Neuroeconomics: The Consilience of Brain and Decision

Paul W. Glimcher* and Aldo Rustichini2

Economics, psychology, and neuroscience are converging today into a single, unified discipline with the ultimate aim of providing a single, general theory of human behavior. This is the emerging field of neuroeconomics in which consilience, the accordance of two or more inductions drawn from different groups of phenomena, seems to be operating. Economists and psychologists are providing rich conceptual tools for understanding and modeling behavior, while neurobiologists provide tools for the study of mechanism. The goal of this discipline is thus to understand the processes that connect sensation and action by revealing the neurobiological mechanisms by which decisions are made. This review describes recent developments in neuroeconomics from both behavioral and biological perspectives.

The full understanding of utility will come from biology and psychology by reduction to the elements of human behavior followed by a bottom-up synthesis, not from the social sciences by top-down inference and guesswork based on intuitive knowledge. It is in biology and psychology that economists and social scientists will find the premises needed to fashion more predictive models, just as it was in physics and chemistry that researchers found the premises that upgraded biology. (p. 206) (1)

Consider the famous St. Petersburg paradox (2). Which of the following would you prefer, $40 or a lottery ticket that pays according to the outcomes of one or more fair coin tosses: heads you get $2 and the game ends, tails you get another toss and the game repeats, but now if the second toss lands heads up you get $4, and so on. If the nth toss is the first to land heads up, you get

2nd dollars. The game continues, however long it takes, until the coin lands heads up. We can assess the average objective, or expected, value of this lottery by multiplying the probability of a win on each flip by the amount of that win:

\[
\text{Expected value} = (0.5 \times 2) + (0.25 \times 4) + (0.125 \times 8) + \ldots = 1 + 1 + 1 + \ldots
\]

This simple calculation reveals that the expected value of the lottery is infinite even though the average person is willing to pay less than $40 to play it. How could this be?

For an economist, any useful explanation must begin with a set of assumptions that renders behavior formally tractable to coherent theoretical and mathematical analysis. Economists therefore explain this behavior by assuming that the desirability of money does not increase linearly, but rather grows more and more slowly as the total amount at stake increases. For example, the desirability of a given amount might be a power function
the covered tray are stronger and thus control synapses that elicit an orienting response to behavioral circuits compete. In this case the under these circumstances the two different
ological explanation (3) presumes that subjective utilities are computed with regard to somewhat arbitrary and idiosyncratic monetary reference points, or frames, set by yet other psychological processes (3). Psychologists use observations like these to argue that human choosers are endowed with a particularly strong fear of losses and that they weigh the merits of all possible gains and losses relative to a psychological benchmark: The psychological approach seeks empirically to describe minimally complex behavioral tendencies, modules, or heuristics that can account for the actions of human choosers.

A traditional neurobiological perspective uses yet another approach: A hungry bird is shown a tray that contains five millet seeds and repeatedly permitted to fly to the tray and eat the seeds. At a neurobiological level, the study of this behavior begins with the assumption that the visual stimulus of the five seeds must somehow propagate through the sensory system of the animal to trigger activation in orienting circuits that move the bird to the seeds. Next, the same bird is permitted to fly to a second tray covered by a piece of paper. When the bird displaces the cover, half of the time it reveals 12 seeds and half of the time it reveals nothing. Mechanistically, the visual stimulus must again trigger an orienting response, and presumably in this case the strength with which visual signals connect synaptically to the orienting circuits reflects both the number of seeds that the bird earns and the likelihood that seeds will be found under the paper. Lastly, both trays are presented, and the bird is observed to fly toward the tray that may contain 12 millet seeds. A standard neurobiological explanation (4, 5) presumes that under these circumstances the two different behavioral circuits compete. In this case the synapses that elicit an orienting response to the covered tray are stronger and thus control circuitry that can account for the simplest measurable elements of behavior. The neurobiological explanation specifies the minimal neural circuitry required to account for the observed behavior of the bird.

What is striking about explanations of choice behavior by economists, psychologists, and neurobiologists is the different levels at which they operate. The economic approach attempts to describe globally all choice behavior with a single logically consistent formalism. The psychological approach examines the ways in which subjective and objective estimates of value differ and posts psychological modules that can account for these observed behavioral preferences. The neurobiological explanation starts with the simplest possible neural

- **Subjective Desirability**
- **Objective Measure**

![Figure 1.](image-url)

**Value of a Gain**
- **dollar, calories, milliliters**

---

**Subjective Desirability**
- Bernoulli’s notion of subjective value or utility. The black line plots the typical relationship between objective and subjective valuations of an action. As the objective value of a gain increases, the subjective desirability, or utility, grows more slowly. Bernoulli demonstrated that this relationship could account for the observation that humans are typically risk-averse. The solid gray line plots a condition in which subjective value grows more quickly than objective value, a preference structure that would yield risk-seeking behavior.

---

**Linking the Decision Sciences**

**Subjective desirability.** The central concept in modern economic theory is the notion of subjective utility: Preferences must be described as subjective properties of the chooser. Surprisingly, the notion that preferences are represented in the nervous system, that these preferences are subjective, and that they guide the production of action has only recently entered the neurobiological mainstream. We believe that this has been a critical flaw in neurobiological studies, because it is essential that economics, psychology, and neuroscience acknowledge a common phenomenological base to achieve a reductive unification of the decision sciences. The concepts that guide the behavioral study of decision-making must also guide the mechanistic study of that process.

In part, this preference-free approach may have arisen from neurobiology’s roots in the stimulus-response physiology of the
early twentieth century (10). Working from Descartes’ philosophy, Sherrington (11) proposed that physiologists should work to link stimulus and response directly through what Pavlov (12) would later call a “definite nervous path.” Scientists like Sherrington and Pavlov proposed that it was the role of neuroscience to chart these stimulus-response connections through the nervous system.

A critical step beyond this initial framework was a recent effort to explain more complicated behaviors and to focus on actions for which deterministic sensory-to-movement mapping approaches were insufficient. Newsome and his colleagues (13, 14) made that step in the late 1980s when they examined perceptual decision-making by monkeys viewing ambiguous sensory stimuli. In those experiments, monkeys stared at a display of chaotically moving spots of light. On training trials, a subset of the spots moved coherently in a single direction, whereas the remaining spots moved randomly (15). The direction of this coherent motion indicated which of two possible saccadic eye movements would yield a fruit juice reward, and at the end of each trial animals were free to make a saccade. If they made the correct movement, they then received the reward. On a critical subset of trials, however, monkeys viewed displays in which none of the dots moved in a single coherent direction, and thus the display provided no information from which the location of the rewarded eye movement could be deduced. Under these conditions, Newsome and his colleagues found that the firing rates of single neurons in the middle temporal visual area (area MT) were still correlated with the behavior of the animals, even when that behavior could not be predicted from the properties of the stimulus. Newsome and his colleague Shadlen’s subsequent studies revealed the basic neurobiological substrate for perceptual decision-making and showed convincingly that this circuit could not be modeled simply as a single “definite nervous path” from stimulus to response (16).

This work, in turn, accelerated studies of the posterior parietal cortex, an area interposed between many of the sensory circuits and motor circuits of the primate brain, which appeared to play a critical role in the perceptual decision-making Newsome has studied (17, 18). Platt and Glimcher (19) made an important advance when they extended Newsome’s approach by proposing that posterior parietal cortex might play a role in decision-making in an economic sense and that it might encode the desirabilities of making particular movements.

In Platt and Glimcher’s experiments, trained rhesus monkeys were allowed to participate in repeated rounds of a simple lottery while the activity of nerve cells in the posterior parietal cortex was monitored. At the beginning of each round, two yellow spots were illuminated on a screen, one to the left and one to the right of where the monkey was looking. This began the lottery phase of the round, a period during which the monkey did not know whether the left or right light would be offered as a prize at the end of that round. At the end of this phase, a third light changed color to red or green, indicating which of the two initial lights had been randomly selected to yield a fruit juice reward on that particular round. The monkey received the fruit juice if he oriented to the selected light at the end of the round. While monkeys played hundreds of rounds of this game, Platt and Glimcher systematically varied either the relative probabilities that the left or right lights would be selected at the end of each round or the size of the reward associated with each. These two variables were selected because economic theories assess desirability by combining the value and likelihood of gain in some subjective manner. Platt and Glimcher found that some parietal neurons did indeed encode the value and likelihood of reinforcement during the lottery phase of each round. Under these conditions, the brains of the monkeys explicitly encoded something very much like the economically defined expected value or expected utility of each light in this simple lottery task.

Subsequent studies of human decision-making using functional magnetic resonance imaging (fMRI) have yielded similar conclusions. Koutson and colleagues (20) have shown, for example, that activity in the human striatum is correlated with the magnitude of the monetary reward subjects earn during lotteries, and Paulus and colleagues (21) have shown a similar result in the human posterior parietal cortex. In a particularly interesting study, Breiter and colleagues (22) (Fig. 2) presented human subjects on sequential rounds with one of three possible lotteries. In lottery one (the good lottery), they faced equal chances of winning $10, $2.50, or $0. In lottery two (the intermediate lottery), they faced equal chances of winning $2.50, winning $0, or losing $1.50. In lottery three (the bad lottery), they faced equal chances of winning $0, losing $1.50, or losing $6. At the beginning of each round, the subjects were told which lottery they would be playing, and the average activity in many brain areas was simultaneously measured. After that measurement was complete, the lottery was actually played and the humans were then told how much real money they had earned on that round. Importantly, all three of these lotteries present a one-third possibility of winning $0, but they do so under different conditions. In the good lottery, winning $0 is the worst possible outcome, whereas in the bad lottery it is the best. The psychologists Kahneman and Tversky (23) have shown that, when a human participates in this good lottery, they find winning $0 to be an intensely negative outcome whereas when a human participates in the bad lottery, they find winning $0 to be a positive outcome; subjective utilities are computed with regard to a reference frame. Breiter and colleagues found that the activity of the sublenticular extended amygdala encoded the desirability of each lottery, taking into account this behaviorally described framing effect.

Other recent neurobiological studies have revealed yet other neurally encoded variables that include the log likelihood that a given eye movement will result in a reward (24, 25), the very closely related integral of perceptual signals indicating which saccade will be rewarded in the Newsome task (26, 27), the average rate at which a saccade has been rewarded in the recent past (28), the instantaneous likelihood, or hazard, that a reinforced saccade will be instructed (29), and combinations of these variables (30).

Strategic thinking. All of these results suggest that classical utility theory can be used as a central concept for the study of choice in economics, psychology, and neuroscience. In the middle of the twentieth century, however, economists pushed utility theory beyond this boundary, enhancing it to include the study of the strategic interactions which arise when decision makers confront intelligent opponents. Extending the concept of subjective utility, VonNeumann and Morgenstern (31) and Nash (32, 33) developed a formal utility-based economic approach in the theory of
games. Recently, Lee and his colleagues (34) and Dorris and Glimcher (35) have begun to
link the neurobiological corpus to this literature by examining the activity of single
neurons in awake-behaving monkeys engaged in decision-making during strategic
conflicts. In Dorris and Glimcher’s study, two opponents face each other, an employer
and an employee. On each round of the game the employee must decide whether to go to
work, in which case he earns a fixed wage, or whether to shirk, in hopes of earning his
wage plus a bonus. The goal of the employee is simply to maximize his gains in terms of
salary and bonus. The employer, on the other
hand, must decide between trusting his em-
ployee to arrive for work or spending money
to hire an inspector who can actually check
and see whether the employee arrived for
work that day. The goal of the employer is to
spend as little as possible on inspections while
maximizing the employee’s incentive to work.

The inspection game is of particular interest to game
theorists and economists because rational strategies for
utility maximization during this strategic conflict lead
to predictable outcomes according to an equilibrium
theory originally developed by John Nash in the 1950s.
Nash (36) equilibrium theory describes how, when the
cost of inspection to the employer is set high, the
efficient strategy for both players converges on a solution in which the employ-
ee manages to shirk fairly often. Conversely, a low in-
spection cost to the employer defines a theoretical equilibrium solution in which
shirk rates are low. One of the fundamental insights this formal analysis reveals is that at
a mixed strategy equilibrium, a situation in which a rational player should distribute his
actions amongst two or more alternatives in an unpredictable fashion, the desirability of
the two or more actions in equilibrium must be equivalent. The Nash approach argues,
especially, that a behavioral equilibrium occurs when the desirability of working and
shirking are rendered equal by the behavior of one’s opponent irrespective of how often
that equilibrium requires that one work. When Dorris and Glimcher examined the
activity of neurons in the posterior parietal cortex of monkeys playing the inspection
game, they found neurons that carried a signal that behaved like relative expected util-
ity. When the monkey’s behavior was well predicted by the Nash equations, neural ac-
tivity was fixed at a single level irrespective of the frequency with which the monkey
chose to make a particular response, even though these same neurons were strongly
modulated by changes in the value of actions during lottery tasks.

Research on human-human strategic interactions that are well described by
classical game theory are also now under
way in a number of laboratories (37). Like
the earlier fMRI studies of simple decision-
making tasks, these experiments are also beginning to shape the common ground
between economics, psychology, and neu-
roscience. Taken together, these findings
suggest that at least under some circumstances decisions may actually be made in
the primate neuro-architecture in a manner long suspected by economists and now
being actively analyzed by psychologists and neuroscientists: Neural circuits may
compute and represent the desirability of making a response. Economics, psycholo-
y, and neuroscience do seem to be
converging around a common conceptual
framework. All three disciplines are begin-
ning to acknowledge that decision-making involves the representation of subjective
desirabilities. The challenge that this con-
vergence around a single concept poses, however, is to leverage the intersection of these
disciplines to explain choice behavior that cannot be described with the common
framework of utility theory: these
are classes of behaviors which have sty-
mied traditional economics and which have lain far beyond the reach of traditional
neuroscience. If it is to be of value, the goal
of a unified decision science will have to be
to use all three sets of approaches simul-
taneously to gain traction in this new
territory.

Beyond Classical Concepts
Choice under risk. As we have seen, the
introduction of the concept of expected utility
solved the St. Petersburg’s puzzle and formed
the core of neoclassical economics. Sub-
sequent puzzles and paradoxes, however, have
plagued this solution. In Ellsberg’s (1961)
paradox (38) (Fig. 3) you are presented with
an urn, and you are told that it contains 90
balls. Of these, 30 are blue, and 60 are either
red or yellow; any proportion is possible.
You are then offered a choice between a
lottery that pays $100 if a blue ball is drawn
(a 1/3 probability) and one that pays $100 if a
red ball is drawn. The probability of a red
draw is unspecified or ambiguous: It is a
choice between an event with a known
probability and an event with an unknown
probability. Under these circumstances, peo-
ple typically choose the first lottery, which
wins if a blue ball is drawn. According to
expected utility theory they could only do so
if they believe that there are fewer than 30
red balls in the urn or, equivalently, that
there are more than 30 yellow balls. Then
(before any balls are actually drawn, but with
the same urn standing in front of you) you are
asked to choose again, this time between a lottery
that pays $100 on either blue or yellow and one
that pays $100 on either red or yellow. Now the likelihood
of winning is clear in the
second case (a 2/3 proba-
bility of winning $100) but
unclear in the first case (a
probability between 1/3 and 1). People this time
typically choose the second
lottery. The first lottery
seems less attractive, be-
cause there might be too
few yellow balls. Is there
anything wrong with this
behavior? If expected util-
ity theory is correct, then there certainly is:
You cannot think that there are too few and
too many yellow balls in the urn at the same
time.

The Ellsberg paradox is just one of many
demonstrations presented in the last half of
the twentieth century that were considered
formal falsifications of expected utility
theory. An even earlier example is Allais’
paradox (39), based on the idea that a certain
outcome may be perceived as more desir-
able, in a qualitatively different way, than
any random outcome, even if very likely
(40). These examples proved that expected utility theory as originally proposed could
not be globally correct; at best it could only
predict choices under some circumstances.
This has led economists and social psychol-
ologists both to attempt modifications to
expected utility theory and to replace it
outright. Both the modifications and replace-
ments have provided important and econom-
ically powerful insights into choice behavior
but have not yet provided a global theory of
choice that can truly replace expected utility theory.

The emerging discipline of neuroeconomics offers a new strategy both for testing existing models of all types and for developing new models with empirical techniques. If we succeed in understanding mechanistically how choices that violate expected utility theory are made at a neural level, then a new global theory of choice will be developed. To that end, a number of laboratories are now beginning to reexamine the conditions under which expected utility theory fails.

The reason that expected utility theory fails under some conditions may be that choosers use more than one evaluative mechanism at a neurobiological level (41). For example, in Dickhaut et al. (42), the processes involved when a certain outcome is one of the options are different from those involved when only random outcomes are at stake, providing an explanation of the Allais’ paradox cited above.

Under many conditions these mechanisms may work together to yield choices similar to those predicted by expected utility theory but may produce odd results when used in isolation, in novel combinations, or in situations for which they are ill suited. Recent work by Damasio and colleagues [for example, (43)] on the class of behavioral paradoxes from which the Ellsberg example is drawn seem to support this conclusion. These studies suggest that an ambiguity-sensitive mechanism associated with the expression of emotion may reside, at least in part, in the ventromedial prefrontal cortex (VMPFC) (Fig. 4) and may be responsible for choice under some but not all conditions. These researchers and others have shown that patients with damage to this area have an impaired ability to make some classes of decisions and have difficulties planning their work and choosing friends. Further, the actions these individuals do elect to pursue often lead to financial as well as personal losses. Yet despite these specific failures, patients with damage to the VMPFC show normal performance on multiple-choice tests of intelligence.

These observations and others like them have led Damasio to propose that the inability of patients with VMPFC lesions to make advantageous decisions under some circumstances is caused by damage to an emotional mechanism that stores and signals the value of future consequences of an action, the somatic marker hypothesis. The hypothesis proposes that, because they lack this emotional mechanism, the patients must rely on other brain mechanisms that achieve a different analysis of the numerous and often conflicting options involving both immediate and future consequences. This other mechanism, operating alone, is hypothesized to produce decisions that are less efficient and slower than those produced by a normal, intact, system.

The importance of the emotion-related VMPFC for regular decision-making has been confirmed by experiments where subjects were asked to make choices among a group of alternatives that carry a monetary reward (typically by selecting one card at a time from four different decks of cards), but for which the probability of reward is unspecified (44). This is precisely the ambiguous situation that produces Ellsberg’s paradox. Under these conditions, patients with VMPFC lesions seem to lack a mechanism at a neurobiological level (41).

Under these conditions, the VMPFC did not show any activation; it was actually less active when choices were being made than when subjects waited between trials. These results suggest that emotional circuits may be important in learning and processing information, rather than in selecting among alternatives.

Together, these data may begin to explain, in a mechanistic way, how information is analyzed when at least one class of behavior which is not predicted by the expected utility theory is produced. The process of learning and evaluating feedback may involve emotion-related areas. Ambiguity aversion, whether advantageous or disadvantageous in a particular situation, may become explicable as we learn more about the computations that brain areas like the VMPFC perform.

**Strategic cooperation.** As with the Ellsberg paradox, challenges have also been raised recently to classical game theory. In a path-breaking study, Guth et al. (47) analyzed the behavior of subjects playing the ultimatum game. In this game, a first player, the proposer, has $10 to split with a second player. He can offer any amount between zero and $10. The second player is informed of the offer and can accept or refuse. If she accepts, the split is made. If she refuses, both players get nothing. The prediction of a restrictive concept of game theory, the subgame perfect equilibrium, is that for any positive amount offered by the proposer, the second player knows that she faces a choice between gaining nothing (if she refuses the offer) or something (if she accepts). The proposer should therefore always offer the minimum possible split to player two, who should always accept. Contrary to this prediction, the robust experimental finding is that low offers (typically S2 or even S3) are consistently refused. The second player appears to prefer, under these conditions, to gain nothing. Anticipating this, proposers typically avoid low offers.

Although expected utility theorists have proposed some explanations for this behavior, it may well be that by analyzing the neural circuits active during the ultimatum game we may be able to both explain the causes of this behavior and to predict it. Studying the ultimatum game in subjects undergoing brain scans, Sanfey et al. (48) found that offers refused by the second players activated specific brain circuits in those players, and interestingly these brain circuits are also associated with emotional arousal: the anterior insula (AI, associated with disgust, both physical and emotional), the dorsolateral prefrontal

*Fig. 4. Medial view of the left half of a human brain, with the front of the brain on the right side of the image. The human ventromedial prefrontal cortex is shown in red.*
cortex (DLPFC, associated with goal maintenance and executive control), and the anterior cingulate cortex (ACC, associated with detection of cognitive conflict). Also significant is the correlation of activation with choices: An activation of the AI is positively correlated with rejection, suggesting that an emotional arousal associated with a low offer is correlated with rejection. The overall picture is that offers we might consider unfair may activate emotional circuits of the brain involved in the decision to reject an offer. If we can come to fully understand how these circuits reach this conclusion, then a behavior that was difficult for classical game theory to predict may become fully explicable with the synthetic approach that neuroeconomics provides.

A similar line of investigation has examined interplayer cooperation during single rounds of the trust game. In this game, two players move sequentially. The first player can decide to transfer a sum of money out of an initial endowment that she receives into an investment pool that immediately triples in value. The second player then gains control of the investment and can divide it between the two players in any way he chooses. In this game, the only Nash equilibrium choice for the first player is to transfer nothing into the investment. Were she to make any transfer, the second player should take all of the money for himself. However, in real experiments the first player typically does transfer a significant amount into the investment, and the second player reciprocates by returning part of the pool. McCabe et al. (49) had subjects play the trust game both against a human opponent and against a computer program which, they were told, would play a human-like strategy. Under these conditions McCabe and colleagues found that subjects were more likely to cooperate with real humans than with computers and that cooperators have a significantly different brain activation in the two conditions. Cooperation is associated with activation of the anterior paracingulate cortex, a brain region associated with (50, 51) interpreting and monitoring the mental state of others.

Although these studies are in their early stages, they suggest the existence of specific brain components that make specialized contributions to decision making. The challenge that these studies face is to derive detailed computational models of the neural mechanisms, which will make neuroeconomics models broadly predictive as well as explanatory.

Summary
Economics, psychology, and neuroscience are converging into a single, unified field aimed at providing a theory of human behavior. In this enterprise, the method and the standard set by neuroscience is the final goal: a reconstruction of the process and mechanism that goes from a stimulus presented to the subject to his final action in response. Economics provides the conceptual structure and the object of the analysis. In this emerging view, people are seen as deciding among options on the basis of the relative desirability of each option. This is true when they are in isolation as well as when they are in strategy (interaction with few persons) and market (interaction with a large number) environments. The recent research we have been surveying describes how desirability is realized as a concrete object, a neural signal in the human and animal brain, rather than as a purely theoretical construction. Desirability is computed and is represented in the brain, and we now have the means to test, measure, and represent this activation.

But the complete reconstruction of the decision process, and hence of human behavior, is not going to be easy, because two of the cornerstones of economic analysis, subjective utility theory and Nash equilibrium, provide, even from the descriptive point of view, an incomplete picture. For example, desirability as represented by the simple economic formalism of expected utility may be appropriate only in simple conditions, where ambiguity is excluded. A more general notion is needed, and as we have seen, is beginning to be investigated and developed by psychologists and economists working together. The goal of the emerging neuroeconomic program will have to be a mechanistic, behavioral, and mathematical explanation of choice that transcends the explanations available to neuroscientists, psychologists, and economists working alone. Although it is unclear today how complete this explanation will ultimately be, neuroeconomic approaches have already begun to yield substantial fruit and to fuse natural and social scientific approaches to the study of human behavior.

References and Notes
14. See also (50) and (51) for a parallel set of studies that presaged and continues to parallel many of these developments.
39. For a review of these issues, see D. Luce, Utility of Gains and Losses, vol. 8 of Scientific Psychology Series (L. Erlbaum, Mahwah, NJ, 2000).
42. See for example A. Bechara, H. Damasio, D. Tranel, A. Damasio, Science 275, 1293 (1997).
44. A. Rustichini, J. Dickhaut, P. Chirivatto, K. Smith, J. A. Ardo, Games and Economic Behavior, in press.
51. K. Conover, P. Shizgal, in Games and Economic Behavior, in press.