Review TRENDS in Cognitive Sciences Vol.10 No.3 March 2006

Neuroeconomics: cross-currents in research on decision-making

Alan G. Sanfey[1], George Loewenstein[2], Samuel M. McClure[3] and Jonathan D. Cohen[3]

[1]Department of Psychology, University of Arizona, Tucson, AZ 85721, USA
[2]Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[3]Department of Psychology, and Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, NJ 08544, USA

Despite substantial advances, the question of how we make decisions and judgments continues to pose important challenges for scientific research. Historically, different disciplines have approached this problem using different techniques and assumptions, with few unifying efforts made. However, the field of neuroeconomics has recently emerged as an inter-disciplinary effort to bridge this gap. Research in neuroscience and psychology has begun to investigate neural bases of decision predictability and value, central parameters in the economic theory of expected utility. Economics, in turn, is being increasingly influenced by a multiple-systems approach to decision-making, a perspective strongly rooted in psychology and neuroscience. The integration of these disparate theoretical approaches and methodologies offers exciting potential for the construction of more accurate models of decision-making.

Introduction

The question of how we make, and how we should make, judgments and decisions has occupied thinkers for many centuries, with different disciplines approaching the problem with characteristically different techniques. A very recent approach, popularly known as neuroeconomics, has sought to integrate ideas from the fields of psychology, neuroscience and economics in an effort to specify more accurate models of choice and decision (for reviews from the perspective of economics, see [1,2]).

How profitable the neuroeconomic approach will be is still unclear. Predictably, there are strong opinions on both sides. On the one hand, its strongest advocates (aided by some exaggerated media reporting) have presented neuroeconomics as a new paradigm that will eventually replace the classical approaches. On the other, skeptics in both communities have argued that economic models and neuroscientific techniques reflect disparate levels of analysis that have little to offer one another. Economists have, historically, been skeptical of the ability of 'process measures' to contribute to our understanding of economic and social behavior [3]; and neuroscientists commonly view economics as too abstract and removed from the mechanisms of interest in the brain.

Although we are perhaps not as optimistic as the most ardent believers in neuroeconomics when it comes to the time-line of progress, we do believe that the field has real potential for making important contributions to our understanding of decision-making, above and beyond what has and will continue to be learned from work within each discipline independently. This is because neuroeconomics is able to draw upon the complementary strengths of its contributing disciplines. In fact, the benefits of increasing contact between neuroscience, psychology and economics are already apparent.

The central argument of this article is that economics, psychology and neuroscience can each benefit from taking account of the insights that the other disciplines have to offer in understanding human decision-making. In the following, we first address how neuroscience can, and already has, benefited from economics' unitary perspective. We then discuss how economics can, and has begun to, be enriched by taking account of cooperation and competition between multiple specialized neural systems, before closing with some thoughts on potentially fruitful future research directions.

The view from economics: one unified theory

Economics contributes to the joint endeavor of neuroeconomics by bringing its insights into the diverse outcomes that can arise from the strategic and market interactions of multiple agents, and through a set of precise, formal, mathematical models to describe these interations and outcomes. However, the aspect of economics that may prove most useful to neuroscientists (and, indeed, that has already begun to bear fruit) is its embracing of a unified theoretical framework for understanding human behavior – namely the idea that behavior can be interpreted as choosing alternatives with the goal of maximizing utility.

The unitary perspective of economics can be seen in the assumptions that it makes about the two essential dimensions of decision-making: *choice* (the evaluation of options and selection of actions), where economics assumes a consistent, stable set of preferences; and *judgment* (information processing and probability estimation), with the assumption of a general reasoning system applicable to a wide range of problems. These assumptions have been criticized, as will be discussed later, but the concept of decisions being made by comparing the utility signals for each of the decision

---

Corresponding author: Sanfey, A.G. (asanfey@u.arizona.edu).

www.sciencedirect.com 1364-6613/$ - see front matter © 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.tics.2006.01.009

alternatives has led to some real developments and has played an increasingly important role in guiding research investigating the underlying brain mechanisms. Additionally, it is possible that well-established ideas from economics will shed light on one of the least developed, but most important, riddles for neuroscience: how the multiple, diverse and specialized neural systems of the brain coordinate their activities to solve complex and often novel problems and give rise to coherent, goal-directed behavior (see Box 1).

How economics can inform neuroscience: the benefits of a unitary perspective

In recent years, a growing number of neuroscientists have turned to economic models as a framework both for interpreting results and guiding new experimental work (e.g. [4,5]). This follows a small, but growing, tradition in neuroscience in which optimal performance is defined for a given behavioral domain, and is then used for constructing theories about underlying neural function. Even as it is recognized that the brain (and consequent behavior) does not operate perfectly optimally, there are several reasons why these assumptions can nevertheless be valuable. First, although complex forms of behavior might not be optimal, simpler evolutionarily conserved mechanisms might prove to be closer to optimal, or at least to have been so in the environment in which they evolved. Second, an assumption of optimality can be a crucial step in the development of formal theory, as it is often easiest to define and precisely characterize the optimal behavior of a system. Formal theory, in turn, enables the generation of precise, testable predictions about the system's behavior. Finally, even when behavior (or neural function) turns out to be suboptimal, defining optimal performance can provide a useful benchmark against which to compare actual behavior. Identifying ways in which behavior systematically deviates from optimality can then generate new insights into underlying mechanisms. The use of the Expected Utility (EU) model [6] is one example of this approach, and has been productively applied to research on the neural bases of reward and decision-making.

According to the classical EU model, utility is computed as the product of the *value* and the *probability* of each potential outcome (see Figure 1). Using this as a starting point, research has sought to distinguish and identify the neural substrates for each of these components and to study their interaction. Although this is still a relatively new direction in neuroscience research, some useful progress has been made.

Box 1. Brain systems and economics

One potential area where economics can contribute is in understanding the dynamic processes by which the brain coordinates its diverse systems to perform new, complex tasks. This problem has received remarkably little attention in neuroscience research, although it is well-explored terrain in economics [76,77].

There are striking parallels between the brain and a modern corporation. Both can be viewed as complex systems transforming inputs into outputs. Both involve the interaction of multiple, highly similar, agents (neurons are similar to one another, just as are people), which, however, are specialized to perform particular functions. Thus, in corporations, units often take the form of departments that perform functions such as research, marketing, and so on. Similarly, the brain has systems specialized for different functions. As in a corporation, these functions may be more or less spatially segregated in the brain, depending upon the processing requirements of the specific functions and their interactions.

Furthermore, there is hierarchical structure in both brains and corporations. Both rely on 'executive' systems that make judgments about the relative importance of tasks and decide how to mobilize specialized capabilities to perform those tasks. Several neuroimaging studies have demonstrated that brain areas involved in executive function are actively engaged during the performance of a novel and demanding task, and then show progressively less activity as the task becomes more automatic [78,79]. These changes closely parallel the increase in activity in more specialized brain areas, as well as increases in performance speed. Similar improvements in speed and efficiency are observed in industry. Studies of shipbuilding during World War II found that labor productivity rose at an annual rate of 40 percent, and that speed of production increased even more rapidly. Although it often took as much as 300 days for a yard to deliver its first ship, by 1943 delivery times were often less than one month and even as short as five days [80]. Presumably such changes were accompanied by diminished reliance on administrative and executive involvement, as is observed in the brain.

It remains to be determined whether these similarities go deeper than mere analogy, and reflect fundamental principles of aggregate behavior that apply across different levels of analysis. However, whether or not the principles of interaction are the same, the theoretical methods that have been developed within economics to analyze and construct models of interactions among economic units, such as principal-agent theory, general equilibrium theory, and the analysis of the firm, are likely to be useful in studying the interactions among brain systems that determine individual behavior.

Neural basis of value and reward

Several decades of research have focused on the neural bases of reward and punishment (the *value* function of the EU model), identifying several systems that are consistently responsive to reinforcement. However, we still have a relatively weak understanding of the underlying neural computations and their influence on decision-making. Early studies of brain reward systems were concerned primarily with establishing the generality of their function and their neurochemical bases. These experiments demonstrated that animals will withstand electric shock, exert significant physical effort, and even reduce food intake to obtain electrical stimulation in appropriate brain areas [7]. Furthermore, the reward value of stimulation was shown to be reduced when dopamine receptor binding was inhibited, particularly in the ventral striatum [8].

Inspired at least in part by an awareness of economic theory, more recent experiments have recognized the importance of distinguishing between the magnitude and the probability of reward, and have tested directly for their separate influence on reward-related brain activity. Single-cell recordings from dopamine neurons [9] and neurons in the orbitofrontal cortex [10,11], striatum [12] and posterior cingulate cortex [13] have shown that neural responses scale reliably with reward magnitude. These results have also been observed in human subjects: activity changes in many of these same brain areas scale directly with the magnitude of earned monetary reward [14–17], the appetitive value of food reward [18,19], as well

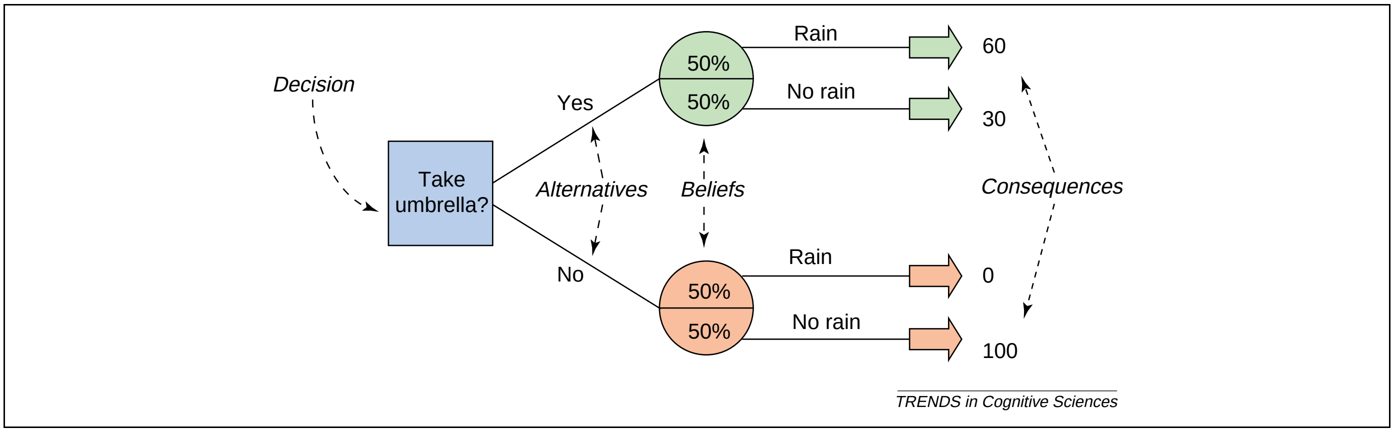

Figure 1. Subjective Expected Utility theory posits two fundamental characteristics of an alternative that must be combined before reaching a decision, namely the value of the alternative and the probability that this value will be attained. The value of each alternative is weighted by its attendant probability, and the option with the highest probability-weighted values (or 'utility') is then chosen. An illustrative example is shown above. Imagine you are trying to decide whether or not to bring an umbrella as you walk to work one morning. Utility theory posits that we must weigh the *probability* (in this case set at $p=0.5$, perhaps gleaned from previous experience or from the morning weather forecast), with the *value* of each of the four possible outcomes. These four outcomes are assigned a subjective rating (0=worst outcome, 100=best outcome), which of course varies depending on the individual decision-maker's preferences, and a calculation is performed multiplying each outcome by its relevant probability, and then combining across alternatives. Therefore, in this example, the utility for the option of taking the umbrella is $[(0.5*60)+(0.5*30)]=45$, whereas the utility for not taking one's umbrella is $[(0.5*100)+(0.5*0)]=50$. The utility maximizing decision-maker should therefore opt to leave the umbrella behind in these circumstances.

as with more abstract, social, rewards [20]. Furthermore, brain stimulation experiments offer a precise mechanism for investigating how reward magnitude might be encoded, suggesting that value is encoded as the integral of excitatory inputs to reward-related brain areas [21].

There is mounting evidence that the mesencephalic dopamine system plays a crucial role in value assessment by signaling errors in reward prediction, which are used to augment reward-producing behaviors both by generating learning signals [22], and by adaptively updating goal states and attentional focus in working memory [23]. Similarly, it has been suggested that the norepinephrine (noradrenaline) system regulates the balance between exploitation (seeking to maximize utility from a given source of reward) and exploration (seeking new sources) through its influence on mechanisms of learning and attention [24,25]. Although computational models have been developed to describe functions of the dopamine and norepinephrine systems, their interaction remains relatively unexplored. Neuroscientists could benefit from interacting with economists who have been empirically investigating similar issues, and are developing theoretical models that provide useful frameworks for integrating these different lines of neuroscientific research (see [26] on learning from feedback and [27] on the tradeoff between exploitation and exploration).

Although most neuroimaging work to date has focused on positive rewards, recent findings also suggest that there are complementary mechanisms for evaluating negative utility. For example, there is evidence that the anterior cingulate cortex (ACC) responds to a variety of signals that indicate negative utility (e.g. performance costs such as response conflict, errors, negative feedback and pain [28–34]). Across these results, there is evidence that increased activity in ACC correlates with the magnitude of anticipated consequences, and current efforts are beginning to test the extent to which these relationships conform to predictions of economic theory (such as Prospect Theory [35]). For example, one important finding from behavioral economics is that people evaluate the outcome of decisions based on a flexible reference point. This predicts that the neural systems responsible for utility assessment should be responsive to relative gains and losses rather than to absolute levels of rewarding and punishing stimuli. Recent studies, using both fMRI [36] and scalp electrical recordings [37] have provided support for this prediction.

Neural basis of probability estimation

Probability has been less well studied than reward, although several recent neuroimaging studies have begun to address this aspect. For example, one study has shown that activity in medial prefrontal cortex is inversely related to the probability of obtaining monetary reward [38]. Other research has observed that activity in (human) brain reward areas depends on uncertainty in the timing of reward delivery [39].

Neural basis of the utility signal

Studies have also begun to examine the interaction between value and probability in the computation of utility and the execution of decision-making behavior. This work builds on the recent discovery of brain areas that appear to be directly related to decision-making, including the lateral intraparietal area (LIP) and regions of frontal cortex. Neurons in these areas closely track the dynamics of decision-making in simple two alternative forced choice tasks: response-selective neurons exhibit a progressive increase in firing rate following stimulus presentation, and the time at which they cross a threshold of activity predicts the timing of the behavioral response. This has been interpreted as evidence that these neurons integrate choice-relevant information and implement a fundamental decision-making mechanism [40–43] that approximates the optimal algorithm [44–46]. These neurons are also directly sensitive to manipulations of utility, exhibiting combined effects of value and probability in a manner predicted by the Expected Utility model [47]. Additionally, there have been human imaging studies testing for activity that scales directly with expected utility, manipulating both reward magnitude and probability [36,38]. This research has found that activation in striatal reward areas correlates with expected utility, demonstrating that utility-like measures are found outside motor preparation areas such as LIP.

The investigations described above reflect the influence that economic theory is already having on neuroscience studies. An independent metric for assessing reward value is crucial for a system that must often choose between rewards delivered in different modalities (e.g. do we take extra vacation time or an additional paycheck?), and the neural systems outlined above may well provide the basis for such a signal.

The view from neuroscience: multiple systems

Psychology and neuroscience also bring much to the neuroeconomics table, contributing a rich tradition of empirical research and increasingly precise methods for studying behavior and the neural mechanisms by which it is governed. Of particular relevance to economics is the growing insight into mechanisms that are responsible for the assessment of utility and execution of decision-making behavior, as outlined in the previous section. However, perhaps the single most important perspective that neuroscience brings is to challenge the core assumption in economics that behavior can be understood in terms of unitary evaluative and decision-making systems.

How neuroscience can inform economics: the benefits of a multiple-system approach

As noted in the first section of this article, economic theory assumes that people choose between alternative courses of action based on a rational evaluation of the consequences, and economists have developed detailed theoretical models for dealing with many decision situations: for example, the Expected Utility model (described above) for decisions under risk, and the Discounted Utility model for decisions with consequences spread over time. These models have the virtue that they are formally explicit, analytically tractable, and can be used to make quantitatively precise predictions about decision-making in a wide variety of circumstances. As such, they have provided a strong and unifying foundation for the development of theory about decision-making, with an assumption that decisions reflect the operation of a unitary all-purpose information processor.

However, psychological research on judgment and decision-making has produced a wealth of evidence demonstrating that, in practice, these models do not provide a satisfactory description of human behavior [48]. There is a long legacy of research within psychology, strongly supported by findings from neuroscience, to suggest that human behavior is not the product of a single process, but rather reflects the interaction of different specialized subsystems. Although most of the time these systems interact synergistically to determine behavior, at times they compete, producing different dispositions towards the same information. A major cause of these observed idiosyncrasies of behavior that have been used to challenge the standard economic model might be that these decisions do not emerge from a unitary process, but rather from interactions between distinguishable sets of processes.

Multiple processes in decision-making

The most general distinction, and the most important for neuroeconomics, is one that psychologists make between automatic and controlled processes [49,50]. Automatic processes are fast and efficient, can often be carried out in parallel, but are highly specialized for domain-specific operations and therefore relatively inflexible. They are thought to reflect the operation of highly over-trained (and, in some cases, possibly ‘hardwired’) mechanisms. However, humans also have a capability for controlled processing underlying our higher cognitive faculties. Controlled processes are highly flexible, and thus able to support a wide variety of goals, but are relatively slow to engage and rely on limited capacity mechanisms – that is, they can support only a small number of pursuits at a time. Furthermore, the operations involved in controlled processes (such as reasoning) are often accessible to introspective, explicit description, whereas those involved in automatic processes (such as recognizing a face) are usually much less so. An example commonly given of the distinction between these processes is driving a stick-shift car: the novice is thought to rely on controlled processing, requiring focused concentration on a sequence of operations that can be articulated but that are effortful and easily disrupted by distraction. By contrast, the well-practiced driver, relying on automatic processes, can carry out the same task efficiently while engaged in other activities (e.g. conversing), but might no longer be able to articulate clearly the individual operations involved.

The distinction between controlled and automatic processing is probably best thought of as a continuum, rather than a qualitative dichotomy [51,52]. Nevertheless, it has proven extremely useful in characterizing the dynamics of behavior involving competing processes. This distinction has appeared prominently in the field of decision research, where investigators refer to two types of evaluative systems: System 1 and System 2 [53,54]. System 1 is automatic and heuristic-based; quickly proposing intuitive answers to problems as they arise. System 2, which corresponds closely with controlled processes, monitors the quality of the answer provided by System 1 and sometimes corrects or overrides these judgments.

Behavioral evidence for multiple systems

The description of System 2 bears a close resemblance to the rational, general-purpose processor presupposed by standard economic theory. However, a bounty of experimental findings suggest that controlled processing accounts for only part of our overall behavioral repertoire, and in some circumstances can face stiff competition from domain-specific automatic processes.

There is now an extensive literature describing the many ways in which human decision-making violates the principles of rationality as defined by the Expected Utility (EU) [55] and Discounted Utility (DU) [56] models. Early work in economics revealed situations (e.g. Ellsberg and Allais paradoxes) whereby behavior violated key axioms of the EU model. More recently, the 'Heuristics and Biases' approach in psychology has documented many instances of deviations from economic rationality [57]. For example, most people are reluctant to play a gamble with 50% chance of winning $25 and 50% chance of losing $20, despite the gamble's overall positive expected value. This illustrates the phenomena of 'loss aversion', whereby people often place disproportionate weight on losses relative to gains of similar absolute value [58].

When it comes to choice over time, there is also ample evidence of violations of the DU model. Perhaps most importantly, there is strong evidence that discounting is much steeper for short time delays than for longer delays, a phenomenon known as 'hyperbolic time discounting' [59]. For example, offered a choice between $10 today and $11 in a week, many people are likely to choose the immediate $10. However, offered the choice between $10 in a year and $11 in a year and a week, most people would chose the $11, now considering the extra week of wait inconsequential. From the economist's perspective, however, this implies a reversal of preference (i.e. whether an extra dollar is worth a week's wait or not), and therefore does not conform to the rational model [56].

Responding to the documented limitations of EU and DU, a variety of alternative models of decision-making under uncertainty and intertemporal choice have been constructed (including hyperbolic time discounting and Prospect Theory). These models have made notable progress in describing real-world decision-making [60], and even in designing policy interventions to change behavior [61].

Additionally, economists have begun to construct models that play on the automatic vs. controlled distinction. For example, Benhabib and Bisin (J. Benhabib and A. Bisin, unpublished) propose a model of decision-making in which automatic processes are initially allowed to determine behavior, but controlled processes are activated whenever the costs from letting the automatic processes carry on become too large. They apply this framework to a dynamic saving-consumption model, and describe how its predictions differ from traditional economic theory. The key insight here is that, rather than thinking about human behavior as being governed by a unitary, general purpose mechanism, it can often be better described in terms of the interaction – and sometimes competition – between different subsystems that might favor different alternatives for a given decision.

However, although these new models of decision-making have made great strides towards providing descriptively realistic models of behavior, as yet they have not provided insight into the actual mechanisms responsible for the deviations from the normative models. Recently, neuroimaging studies have begun to identify the contribution that different systems make to decision-making, in particular when this involves competition between controlled processing and emotional responses – a special case of automatic processes that we consider in the next section.

Neural evidence for multiple systems

With the advent of neuroeconomics, discoveries about the neural mechanisms involved in perception, attention, learning and action selection have begun to drive the development of new, mechanistically explicit models of decision-making. In several instances, where these models have been expressed formally, they have also begun to make contact with economic theory [62]. In some cases, this has provided validation of some of the basic principles of economic theory (see previous section), whereas in others it has begun to provide insight into how and why human behavior deviates from optimality as defined by economic models. Psychological research, buttressed by recent neuroscientific findings, has begun to identify separable systems that contribute to decision-making and behavior, systems that for the most part work cooperatively, but sometimes compete. Recent neuroscientific research has begun to characterize the engagement and disposition of these neural systems under a variety of conditions in which behavior seems to deviate from the expectations of economic theory.

Systems for emotion and deliberation. Perhaps the distinction with the greatest immediate ramifications for economic theories is between systems supporting emotion and those supporting deliberation, which closely parallels the distinction between automatic and controlled processes. The nature of emotions has been the subject of intense inquiry and debate for entire fields of science, a full consideration of which is well beyond the scope of this article. For present purposes, we will use 'emotion' to refer to low-level psychological processes engaged by events that elicit strong valenced and stereotyped behavioral responses (e.g. fear is a response to threatening stimuli that leads to freezing or withdrawal). Accordingly, emotions are rapid, highly automatic responses to specific stimuli or events, well adapted to some circumstances but not to others. These contrast with the capacity for controlled processing discussed above – the ability to respond flexibly to circumstances, to deliberate rationally about the long-term consequences of our behavior, and to plan behavior accordingly.

Although emotional processes, like other automatic processes, share common neural substrates with controlled processes, it is becoming increasingly clear from neuroimaging studies that these different types of processes do involve some distinguishable neural components. There is a general consensus that high-level, deliberative processes, such as problem-solving and planning, consistently engage anterior and dorsolateral regions of prefrontal cortex as well as areas of posterior parietal cortex [63–66]. By contrast, automatic processes appear to rely heavily on more posterior cortical structures, as well as subcortical systems. Emotional processes, in particular, seem reliably to engage a set of structures classically referred to as the limbic system, which includes brainstem reward-processing structures (such as the ventral tegmental area), areas of the midbrain and cortex to which they project (such as the nucleus accumbens and ventromedial frontal, orbitofrontal, and anterior cingulate cortex), and several other areas such as the amygdala and insular cortex [67].

Strategic interactions. The neural mechanisms responsible for deliberation and emotion are clearly closely interrelated (e.g. [23,24,68]). Nevertheless, distinguishing between these could help shed light on many of the most basic patterns uncovered by behavioral economics, such as nonlinear probability weighting and loss aversion, as well as several other behaviors that challenge the standard economic model, including many that involve market and non-market interactions between individuals. An example of the latter is behavior in a well-studied decision task known as the Ultimatum Game. In this game, participants must decide either to accept or reject an offer of money made by a partner, whose task is to divide a sum of money between the two. The twist is that if the offer is rejected, then neither player receives anything. Whereas standard economic models would predict that all non-zero offers should be accepted, on the premise that some money is better than none, in fact about half of the unfair offers are typically rejected. This is the case even when the game involves single (one-shot) confidential interactions, in which bargaining or the establishment of a reputation cannot explain the findings. A neuroimaging study examining the Ultimatum Game [69] found two brain regions that were particularly active when the participant was confronted with an unfair offer, the anterior insula and the dorsolateral prefrontal cortex (dlPFC) (Figure 2). Activation in these areas has been shown to correlate with emotional and deliberative processing, respectively, and it was found that if the insular activation was greater than the dlPFC activation, participants tended to reject the offer, whereas if the dlPFC activation was greater, they tended to accept the offer. This offers neural evidence for a two-system account of decision-making in this task.

The conclusion that strategic interactions between individuals involves an interplay between emotion and deliberation is underlined by research that has observed activation in the caudate nucleus related to satisfaction derived from punishing norm violations [70], and in the ventral striatum related to exhibitions of trust in an investment game [71].

Decisions over time. Hyperbolic time discounting provides another illustration of the benefits of investigating emotion–deliberation interactions. One hypothesis that has been advanced to explain this phenomenon is that it reflects the operation of two fundamentally different mechanisms, one affective, which heavily values the present and steeply discounts all future opportunities, and the other deliberative, which discounts options more consistently across time [72,73]. However, it has been difficult, from behavioral data alone, to provide evidence for separate mechanisms, or to motivate them from purely theoretical considerations. Recently, neuroimaging has produced evidence for this distinction [74], showing that choices involving the option of an immediate reward actively engage the ventral striatum, as well as medial and orbitofrontal areas – areas rich in dopaminergic innervation and, as noted above, consistently associated with the evaluation of reward. These are areas of the brain that appear to have been highly conserved across evolution, and their steep form of discounting might reflect adaptations to the computation of utility that occurred in the absence of neural structures capable of sophisticated planning for the future. By contrast, areas of frontal and parietal cortex commonly associated with more abstract forms of reasoning and planning were consistently involved, independently of when the reward became available (Figure 3). Furthermore, for choices pitting an immediate reward against a greater but later reward, the relative activity in these different areas was closely correlated with the choice made, with greater fronto-parietal activity associated with selection of the deferred but more valuable option, suggesting that this system appears to value rewards in the future in a more even-handed fashion.

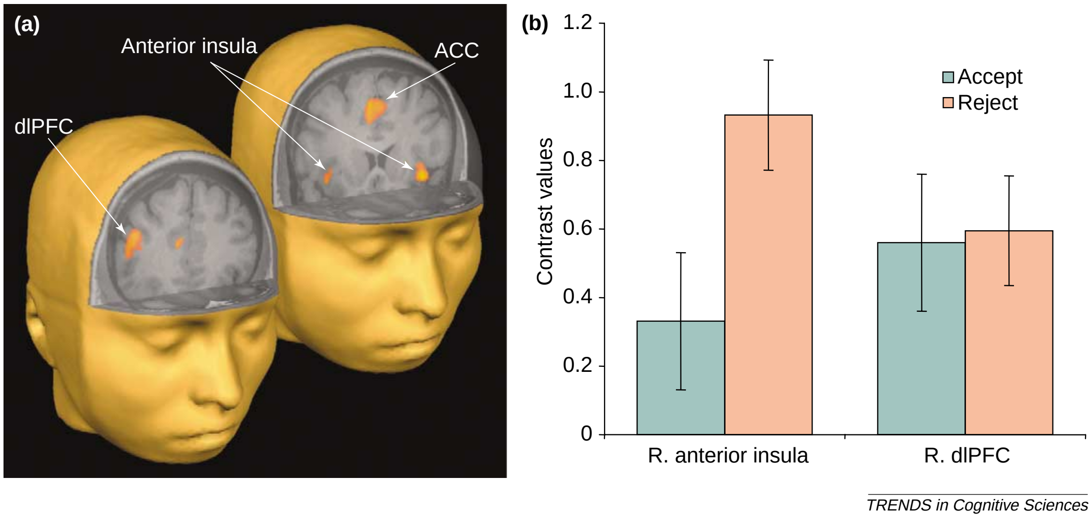

Figure 2. Deciding whether to accept or reject an unfair offer in the Ultimatum Game leads to conflict between emotional, 'reject', and cognitive, 'accept', systems. **(a)** Activation related to the presentation of an unfair offer from another human in the Ultimatum Game, showing activation of bilateral anterior insula and anterior cingulate cortex (ACC, right), and activation of right dorsolateral prefrontal cortex (dlPFC, left). Areas in orange showed greater activation following unfair as compared with fair offers ($P<0.001$). **(b)** Right anterior insula and right dlPFC activation for all unfair offer trials, categorized by subsequent acceptance or rejection. Trials in which the offer was subsequently rejected had significantly higher anterior insula activation than trials where the offer was subsequently accepted. Adapted from [69].

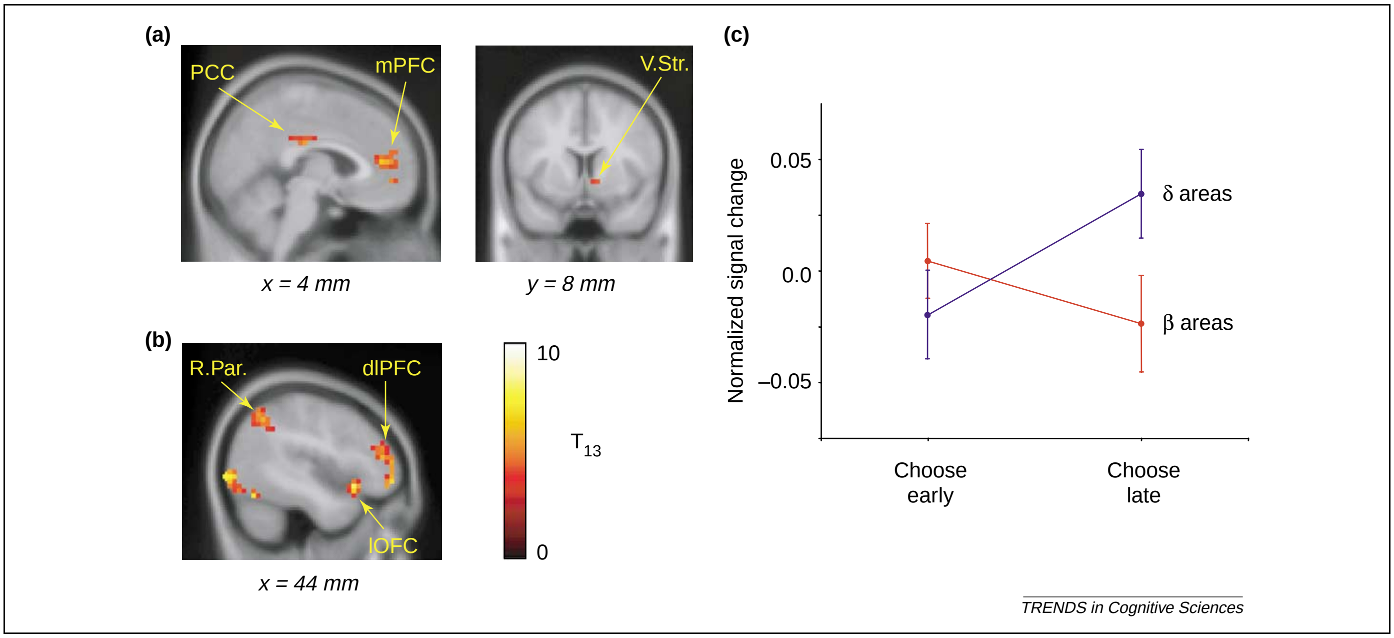

Figure 3. Beta and delta systems in intertemporal choice. When people choose between two rewards separated in time, two different brain systems are involved. **(a)** The first, termed the β system, responds preferentially to the presence of reward available in the immediate future. The β system comprises brain areas associated with the midbrain dopamine system and relatively automatic reward processing, including the ventral striatum (V.Str.) and medial prefrontal cortex (mPFC). **(b)** The second system, the δ system, comprises brain areas implicated in cognition, including the dorsolateral prefrontal cortex (dlPFC) and right posterior parietal cortex (R. Par.). **(c)** When intertemporal choices involve deciding between an immediate and delayed reward, choice tends to reflect the relative activity in the β and δ systems. When people select the later, larger reward, activity tends to predominate in the δ system. Adapted from [74].

Multiple systems; multiple interactions

The above findings provide an initial toehold for measuring physical mechanisms responsible for utility evaluation and decision-making in the brain. Such studies offer the promise that we will be able to identify and precisely characterize these mechanisms, and the factors that influence their engagement and interaction. Even at this early stage, however, results highlight the fact that decision-making involves the interaction among multiple subsystems governed by different parameters and possibly even different principles. As is the case for the distinction between automatic and controlled processes, economists have already begun to propose dual-process models that incorporate a distinction similar to that between emotion and deliberation ([72,75]; Fudenberg, D. and Levine, D. unpublished). These models have mainly been applied to intertemporal choice, but in one case (Loewenstein, G. and O'Donoghue, T. see http://papers.ssrn.com/sol3/papers.cfm?abstract_id=539843) have also been applied to decision-making under uncertainty and to interpersonal interactions of the type examined in the Ultimatum Game.

The future

This article has reviewed two general ways in which the neuroeconomic endeavor can make important contributions to research on decision-making – firstly, the incorporation into neuroscience and psychology of the formal, rigorous economic modeling approach, and secondly, the awareness within the economic community of the evidence for multiple-systems involved in decision-making. One current challenge is to ensure that researchers are communicating productively; often, terms such as 'choice', 'judgment' and 'decision' are used in different ways by different fields. A useful endeavor would be to arrive at a common language, and perhaps a common set of 'decision tasks', to ensure that the hitherto heartening level of collaboration across these diverse fields continues in a productive fashion. With increasing collaboration among researchers from across the fields of interest, it also seems certain that other interesting avenues of research will open up in the near future (see Box 2 for some potential questions of interest).

Box 2. Questions for future research

- How (and where) are value and probability combined in the brain to provide a utility signal? What are the dynamics of this computation?
- What neural systems track classically defined forms of expected and discounted utility? Under what conditions do these computations break down?
- To what extent does tracking utility computations generalize to more complex decision tasks?
- How is the utility metric different from neural signals of attention and motor preparation?
- How is negative utility signaled? Is there a negative utility prediction signal comparable to the one for positive utility?
- How are rewards of different types mapped onto a common neural currency like utility?
- How do systems that seem to be focused on immediate decisions and actions interact with systems involved in longer-term planning (e.g. making a career decision)? For example, does an unmet need generate a tonic and progressively increasing signal (i.e. a mounting 'drive'), or does it manifest as a recurring episodic/phasic signal with increasing amplitude?
- Under what circumstances do these various systems cooperate or compete? When there is competition, how and where is it adjudicated?
- Psychologists, neuroscientists and behavioral economists all seem to agree that various automatic forms of behavior (including emotional responses) reflect the operation of a multiplicity of mechanisms. However, do higher-level deliberative processes rely similarly on multiple mechanisms, or a single, more tightly integrated (unitary) set of mechanisms?

Neuroeconomics is still far from providing a full set of answers to these questions, but it seems reasonable to expect that, as it converges on the mechanisms of utility evaluation and of decision-making, and characterizes these in terms of formal models that can inform economic theory, all of the fields involved stand to gain.

References

1 Glimcher, P.W. and Rustichini, A. (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306, 447–452
2 Camerer, C. *et al*. (2005) Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* 43, 9–64
3 Bruni, L. and Sugden, R. The road not taken: two debates about the role of psychology in economics. *Econ. J.* (in press)
4 Shizgal, P. (1997) Neural basis of utility estimation. *Curr. Opin. Neurobiol.* 7, 198–208
5 Glimcher, P.W. (2003) *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*, MIT Press
6 von Neumann, J. and Morgenstern, O. (1947) *Theory of Games and Economic Behavior*, Princeton University Press
7 Olds, J. (1977) *Drives and Reinforcements: Behavioral Studies of Hypothalamic Function*, Raven Press
8 Colle, L.M. and Wise, R.A. (1988) Effects of nucleus accumbens amphetamine on lateral hypothalamic brain stimulation reward. *Brain Res.* 459, 361–368
9 Tobler, P.N. *et al*. (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645
10 Tremblay, L. and Schultz, W. (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708
11 Roesch, M.R. and Olson, C.R. (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304, 307–310
12 Cromwell, H.C. and Schultz, W. (2003) Effects of expectations for different reward magnitudes on neural activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838
13 McCoy, A.N. *et al*. (2003) Saccade reward signals in posterior cingulate cortex. *Neuron* 40, 1031–1040
14 Knutson, B. *et al*. (2001) Dissociation of reward anticipation and outcome with event related fMRI. *Neuroreport* 12, 3683–3687
15 O'Doherty, J. *et al*. (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102
16 Delgado, M.R. *et al*. (2003) Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn. Affect. Behav. Neurosci.* 3, 27–38
17 Elliott, R. *et al*. (2003) Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J. Neurosci.* 23, 303–307
18 McClure, S.M. *et al*. (2004) Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387
19 Kringelbach, M.L. *et al*. (2003) Activation of human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb. Cortex* 13, 1064–1071
20 Rilling, J.K. *et al*. (2004) Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* 15, 2539–2543
21 Mark, T.A. and Gallistel, C.R. (1993) Subjective reward magnitude of medial forebrain stimulation as a function of train duration and pulse frequency. *Behav. Neursci.* 107, 389–401
22 Montague, P.R. *et al*. (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947
23 Braver, T.S. and Cohen, J.D. (2000) On the control of control: the role of dopamine in regulating prefrontal function and working memory. In *Attention and Performance* (Monsell, S. and Driver, J., eds), pp. 713–737, Academic Press
24 Aston-Jones, G. and Cohen, J.D. (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450
25 Yu, A.J. and Dayan, P. (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692
26 Camerer, C.F. and Ho, T. (1999) Experience-Weighted Attraction (EWA) learning in normal-form games. *Econometrica* 67, 827–874
27 Merlo, A. and Schotter, A. (1999) A surprise-quiz view of learning in economic experiments. *Games Econ. Behav.* 28, 25–54
28 Falkenstein, M. *et al.* (1995) Event related potential correlates of errors in reaction tasks. In *Perspectives of Event-Related Potentials Research* (Karmos, G. *et al*., eds), pp. 287–296, Elsevier
29 Carter, C.S. *et al*. (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749
30 Holroyd, C.B. and Coles, M.G. (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709
31 Miltner, W.H. *et al*. (2003) Implementation of error-processing in the human anterior cingulate cortex: a source analysis of the magnetic equivalent of the error-related negativity. *Biol. Psychol.* 64, 157–166
32 Peyron, R. *et al*. (2000) Functional imaging of brain responses to pain. A review and meta-analysis. *Clin. Neurophysiol.* 30, 263–288
33 Gehring, W.J. and Willoughby, A.R. (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282
34 Yeung, N. and Sanfey, A.G. (2004) Independent coding of reward magnitude and valence in the human brain. *J. Neurosci.* 24, 6258–6264
35 Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decisions under risk. *Econometrica* 47, 262–291
36 Breiter, H.C. *et al*. (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639
37 Holroyd, C.B. *et al*. (2004) Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology* 41, 245–253
38 Knutson, B. *et al*. (2005) Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812
39 Berns, G.S. *et al*. (2001) Predictability modulates human response to reward. *J. Neurosci.* 21, 2793–2798
40 Schall, J.D. (2001) Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* 2, 33–42
41 Shadlen, M.N. and Newsome, W.T. (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936
42 Roitman, J.D. and Shadlen, M.N. (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489
43 Sugrue, L.P. *et al*. (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787
44 Usher, M. and McClelland, J.L. (2001) The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592
45 Gold, J.I. and Shadlen, M.N. (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16
46 Brown, E.T. *et al*. (2005) Simple neural networks that optimize decisions. *Int. J. Bifurcat. Chaos* 15, 803–826
47 Platt, M.L. and Glimcher, P.W. (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238
48 Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131
49 Posner, M. and Snyder, C. (1975) Facilitation and inhibition in the processing of signals. In *Attention and Performance V* (Rabbitt, P.M.A. and Dornic, S., eds), pp. 669–682, Academic Press
50 Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psych Rev* 84, 1–66
51 Kahneman, D. and Treisman, A. (1984) Changing views of attention and automaticity. In *Varieties of Attention* (Parasuraman, R. and Davies, D.R., eds), pp. 29–61, Academic Press
52 Cohen, J.D. *et al*. (1990) On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361
53 Sloman, S.A. (2002) Two systems of reasoning. In *Heuristics and Biases: The Psychology of Intuitive Judgment* (Gilovich, T. and Griffin, D., eds), pp. 379–396, Cambridge University Press
54 Kahneman, D. (2003) A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720

55 Starmer, C. (2000) Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *J. Econ. Lit.* 38, 332–382

56 Frederick, S. *et al*. (2002) Time discounting and time preference: a critical review. *J. Econ. Lit.* 40, 351–401

57 Kahneman, D. *et al*. (1982) *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press

58 Kahneman, D. *et al*. (1991) The endowment effect, loss aversion, and the status quo bias: anomalies. *J Econ Perspec* 5, 193–206

59 Ainslie, G. and Haslam, N. (1992) Hyperbolic discounting. In *Choice over Time* (Loewenstein, G. and Elster, J., eds), pp. 57–92, Russell Sage

60 Camerer, C.F. (2000) Prospect theory in the wild: evidence from the field. In *Choices, Values, and Frames* (Kahneman, D. and Tversky, A., eds), pp. 288–300, Cambridge University Press

61 Benartzi, S. and Thaler, R.H. (2004) Save more tomorrow: using behavioral economics to increase employee saving. *J. Polit. Econ.* 112, S164–S187

62 Montague, P.R. and Berns, G.S. (2002) Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284

63 Duncan, J. *et al*. (1996) Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognit. Psychol.* 30, 257–303

64 Smith, E.E. and Jonides, J. (1999) Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661

65 Miller & Cohen. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202

66 Stuss, D.T. and Knight, R.T. (2002) *Principles of Frontal Lobe Function*, Oxford University Press

67 Dalgleish, T. (2004) The emotional brain. *Nat. Rev. Neurosci.* 5, 583–589

68 Ochsner, K.N. and Gross, J.J. (2005) The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249

69 Sanfey, A.G. *et al*. (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758

70 de Quervain, D.J.F. *et al*. (2004) The neural basis of altruistic punishment. *Science* 305, 1254–1258

71 King-Casas, B. *et al*. (2005) Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83

72 Shefrin, H.M. and Thaler, R.H. (1988) The behavioral life-cycle hypothesis. *Econ. Inq.* 26, 609–643

73 Loewenstein, G. (1996) Out of control: visceral influences on behavior. *Org Behav Hum Dec Proc* 65, 272–292

74 McClure, S.M. *et al*. (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507

75 Bernheim, D. and Rangel, A. (2004) Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94, 1558–1590

76 Williamson, O. (1967) Hierarchical control and optimum firm size. *J. Polit. Econ.* 76, 123–138

77 Williamson, O. (1979) Transaction cost economics: the governance of contractual relations. *J. Law Econ.* 22, 233–261

78 Haier, R.J. *et al*. (1992) Regional glucose metabolic changes after learning a complex visuospatial/motor task: a PET study. *Brain Res.* 570, 134–143

79 Karni, A. *et al*. (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc. Natl. Acad. Sci. U. S. A.* 95, 861–868

80 Thompson, P. (2001) How much did the Liberty Shipbuilders learn? New evidence for an old case study. *J. Polit. Econ.* 109, 103–137