuroscience, and then analyze their consequences for the philosophical conception of irrationality.

2. Neuroeconomics and The Brain

The Rational Deliberator turns out to be a well-camouflaged Adaptive

Responder

(Clark 33)

Neuroscience is the scientific study of the nervous system. Molecular, cellular, behavioral, and cognitive mechanisms are revealed through different means; functional imaging technologies study degrees of activation and locations; single-cell recordings analyze the activity of individual cells; lesion studies try to determine the function of certain brain areas by investigating brain-impaired subjects (see D'Esposito, for an accessible presentation); and computational neuroscience uses computer simulations to support or invalidate hypotheses about brain mechanisms (Eliasmith). Obviously, these researches should shed light on the nature of decision-making. Recently, a conjunction between the neuroscientific study of decision-making and experimental economics has led to the creation of a new field now called *neuroeconomics* (Glimcher; McCabe; Zak). Although many definitions of the field exist, I will limit myself to the one given by philosopher Don Ross: "the program for understanding the neural basis of the behavioral response to scarcity" (Ross 330). Ross's definition connects clearly this field to the traditional endeavour of economics (cf. Robbins's characterization of economics in first section).

Generally speaking, neuroeconomic research proceeds according to the following methodology:

(1) Choosing a formal model of rationality, a decision- or a game-theoretic situation, and then deducing what decisions agents should make.

(2) Testing the model behaviorally to see if subjects follow normative standards.

(3) Identifying the brain areas and neural mechanisms that underlie choice behavior.

(4) Explaining why subject follow/fail to follow normative standards

Of course, many variations are possible at each step: the formal model may be an alternative theory, or tackle a question not addressed by rational-choice theory; subjects may be of different age, sex or cultural background, or they may be subjects who incur cerebral impairment, cognitive deficit, etc. Neuroeconomics research thus proceeds by comparing formal models with behavioral data, and by identifying neural structures causally involved in economic behavior.

2.1 *Homo Neuroeconomicus*

Neuroeconomics explains decision-making as the product of brain processes involved in the representation, anticipation, valuation and selection of choice opportunities. It breaks down the whole process of decision into mechanistic components: certain brain areas may represent the value of the outcome of an action before decision, other ones may represent the value of the action *per se*, and yet other ones may represent these values at the time of the decision. Although such dispersion of

data may appear confusing, economic psychology provides us with a useful framework for understanding the mechanics of rationality at the neural level in a coherent manner. Kahneman and his collaborators suggest that the concept of utility should be divided in subspecies (Kahneman, Wakker and Sarin). While *decision* utility is important (the expected gains and losses, or cost and benefits), decision-makers also value *experienced* utility (the hedonic, pleasant or unpleasant affect), *predicted* utility (the anticipation of experienced utility) and *remembered* utility (how experienced utility is remembered after the decision, e.g. as regretting or rejoicing). Neuroeconomics should identify neural structures and processes associated with these variables or, if necessary, suggest another typology. This *distributed account of utility*, as I will call it here, is a useful tool for organizing the numerous findings in this burgeoning field.

Neuroeconomics and the distributed account of utility can, for instance, provide a more precise explanation of loss-aversion, a robust finding in psychology. Subjects usually give to the loss of $10 a higher impact than a $10 gain. Tversky and Kahneman attribute this aversion to a bias in the representation of the values of gain and loss (Tversky and Kahneman). Instead of postulating abstract cognitive heuristics, neuroeconomics explains loss-aversion as the interaction of neural structures involved in the anticipation, registration and computation of the hedonic affect of a risky decision. The amygdala, a structure involved in fear, emotional learning and memory modulation, registers the emotional impact of the loss; the ventromedial prefrontal cortex predicts that a loss will result in a given affective impact; and midbrain dopaminergic neurons compute the probability and magnitude of the loss (Naqvi, Shiv and Bechara; Tom et al.). Subjects are thus loss-averse because they tend to have or already had a negative response to losses (experienced utility). When they expect a loss to occur (decision

utility), they anticipate their affective reaction (predicted utility). They might be also attempting to minimize their post-decision feeling of regret (remembered utility).

Similar researches can also illuminate ambiguity-aversion. In many experimental settings, subjects have a strong preference for *risky* prospects (those for which the probabilities are known) over *ambiguous* one (those for which the probabilities are unknown). For instance, let's imagine two decks of 20 cards. There are 10 red cards and 10 blue cards in the first one (risky deck), while there is an unknown proportion of blue to red cards in the second one (ambiguous deck). Agents win $1 each time they pick a red card. Despite a 50–50 chance of winning in both cases, subjects have a marked preference for the risky deck. According to decision theory, there is no reason to prefer one deck to another, but neuroeconomic studies showed that in this case of decision under ambiguity, a stronger activation is found in many areas, especially the amygdala (Huettel et al.). Although decision theory treats ambiguity as a special case of risky decision-making, ambiguous and risky decision-making are supported by two distinct mechanisms. It thus seems that ambiguity-aversion happens because people have a stronger negative affective reaction to ambiguity than risk.

One of the most robust finding in neuroeconomics concerns *decision* utility, the calculation of cost and benefits. According to many findings, this process is realized by dopaminergic systems, a network of structures in 'older' brain areas highly involved in motivation and valuation (Montague and Berns; Berridge). Dopaminergic neurons respond selectively to prediction errors, either the presence of unexpected rewards or the absence of expected rewards. In other words, they detect the discrepancy between predicted and experienced utility. Moreover, dopaminergic neurons learn from their mistakes: they learn to predict future rewarding events from prediction errors, and the

product of this learning (a 'behavioral policy') can then bias action choice. Computational neuroscience identifies a class of reinforcement learning algorithms that mirror the activity of dopaminergic activity (Niv, Duff and Dayan; R. E. Suri and W. Schultz). It is suggested that dopaminergic neurons broadcast in different brain areas a reward-prediction error signal similar to those displayed by temporal difference (TD) algorithms developed by computer scientists (Sutton and Barto "A Temporal-Difference Model of Classical Conditioning"; Sutton and Barto *Reinforcement Learning : An Introduction*). These algorithms are plausible descriptions of neural mechanisms of decision-making implemented in dopaminergic systems. They are not only involved in basic reward prediction, such as food, but also with abstract stimuli like art, brands, love or trust (Montague, King-Casas and Cohen 420).

It might seem unusual to understand decision-making from a reinforcement learning perspective. However, reinforcement learning is not stimulus-response association analogous to classical behaviorism, but is rather the learning of "how to map situations to actions (…) so as to maximize a numerical reward signal" (Sutton and Barto *Reinforcement Learning : An Introduction* 3). Reinforcement learning like TD entails *prediction*: prediction is a cognitive, computational notion, not a behavioral one like *association*. Reinforcement learning, as Selten once noted (16), is thus a basic mode of behavior that cannot be ignored by a theory of rationality.

The main contribution of neuroeconomics to decision theory so far is a new picture of decision-makers as adaptive and affective agents. *Homo Neuroeconomicus* is a fast decider that relies less on logic and more on a complex collection of flexible neural circuits associated with affective responses. Everyday utility maximization is more about feelings and less about the objective outcome of a decision: we use emotions to

anticipate emotions in order to control our behavior toward a maximization of positive emotions and a minimization of negative ones. The neuroeconomic picture of individual rationality is thus affective through and through.

2.2 *Homo Neuroeconomicus* Playing Games

Neuroeconomics provides also new insights into the nature of strategic rationality. Here, the results suggest that strategic decision-making is again a highly affective business. Brain scans of people playing the ultimatum game indicate that unfair offers trigger, in the responders' brain, a 'moral disgust': the anterior insula (associated with negative emotional states like disgust or anger) is more active when unfair offers are proposed (Sanfey et al.). Subjects experience this affective reaction to unfairness only when the proposer is a human being: the activation of the insula is significantly lower when the proposer is a computer. Moreover, this activation is proportional to the degree of unfairness and correlated with the decision to reject unfair offers (Sanfey et al. 1756). Beside the anterior insula, two other areas are recruited in ultimatum decisions: the dorsolateral prefrontal cortex (hereafter DLPFC), involved in cognitive control and goal maintenance, and the anterior cingulate cortex, involved in cognitive conflict and emotional modulation. When an offer is fair, it seems normal to accept it: there is a monetary gain and no aversive feelings. When the offer is unfair, however, the brain faces a dilemma: punish the unfair proposer, or get a little money? The final decision depends on whether the DLPFC or the anterior insula 'wins': Sanfey and his colleagues found that the anterior insula is more active in rejections, while the DLPFC is more active in acceptations. The ACC, itself more active when offers are unfair, behaves as a moderator between the cognitive goal (more money) and the emotional goal (punishing).

Deciding to accept or reject an offer in the ultimatum game is a complex adjustment between multiple goals and values, a complexity that decision theory or the belief-desire model can hardly account for.

Other studies indicate that in similar games where cooperation is common but unexpected by game theory, players enjoy cooperating, what economists refer to as the "warm glow of giving" (Andreoni). In the prisoner's dilemma, players who initiate and players who experience mutual cooperation display activation in nucleus accumbens and other reward-related areas (Rilling et al.). In the trust game, where cooperation is common but not prescribed by game theory, players are ready to lose money for punishing untrustworthy players. De Quervain *et al.* found that both punishing cheaters and anticipating this punishment activate the nucleus accumbens, suggesting that 'revenge tastes sweet' (de Quervain et al.). It follows that fairness, trust and cooperation are common because they have an intrinsic value: even donating to societal causes elicits the same rewarding effect (Moll et al.).

When humans interact with computers, however, these social emotions disappear. In all the aforementioned experimental games, human and computer partners do not elicit the same neural-affective reactions (Rilling et al.). Human partners induce positive or negative emotions, while computer partners induce much less affective reactions. Since computers do not entertain moral values, cultural norms and social emotions like normal flesh-and-bone individuals do, they do not feel the 'warm glow' of cooperation, the 'sweet taste' of revenge or the 'moral disgust' of unfairness. Consequently, a selfish attitude toward an emotionless machine is not seen as morally dubious, just as we do not feel cheated by it.

To sum up, social neuroeconomics shows that money is not the only currency in social exchange. A one-dollar gift is pleasurable, but a one-dollar offer in the ultimatum game elicits negative reaction if the proposer is a human.

3. Neuroeconomics and irrationality

Studies in neuroeconomics are apt to provide meaningful explanations for choice behavior. Loss-aversion and cooperation, oddities from the rational-choice theorist point of view, now make sense as patterns of inference and behavior that follow naturally from the affective dynamics of decision-makers. If rationality is a normative, evaluative concept, one might argue that a theory of decision-making mechanisms can hardly draw a distinction between rational and irrational decisions. Or can it?

Classical models of rationality in fact suffer from the exact same problem. As Sen observed, every choice is said to reveal preferences:

> *The reduction of man to a self-seeking animal depends in this approach on careful definition. If you are observed to choose x rejecting y, you are declared to have "revealed" a preference for x over y. Your personal utility is then defined as simply a numerical representation of this "preference," assigning a higher utility to a "preferred" alternative. With this set of definitions you can hardly escape maximizing your own utility (…). (Sen 322)*

Philosophical theories of rationality that consider rational all those actions explained by beliefs and desires also face this problem. Whatever theory one might choose, a decision can always be described as the product of practical reasoning, whereby the agent infers a

choice from her beliefs and her desires. This problem is not particular to neuroeconomics.

In order to escape a tautological account of rationality, it is important to distinguish *external* and *internal* assessment of rationality. An external assessment of rationality is an evaluation of the effectiveness of a rule or procedure. It assesses the optimality of rules in achieving certain goals. An internal assessment of rationality is an evaluation of the coherence of intentions, actions and plans. An action can hence be rational from the first perspective but not from the second one, and vice versa. Gambling, for instance, can make sense only from the gambler's point of view: despite knowing that the odds are against her (thus externally irrational, if the goal is money maximization), she trades losses (decision utility) against the pleasurable feeling of gambling (experienced utility).

The same reasoning can also apply to neural mechanisms discovered by neuroeconomics. Once precise mechanisms are spelled out, their optimality can be evaluated (external perspective) and their coherence made explicit (internal perspective). From an external perspective, TD algorithms plausibly represent decision-making processes (Montague), both from a descriptive and a normative point of view. Descriptively, TD models closely mimic human, monkey and bee choices (Montague et al.; Glimcher, Dorris and Bayer; Egelman, Person and Montague). Normatively, these models achieve good performance in tasks that require information-processing under uncertainty, such as Backgammon, spatial delayed response tasks and autonomous robot navigation (Suri and Schultz; Perez-Uribe). In the same vein, making a fair offer in the ultimatum game is also a good strategy: offering a $5/$5 split entails a sure gain of $5, while offering a $1/$9 entails a sure loss. Therefore, from the external perspective, the

optimality of neural mechanisms can be assessed, and careful studies can draw a distinction whether, in a particular environment, such-and-such mechanism is effective. TD-learning, for instance, is effective under radical uncertainty, but is beaten by classical rationality under risk.

From an internal perspective, neuroeconomics studies show how externally irrational choices can be internally rational. Loss-aversion and ambiguity-aversion are not optimal under decision theory, but neuroeconomics shows that loss- and ambiguity-averse subjects are rational in the internal sense, because they try to minimize negative feeling (elicited by amygdala activation). In this sense, it can be rational to maximize experienced utility, based on predicted utility. Loss-aversion can also be internally rational because preferring a sure decision is a way to minimize the painful feeling of loss and the post-decision feeling of regret. Various studies indicate that a loss of money can elicit negative emotional response (Delgado, Labouliere and Phelps), and that avoiding an aversive event can generate rewarding signals in the orbitofrontal cortex (Kim, Shimojo and Doherty). A loss-averse agent maximizes experienced utility before, during and after decision. Similarly, players in the ultimatum game and prisoner's dilemma also behave rationally: they maximize positive feelings by being fair and minimize negative ones by avoiding or punishing unfairness. A responder (in the ultimatum game) accepting an unfair split would experience moral disgust, and would be happier from punishing the unfair proposer than from accepting the offer. One of the strengths of neuroeconomics is a more accurate description of subjects' preferences: instead of inferring preferences solely from behavior, imaging and lesion studies allow inferences from brain functioning.

Are agents hence always internally rational, and always have good reasons (maximizing positive feelings, minimizing negative ones)? As Davidson suggested, accounting for irrationality requires partitioning the mind into subparts with a high degree of internal structures, potentially causing genuine internal conflict between values, judgments or motivations (Davidson "Problems of Rationality"). Explanations of practical irrationality such as *akrasia* (not acting on one's own best judgment about what to do) reveal conflicting motivations: I want to finish my homework, I know it is the best thing to do, but instead I watch *Casablanca* on TV, even though I know it is not the best decision. With the distributed account of utility, one might identify several ways to be irrational. A perfectly rational agent would be an agent whose *decision*, *predicted*, *experienced* and *remembered* utility would always be in agreement. Any mismatch between these types of utility could be a variety of irrationality, and neuroeconomics could describe the interactions between different utility-maximization mechanisms. When decision utility is higher than predicted utility, for instance, some mechanisms inside the agent value the outcome of the decision more highly than others that predict an inferior utility. Drug addiction works exactly along the lines of this 'irrational pursuit' (Berridge): dopaminergic neurons 'hijacked' by drugs overvalue the drug rewarding effect, even if prefrontal areas involved in affective prediction do not anticipate a positive hedonic effect. This interaction betweens different processes may explains why addiction often seems internally irrational (its external irrationality is more obvious). John Cheever's narration of his alcoholism in his *Journals* illustrates the incoherence between decision and utility, both predicted and experienced: "*[d]rink, its implements, environments, and effects all seem disgusting. And yet each noon I reach for the whiskey bottle* (Cheever 54). In his case, although the decision utility of drinking does

17

not coincide with its predicted and experienced utility, these two types of utility are here in agreement: Cheever both anticipates and experiences disgust. In other cases, we are irrational when predicted and experienced utility mismatch. Gilbert and Wilson coined the term 'miswanting' to describe the shortcomings of affective forecasting (the prediction of emotional impact). A perfectly rational agent, at time t, would choose to do X at $t+1$ given what she expects her future valuation of X to be. Many studies showed instead that people are subject to many biases that cause miswanting. For instance, people overestimate the length or intensity of the emotional impact of negative events: they recover from tragic incidents faster than they anticipated, and get used to pleasant effects. For instance, almost everybody imagines that being richer implies being happier. Research in psychology indicates, to the contrary, that

> *[p]eople with above-average income are relatively satisfied with their lives but are barely happier than others in moment-to-moment experience, tend to be more tense, and do not spend more time in particularly enjoyable activities. (Kahneman et al. 1908)*

As of today, neuroeconomic research does not explain all these prediction biases, but research progresses in this direction. For now, we know that neural circuits involved in affective forecasting are not the same as those involved in experienced utility (Berridge). A thorough neuroeconomic explanation of miswanting would show how certain mechanisms bias affective forecasting and lead to disappointment. As a matter of a fact, neuroeconomics follow the Davidsonian conception of irrationality: partitioning the mind in sub-personal circuits (Ross 45)

4. Conclusion

Philosophy, economics and psychology have already analyzed and enriched our conception of decision-making. Neuroeconomics, while still in its infancy, has already started to produce empirical results relevant to philosophy of science, philosophy of mind, moral philosophy and other theoretical approaches of rationality. It can enrich the debate on the relevance of folk-psychological categories such as belief and desire; it challenges the idea that there cannot be a science of rationality; it proposes a causal-mechanistic account of decision-making; it shows that judgments of value are highly affective; and finally, it is a first step toward a naturalistic account of rationality. By breaking down utility-maximization into subspecies (experienced, decision, predicted, remembered), and by identifying the brain processes and circuitry that maximize these types of utilities, neuroeconomics offers a finer description of decision-making than rational-choice theory or psychology alone. Neuroeconomics is a fertile 'trading zone' (Thagard) for interdisciplinary exchange between economics, neuroscience, psychology, anthropology, and computer science in the years to come.

Acknowledgement

References

Andreoni, James. 'Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving.' *The Economic Journal* 100.401 (1990): 464-77.

Audi, Robert. *The Architecture of Reason : The Structure and Substance of Rationality.* Oxford ; New York: Oxford University Press, 2001.

Baron, Jonathan. *Thinking and Deciding.* 3rd ed. Cambridge, UK ; New York: Cambridge University Press, 2000.

Berridge, Kent. C. 'Irrational Pursuits: Hyper-Incentives from a Visceral Brain.' *The Psychology of Economic Decisions.* Eds. I. Brocas and J. Carrillo. Oxford: Oxford University Press, 2003. 17-40.

Camerer, Colin. F. 'Psychology and Economics. Strategizing in the Brain.' *Science* 300.5626 (2003): 1673-5.

Cheever, John. 'Journals.' *The New Yorker* 1990.

Clark, Andy. *Being There : Putting Brain, Body, and World Together Again.* Cambridge, Mass.: MIT Press, 1997.

D'Esposito, Mark. 'Functional imaging of neurocognition.' *Seminars in Neurology* 20.4 (2000): 487-98.

Davidson, Donald. *Essays on Actions and Events.* Oxford: Oxford University Press, 1980.

---. ' Paradoxes of Irrationality.' *Problems of Rationality.* Oxford: Oxford University Press, 2004. 169-88.

de Quervain, Dominique. J., et al. 'The Neural Basis of Altruistic Punishment.' *Science* 305.5688 (2004): 1254-8.

Delgado, Mauricio. R., Christa. D. Labouliere, and Elizabeth. A. Phelps. 'Fear of Losing
Money? Aversive Conditioning with Secondary Reinforcers.' *Social Cognitive and
Affective Neuroscience* 1.3 (2006): 250-59.

Egelman, David. M., Christophe Person, and Read Montague. 'A Computational Role for
Dopamine Delivery in Human Decision-Making.' *J Cogn Neurosci* 10.5 (1998):
623-30.

Eliasmith, Chris. 'Computational Neuroscience.' *Philosophy of Psychology and
Cognitive Science, Handbook of Philosophy of Science*. Ed. Paul Thagard. Vol. 4.
Amsterdam: Elsevier, 2007. 313-38.

Gilbert, Daniel. T., and Timothy. D. Wilson. 'Miswanting: Some Problems in the
Forecasting of Future Affective States.' *Feeling and thinking: The role of affect in
social cognition* (2000): 178–97.

Glimcher, Paul W. *Decisions, Uncertainty, and the Brain : The Science of
Neuroeconomics*. Cambridge, Mass. ; London: MIT Press, 2003.

Glimcher, Paul W., Michael C. Dorris, and Hannah M. Bayer. 'Physiological Utility
Theory and the Neuroeconomics of Choice.' *Games and Economic Behavior* 52.2
(2005): 213.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 'An Experimental Analysis of
Ultimatum Bargaining.' *Journal of Economic Behavior and Organization* 3.4
(1982): 367-88.

Huettel, Scott A., et al. 'Neural Signatures of Economic Preferences for Risk and
Ambiguity.' *Neuron* 49.5 (2006): 765-75.

Kahneman, Daniel. 'A Perspective on Judgment and Choice: Mapping Bounded
Rationality.' *America Psychologist* 58.9 (2003): 697-720.

Kahneman, Daniel., et al. 'Would You Be Happier If You Were Richer? A Focusing Illusion.' *Science* 312.5782 (2006): 1908-10.

Kahneman, D., Paul Slovic, and Amos. Tversky, eds. *Judgment under Uncertainty : Heuristics and Biases*. Cambridge ; New York: Cambridge University Press, 1982.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 'Back to Bentham? Explorations of Experienced Utility.' *The Quarterly Journal of Economics* 112.2 (1997): 375-405.

Kim, Hackjin, Shinsuke Shimojo, and John P. Doherty. 'Is Avoiding an Aversive Outcome Rewarding? Neural Substrates of Avoidance Learning in the Human Brain.' *PLoS Biology* 4.8 (2006): e233.

Ledyard, John O. 'Public Goods: A Survey of Experimental Research.' *Handbook of Experimental Economics*. Eds. Jonh H. Kagel and Alvin E. Roth: Princeton University Press, 1995. 111-94.

McCabe, Kevin. 'Neuroeconomics.' *Encyclopedia of Cognitive Science*. Ed. Lynn Nadel. New York, N.Y.: Wiley InterScience, 2005. 294-98.

Moll, Jorge, et al. 'Human Fronto-Mesolimbic Networks Guide Decisions About Charitable Donation.' *PNAS* 103.42 (2006): 15623-28.

Montague, Read., Brooks King-Casas, and Jonathan D. Cohen. 'Imaging Valuation Models in Human Choice.' *Annual Review of Neuroscience* 29 (2006): 417-48.

Montague, Read, and Gregory S. Berns. 'Neural Economics and the Biological Substrates of Valuation.' *Neuron* 36.2 (2002): 265-84.

Montague, P. Read, et al. 'Bee Foraging in Uncertain Environments Using Predictive Hebbian Learning.' *Nature* 377.6551 (1995): 725-28.

Montague, Read. *Why Choose This Book? : How We Make Decisions*. New York:

Penguin Group, 2006.

Naqvi, Nasir, Baba Shiv, and Antoine Bechara. 'The Role of Emotion in Decision Making:

A Cognitive Neuroscience Perspective.' *Current Directions in Psychological*

Science 15.5 (2006): 260-64.

Niv, Yael., Michale O. Duff, and Peter Dayan. 'Dopamine, Uncertainty and Td Learning.'

Behavioral ad Brain Functions 1 (2005): 6.

Oosterbeek, Hessel, Randolph Sloof, and Gijs van de Kuilen. 'Differences in Ultimatum

Game Experiments: Evidence from a Meta-Analysis.' *Experimental Economics* 7

(2004): 171-88.

Perez-Uribe, Andres. *Using a Time-Delay Actor-Critic Neural Architecture with*

Dopamine-Like Reinforcement Signal for Learning in Autonomous Robots.

Lecture Notes in Computer Science, 2001.

Poundstone, William. *Prisoner's Dilemma*. 1st ed. New York: Doubleday, 1992.

Rilling, James K., et al. 'The Neural Correlates of Theory of Mind within Interpersonal

Interactions.' *Neuroimage* 22.4 (2004): 1694-703.

Rilling, James K., et al. 'A Neural Basis for Social Cooperation.' *Neuron* 35.2 (2002):

395-405.

Robbins, Lionel. *An Essay on the Nature and Signifiance of Economic Science*. London

Macmillan, 1932.

Ross, Don. 'The Economics of the Sub-Personal: Two Research Programs.' *Economics*

and the Mind. Eds. Barbara Montero and Mark D. White. New York: Routledge,

2007. 41-57.

Ross, Don. *Economic Theory and Cognitive Science : Microexplanation*. Cambridge,

    Mass.: MIT Press, 2005.

Sally, David. 'Conversations and Cooperation in Social Dilemmas: A Meta-Analysis of

    Experiments from 1958 to 1992.' *Rationality and Society* 7 (1995): 58 – 92

Samuelson, Larry. 'Economic Theory and Experimental Economics.' *Journal of*

    Economic Literature 43 (2005): 65-107.

Sanfey, Alan. G., et al. 'The Neural Basis of Economic Decision-Making in the Ultimatum

    Game.' *Science* 300.5626 (2003): 1755-8.

Searle, John. *Rationality in Action*. The Jean Nicod Lectures. Cambridge, Mass.: MIT

    Press, 2001.

Selten, Reinhard. 'What Is Bounded Rationality ? .' *Bounded Rationality: The Adaptive*

    Toolbox. Eds. G. Gigerenzer and R. Selten. MIT Press: Cambridge, MA, 2001. 13-

    36.

Sen, Amartya K. 'Rational Fools: A Critique of the Behavioral Foundations of Economic

    Theory.' *Philosophy and Public Affairs* 6.4 (1977): 317-44.

Simon, Herbert A. *Models of Bounded Rationality*. Cambridge, Mass.: MIT Press, 1982.

Sterelny, Kim. 'Cognitive Load and Human Decision, or, Three Ways of Rolling the Rock

    up Hill ' *The Innate Mind : Culture and Cognition*. Eds. Peter Carruthers,

    Stephen Laurence and Stephen P. Stich. New York: Oxford University Press,

    2006.

Suri, Roland E., and Wolfram Schultz. 'Temporal Difference Model Reproduces

    Anticipatory Neural Activity.' *Neural Computation* 13.4 (2001): 841-62.

Suri, Roland E., and Wolfram Schultz. 'A Neural Network Model with Dopamine-Like
Reinforcement Signal That Learns a Spatial Delayed Response Task.'
Neuroscience 91.3 (1999): 871.

Sutton, Richard S., and Andrew G. Barto. 'A Temporal-Difference Model of Classical
Conditioning.' *Ninth Annual Conference of the Cognitive Science Society*, 1987.
355-78.

---. *Reinforcement Learning : An Introduction.* Adaptive Computation and Machine
Learning. Cambridge, Mass.: MIT Press, 1998.

Thagard, P. 'Being Interdisciplinary: Trading Zones in Cognitive Science.'
Interdisciplinary Collaboration: An Emerging Cognitive Science. Eds. Sharon J.
Derry, Christian Schunn, D. and Morton Ann Gernsbacher, 2005. 317-39.

Tom, Sabrina, et al. 'The Neural Basis of Loss Aversion in Decision-Making under Risk.'
Science 315.5811 (2007): 515-8.

Tversky, Amos. 'A Critique of Expected Utility Theory: Descriptive and Normative
Considerations.' *Erkenntnis* V9.2 (1975): 163-73.

Tversky, Amos, and Daniel. Kahneman. 'Loss Aversion in Riskless Choice: A Reference-
Dependent Model.' *The Quarterly Journal of Economics* 106.4 (1991): 1039-61.

Zak, Paul. J. 'Neuroeconomics.' *Philosophical Transactions of the Royal Society of
London, Series B, Biological Science* 359.1451 (2004): 1737-48.