Scand. J. of Economics 106(3), 555–579, 2004
DOI: 10.1111/j.1467-9442.2004.00378.x

Neuroeconomics: Why Economics Needs Brains*

Colin F. Camerer
California Institute of Technology, Pasadena, CA 91125, USA
camerer@hss.caltech.edu

George Loewenstein
Carnegie-Mellon University, Pittsburgh, PA 15213, USA
gl20@andrew.cmu.edu

Drazen Prelec
MIT, Cambridge, MA 02139, USA
dprelec@mit.edu

Abstract

Neuroeconomics uses knowledge about brain mechanisms to inform economic theory. It opens up the "black box" of the brain, much as organizational economics opened up the theory of the firm. Neuroscientists use many tools—including brain imaging, behavior of patients with brain damage, animal behavior and recording single neuron activity. The key insight for economics is that the brain is composed of multiple systems which interact. Controlled systems ("executive function") interrupt automatic ones. Brain evidence complicates standard assumptions about basic preference, to include homeostasis and other kinds of state-dependence, and shows emotional activation in ambiguous choice and strategic interaction.

Keywords: Behavioral economics; neuroscience; neuroeconomics; brain imaging

JEL classification: *C*91; *D*81

I. Introduction

In a strict sense, all economic activity must involve the human brain. Yet, economics has achieved much success with a program that sidestepped the

* We thank participants at the Russell Sage Foundation-sponsored conference on Neurobehavioral Economics (May 1997) at Carnegie-Mellon, the Princeton workshop on Neural Economics (December 2000) and the Arizona conference (March 2001). This research was supported by NSF grant SBR-9601236 and by the Center for Advanced Study in Behavioral Sciences, where the authors visited during 1997–1998. David Laibson's presentation at the Princeton conference was particularly helpful, as were comments and suggestions from referees, John Dickhaut, Paul Zak, a paper by Jen Shang, and conversations with John Allman, Greg Berns, Jonathan Cohen, Angus Deaton, Dave Grether, Brian Knutson, David Laibson, Danica Mijovic-Prelec, Read Montague, Charlie Plott, Matthew Rabin, Peter Shizgal and Steve Quartz.

biological and cognitive sciences that focus on the brain, in favor of the maximization style of classical physics, with agents choosing consumption bundles having the highest utility subject to a budget constraint, and allocations determined by equilibrium constraints. Later tools extended the model to include utility tradeoffs with uncertainty and time, Bayesian processing of information, and rationality of expectations about the economy and about the actions of other players in a game.

Of course these economic tools have proved useful. But it is important to remember that before the emergence of revealed preference, many economists had doubts about the rationality of choice. In 1925, Viner (pp. 373–374), lamented that

> "Human behavior, in general, and presumably, therefore, also in the market place, is not under the constant and detailed guidance of careful and accurate hedonic calculations, but is the product of an unstable and unrational complex of reflex actions, impulses, instincts, habits, customs, fashions and hysteria."

At the same time, economists feared that this "unstable and unrational complex" of influences could not be measured directly. Jevons (1871) wrote, "I hesitate to say that men will ever have the means of measuring directly the feelings of the human heart. It is from the quantitative effects of the feelings that we must estimate their comparative amounts." The practice of assuming that unobserved utilities are revealed by observed choices—revealed preference—arose as a last resort, from skepticism about the ability to "measure directly" feelings and thoughts.

But Jevons was wrong. Feelings and thoughts *can* be measured directly now, because of recent breakthroughs in neuroscience. If neural mechanisms do not always produce rational choice and judgment, the brain evidence has the *potential* to suggest better theory.

The theory of the firm provides an optimistic analogy. Traditional models treated the firm as a black box which produces output based on inputs of capital and labor and a production function. This simplification is useful but modern views open the black box and study the contracting practices inside the firm—viz., how capital-owners hire and control labor. Likewise, neuroeconomics could model the details of what goes on inside the consumer mind just as organizational economics models what goes on inside firms.

This paper presents some of the basic ideas and methods in neuroscience, and speculates about areas of economics where brain research is likely to affect predictions; see also Zak (2004), and Camerer, Loewenstein and Prelec (2004) for more details. We postpone most discussion of why economists should care about neuroscience to the conclusion.

II. Neuroscience Methods

Many different methods are used in neuroscience. Since each method has strengths and weaknesses, research findings are usually embraced only after they are corroborated by more than one method. Like filling in a crossword puzzle, clues from one method help fill in what is learned from other methods.

Much neural evidence comes from studies of the brains of non-human animals (typically rats and primates). The "animal model" is useful because the human brain is basically a mammalian brain covered by a folded cortex which is responsible for higher functions like language and long-term planning. Animal brains can also be deliberately damaged and stimulated, and their tissues studied.

Many human physiological reactions can be easily measured and used to make inferences about neural functioning. For example, pupil dilation is correlated with mental effort; see Kahneman and Peavler (1969). Blood pressure, skin conductance (sweating) and heart rate are correlated with anxiety, sexual arousal, mental concentration and other motivational states; see Levenson (1988). Emotional states can be reliably measured by coding facial expressions and recording movements of facial muscles (positive emotions flex cheekbones and negative emotions lead to eyebrow furrowing); see Ekman (1992).

Brain imaging: Brain imaging is the great leap forward in neuroscientific measurement. Most brain imaging involves a comparison of people performing different tasks—an "experimental" task E and a "control" task C. The difference between images taken during E and C shows what part of the brain is differentially activated by E.

The oldest imaging method, electro-encephalogram (or EEG) measures electrical activity on the outside of the brain using scale electrodes. EEG records timing of activity very precisely (~1 millisecond) but spatial resolution is poor and it does not directly record interior brain activity. Positron emission topography (PET) is a newer technique, which measures blood flow in the brain using positron emissions after a weakly radioactive blood injection. PET gives better spatial resolution than EEG, but poorer temporal resolution and is limited to short tasks (because the radioactivity decays rapidly). However, PET usually requires averaging over fewer trials than fMRI.

The newest method is functional magnetic resonance imaging (fMRI). fMRI measures changes in blood oxygenation, which indicates brain activity because the brain effectively "overshoots" in providing oxygenated blood to active parts of the brain. Oxygenated blood has different magnetic properties from deoxygenated blood, which creates the signal picked up by fMRI. Unfortunately, the signal is weak, so drawing inferences requires repeated sampling and many trials. Spatial resolution in fMRI is better than PET (~3 millimeter[3] "voxels"). But technology is improving rapidly.

Single-neuron measurement: Even fMRI only measures activity of "circuits" consisting of thousands of neurons. In single neuron measurement, tiny electrodes are inserted into the brain, each measuring a *single* neuron's firing. Because the electrodes damage neurons, this method is only used on animals and special human populations (when neurosurgeons use implanted electrodes to locate the source of epileptic convulsions). Because of the focus on animals, single neuron measurement has so far shed far more light on basic emotional and motivational processes than on higher-level processes such as language and consciousness.

Psychopathology: Chronic mental illnesses (e.g., schizophrenia), developmental disorders (e.g., autism), and degenerative diseases of the nervous system (e.g., Parkinson's Disease (PD)) help us understand how the brain works. Most forms of illness have been associated with specific brain areas. In some cases, the progression of illness has a localized path in the brain. For example, PD initially affects the basal ganglia, spreading only later to the cortex. The early symptoms of PD therefore provide clues about the specific role of basal ganglia in brain functioning; see Lieberman (2000).

Brain damage in humans: Localized brain damage, produced by accidents and strokes, and patients who underwent radical neurosurgical procedures, are an especially rich source of insights; see e.g. Damasio (1994). If patients with known damage to area X perform a particular task more poorly than "normal" patients, the difference is a clue that area X is necessary to do that task. Often a single patient with a one-of-a-kind lesion changes the entire view in the field (much as a single crash day in the stock markets—October 19, 1987—changed academic views of financial market operations). For example, patient "S.M." has bilateral amygdala damage. She can recognize all facial expressions except fear; and she does not perceive faces as untrustworthy the way others do. This is powerful evidence that the human amygdala is crucial for judging who is afraid and who to distrust. "Virtual lesions" can also be created by "transcranial magnetic stimulation (TMS)", which creates temporary local disruption to brain regions using magnetic fields.

III. Stylized Facts about the Brain

We now review some basic facts about the brain, emphasizing those of special interest to economists. Figure 1 shows a "sagittal" slice of the human brain, with some areas that are mentioned below indicated. It has four lobes—from front to back (left to right, clockwise in Figure 1), *frontal*, *parietal*, *occipital* and *temporal*. The frontal lobe is thought to be the locus

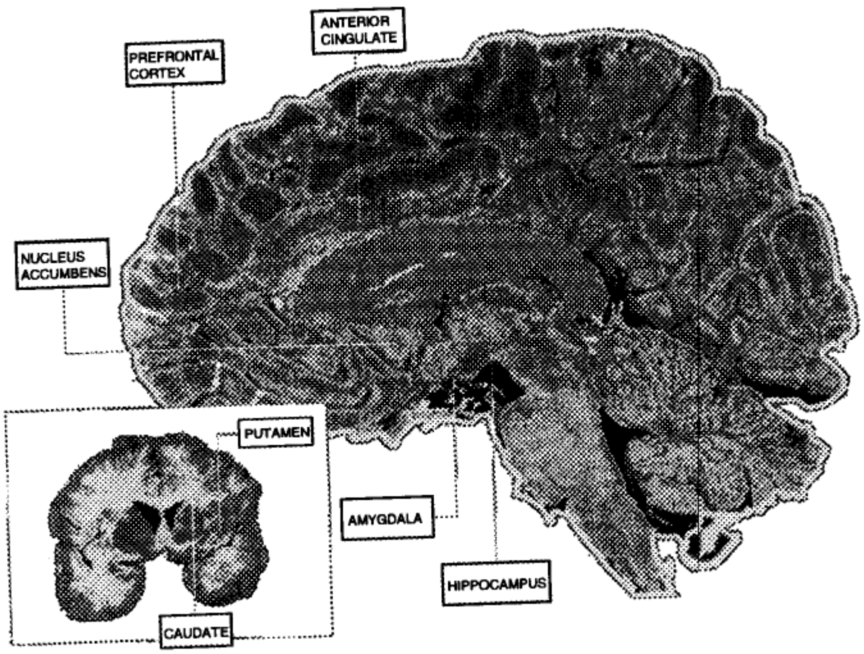

Fig. 1. Human brain (frontal pole left) regions of potential interest to economists

of planning, cognitive control and integration of cross-brain input. Parietal areas govern motor action. The occipital lobe is where visual processing occurs. The temporal lobes are important for memory, recognition and emotion. Neurons from different areas are interconnected, which enables the brain to respond to complex stimuli in an integrated way. When an automated insurance broker calls and says, "Don't you want earthquake insurance? Press 1 for more information" the occipital lobe "pictures" your house collapsing; the temporal lobe feels a negative emotion; and the frontal lobe receives the emotional signal and weighs it against the likely cost of insurance. If the frontal lobe "decides" you should find out more, the parietal lobe directs your finger to press 1 on your phone.

A crucial fact is that the human brain is basically a mammalian brain with a larger cortex. This means human behavior will generally be a compromise between highly evolved animal emotions and instincts, and more recently evolved human deliberation and foresight; see e.g. Loewenstein (1996). It also means we can learn a lot about humans from studying primates (who share more than 98% of our genes) and other animals.

Three features of human brain function are notable: *automaticity, modularity* and *sense-making*. According to a prominent neuroscientist, Gazzaniga (1988) wrote:

"Human brain architecture is organized in terms of functional modules capable of working both cooperatively and independently. These modules can carry out their functions in parallel and outside of conscious experience. The modules can effect internal and external behaviors, and do this at regular intervals. Monitoring all this is a left-brain-based system called the interpreter. The interpreter considers all the outputs of the functional modules as soon as they are made and immediately constructs a hypothesis as to why particular actions occurred. In fact the interpreter need not be privy to why a particular module responded. Nonetheless, it will take the behavior at face value and fit the event into the large ongoing mental schema (belief system) that it has already constructed."

Many brain activities are *automatic* parallel, rapid processes which typically occur without awareness. Automaticity implies that "people"—i.e., the deliberative cortex and the language processing which articulates a person's reasons for their own behavior—may genuinely not know the cause of their own behavior.[1]

Automaticity means that overcoming some habits is only possible with cognitive effort, which is scarce. But the power of the brain to automatize also explains why tasks which are so challenging to brain and body resources that they seem impossibly difficult at first—windsurfing, driving a car, paying attention to four screens at once in a trading room—can be done automatically after enough practice.[2] At the same time, when good performance becomes automatic (in the form of "procedural knowledge") it is typically hard to articulate, which means human capital of this sort is difficult to reproduce by teaching others.

The different brain modules are often neuroanatomically separated (like organs of the body). Some kinds of modularity are really remarkable: The "facial fusiform area" (FFA) is specialized for facial recognition; "somatosensory cortex" has areas corresponding directly to different parts of the body (body parts with more nerve endings, like the mouth, have more corresponding brain tissue); features of visual images are neurally encoded in different brain areas, reproducing the external visual

[1] For example, 40-millisecond flashes of angry or happy faces, followed immediately by a neutral "mask" face, activate the amygdala even though people are completely unaware of whether they saw a happy or angry face; see Whalen, Rauch, Etcoff, McInerney, Lee and Jenike (1998).

[2] Lo and Repin (2002) recorded psychophysiological measures (like skin conductance and heart rate) with actual foreign exchange traders during their work. They found that more experienced traders showed lower emotional responses to market events that set the hearts of less experienced traders pounding. Their discovery suggests that responding to market events becomes partially automated, which produces less biological reaction in experienced traders.

organization of the elements internally ("retinotopic mapping"); and there are separate language areas, Broca's and Wernicke's areas,[3] for semantics and for comprehension and grammar.

Many neuroscientists think there is a specialized "mentalizing" (or "theory of mind") module, which controls a person's inferences about what other people believe, or feel, or might do; see e.g. Fletcher, Happe, Frith, Baker, Dolan, Frackowiak and Frith (1995). Such a module presumably supports a whole range of critical human functions—decoding emotions, understanding of social rules, emotions, language, strategic concepts (bluffing)—and has obvious importance for economic transactions.

Modularity is important for neuroeconomics because it invites tests that map theoretical distinctions onto separate brain areas. For example, if people play games against other people differently than they make decisions (a "game against nature"), as is presumed in economic theory, those two tasks should activate some different brain areas.

However, the modularity hypothesis should not be taken too far. Most complex behaviors of interest to economics require collaboration among more specialized modules and functions. So the brain is like a large company—branch offices specialize in different functions, but also communicate to one another, and communicate more feverishly when an important decision is being made. Attention in neuroeconomics is therefore focused not just on specific regions, but also on finding "circuits" or collaborative systems of specialized regions which create choice and judgment.

The brain's powerful drive toward *sense-making* leads us to strive to interpret our own behavior. The human brain is like a monkey brain with a cortical "press secretary" who is glib at concocting explanations for behavior, and privileges deliberative explanations over cruder ones; cf. Nisbett and Wilson (1977) and Wegner and Wheatley (1999). An important feature of this sense-making is that it is highly dependent on expectations; in psychological terms, it is "top down" as opposed to "bottom-up". For example, when people are given incomplete pictures, their brains often automatically fill in the missing elements so that there is never any awareness that anything is missing. In other settings, the brain's imposition of order can make it detect patterns where there are none; see Gilovich (1991). When subjects listen to music and watch flashing Christmas tree lights at the same time, they mistakenly report that the two are synchronized. Mistaken beliefs in sports streaks, as evidenced by Gilovich, Vallone and Tversky (1985), and seeing spurious patterns in time series like stock-price data ("technical analysis") may come from "too much" sense-making.

[3] Patients with Wernicke damage can babble sentences of words which make no sense strung together. Broca patients' sentences make sense but they often "can't find just the right word".

Top-down encoding also implies the brain misses images it does not expect to see. A dramatic example is "change-blindness". In an amusing study titled "Gorillas in our Midst", subjects watch a video of six people passing a basketball and count the passes made by one "team" (indicated by jersey color). Forty seconds into the film clip, a gorilla walks into the center of the game, turns to the camera, thumps its chest, and then walks off. Although the gorilla cavorts onscreen for a full total of nine seconds, about one-half of the subjects remain oblivious to the intrusion, even when pointedly asked whether they had seen "the gorilla walking across the screen"; see Simons and Chabris (1999).

When the brain does assimilate information, it does so rapidly and efficiently, "overwriting" what was previously believed. This can create a powerful "hindsight bias" in which events seem, after the fact, to have been predictable even when they were not. Hindsight bias is probably important in agency relations when an agent takes an informed action and a principal "second-guesses" the agent if the action turns out badly. This adds a special source of risk to the agent's income and may lead to other behaviors like herding, diffusion of responsibility, inefficiencies from "covering your ass", excessive labor turnover, and so on.

We emphasize these properties of the brain, which are rapid and often implicit (subconscious), because they depart the most from conscious deliberation that may take place in complex economic decisions like saving for retirement and computing asset values. Our emphasis does not deny the importance of deliberation. The presence of other mechanisms just means that the right models should include many components and how they interact.

IV. Topics in Neuroeconomics

Preferences

Thinking about the brain suggests several shortcomings with the standard economic concept of preference.

1. Feelings of pleasure and pain originate in *homeostatic* mechanisms that detect departures from a "set-point" or ideal level, and attempt to restore equilibrium. In some cases, these attempts do not require additional voluntary actions, e.g., when monitors for body temperature trigger sweating to cool you off and shivering to warm you up. In other cases, the homeostatic processes operate by changing momentary preferences, a process called "alliesthesia"; see Cabanac (1979). When the core body temperature falls below the 98.6F set-point, almost anything that raises body temperature (such as placing one's hand in warm water) feels good, and the opposite is true when body temperature is too high. Similarly, monitors for blood sugar levels, intestinal distention and many other variables trigger hunger.

Homeostasis means preferences are "state-dependent" in a special way: the states are internal to the body and both affect preferences *and* act as information signals which provoke equilibration. Some kinds of homeostatic state-dependence are "contagious" across people—for example, the menstrual cycles of females living together tend to converge over time. Perhaps "waves" of panic and euphoria in markets work in a similar way, correlating responses so that internal states become macroeconomic states (as in the "animal spirits", which, in Keynes's view, were a cause of business cycles).

2. Inferring preferences from a choice does not tell us everything we need to know. Consider the hypothetical case of two people, Al and Naucia, who both refuse to buy peanuts at a reasonable price; cf. Romer (2000). The refusal to buy reveals a common disutility for peanuts. But Al turned down the peanuts because he is allergic: consuming peanuts causes a prickly rash, shortens his breath, and could even be fatal. Naucia turned down the peanuts because she ate a huge bag of peanuts at a circus years ago, and subsequently got nauseous from eating too much candy at the same time. Since then, her gustatory system associates peanuts with illness and she refuses them at reasonable prices. While Al and Naucia both revealed an identical disutility, a neurally detailed account tells us more. Al has an inelastic demand for peanuts—you can't *pay him* enough to eat them!—while Naucia would try a fistful for the right price. Their tastes will also change over time differently: Al's allergy will not be cured by repeated consumption, while Naucia's distaste might be easily changed if she tried peanuts once and didn't get sick.

Another example suggests how concepts of preference can be even wider of the mark by neglecting the nature of biological state-dependence: Nobody *chooses* to fall asleep at the wheel while driving. Of course, an imaginative rational-choice economist—or a satirist—could posit a tradeoff between "sleep utility" and "risk of plowing into a tree utility" and infer that a dead sleeper must have had higher u(sleep) than u(plowing into a tree). But this "explanation" is just tautology. It is more useful to think of the "choice" as resulting from the interaction of multiple systems—an automatic biological system which homeostatically shuts down the body when it is tired, and a controlled cognitive system which fights off sleep when closing your eyes can be fatal, and sometimes loses the fight.

For economists, it is natural to model these phenomena by assuming that momentary preferences depend on biological states. This raises a deep question of whether the cortex is aware about the nature of the processes and allocates cognitive effort (probably cingulate activity) to control them. For example, Loewenstein, O'Donoghue and Rabin (in press) suggest that people neglect mean-reversion in biological states, which explains stylized facts like suicide resulting from temporary depression, and shoppers buying more food when they are hungry.[4]

3. A third problem with preferences is that there are *different types* of utilities which do not always coincide. Kahneman (1994) distinguishes four types: remembered utility, anticipated utility, choice utility and experienced utility. Remembered utility is what people recall liking; anticipated utility is what they expect to like; choice utility is what they reveal by choosing (classical revealed preference); and experienced utility is what they actually like when they consume.

It is likely that the four types of utility are produced, to some extent, in separate brain regions. For example, Berridge and Robinson (1998) have found distinct brain regions for "wanting" and "liking", which correspond roughly to choice utility and experienced utility. The fact that these areas are dissociated allows a wedge between those two kinds of utility. Similarly, a wedge between remembered and experienced utility can be created by features of human memory which are adaptive for general purposes (but maladaptive for remembering precisely how something felt), such as repression of memories for severe pain in childbirth and other traumatic ordeals (e.g., outdoor adventures led by author GL).

If the different types of utility are produced by different regions, they will not always match up. Examples are easy to find. Infants reveal a choice utility by putting dirt in their mouths, but they don't rationally anticipate liking it. Addicts often report drug craving (wanting) which leads to consumption (choosing) that they say is not particularly pleasurable (experiencing). Compulsive shoppers buy goods (revealing choice utility) which they never use (no experienced utility). When decisions are rare, like getting pregnant, deciding whether to go to college, signing up for pension contributions, buying a house, or declaring war, there is no reason to think the four types of utility will necessarily match up. This possibility is important because it means that the standard analysis of welfare, which assumes that choices anticipate experiences, is incomplete.

In repeated situations with clear feedback, human learning may bring the four types of utilities together gradually. The rational choice model of consistent and coherent preferences can then be characterized as a limiting case of a neural model with multiple utility types, under certain learning conditions.

4. A fourth problem with preference is that people are assumed to value money for what it can purchase—that is, the utility of income is indirect, and

[4] Biological state-dependence also affects tipping. Most economic models suggest that the key variable affecting tipping behavior is how often a person returns to a restaurant. While this variable does influence tips slightly, a much stronger variable is how many alcoholic drinks the tipper had; see Conlin, Lynn and O'Donoghue (2003).

should be derived from direct utilities for goods that will be purchased with money. But roughly speaking, it appears that similar brain circuitry—dopaminergic neurons in the midbrain—is active for a wide variety of rewarding experiences—drugs, food, attractive faces, humor—and money rewards. This means money may be *directly* rewarding, and its loss painful. This might explain why workaholics and the very wealthy keep working long hours after they "should be" retired or cutting back (i.e., when the marginal utility of goods purchased with their marginal income is very low). Similarly, the immediate "pain of paying" can make wealthy individuals reluctant to spend when they should, and predicts unconventional effects of pricing—e.g. a preference for fixed payment plans rather than marginal-use pricing; see Prelec and Loewenstein (1998).

5. A common principle in economic modeling is that the utility of income depends only on the value of the goods and services it can buy, and is independent of the *source* of income. But Loewenstein and Issacharoff (1994) found that selling prices for earned goods were larger when the allocated good was earned than when it was unearned. Zink, Pagnoni, Martin-Skurski, Chappelow and Berns (2004) also found that when subjects earned money (by responding correctly to a stimulus), rather than just receiving equivalent rewards with no effort, there was greater activity in a midbrain reward region called the striatum. Earned money is literally more rewarding, in the brain, than unearned money. The fact that brain utility depends on the source of income is potentially important for welfare and tax policies.

6. Addiction is an important topic for economics because it seems to resist rational explanation. Becker and Murphy (1988) suggest that addiction and other changes in taste can be modeled by allowing current utility to depend on a stock of previous consumption. They add the assumption that consumers understand the habit formation, which implies that behavior responds to expected future prices.[5]

While variants of this model are a useful workhorse, other approaches are possible. It is relevant to rational models of addiction that every substance to which humans may become biologically addicted is also potentially addictive for rats. Addictive substances appear therefore to be "hijacking" primitive reward circuitry in the "old" part of the human brain. Although this fact does not disprove the rational model (since

[5] Evidence in favor of the rational-addiction view is that measured price elasticities for addictive goods like cigarettes are similar to those of other goods (roughly −0.5 and −2), and there is some evidence that current consumption does respond to expected future prices; cf. Gruber and Koszegi (2001) and Hung (2001). However, data limitations make it difficult to rule out alternative explanations (e.g., smokers may be substituting into higher-nicotine cigarettes when prices go up).

recently-evolved cortex may override rat-brain circuitry), it does show that rational intertemporal planning is not necessary to create the addictive phenomena of tolerance, craving and withdrawal. It also highlights the need for economic models of the primitive reward circuitry, which would apply equally to man and rat.

Another awkward fact for rational-addiction models is that most addicts quit and relapse regularly. And while rational addicts should buy drugs in large quantities at discounted prices, and self-ration them out of inventory, addicts usually buy in small packages; cf. Wertenbroch (1998). These facts suggest a struggle between a visceral desire for drugs and cortical awareness that drug use is a losing proposition in the long run; relapse occurs when the visceral desire wins the struggle.

It is also remarkable that repeated drug use conditions the user to expect drug administration after certain cues appear (e.g., shooting up in a certain neighborhood or only smoking in the car). Laibson (2001) created a pioneering formal model of cue-dependent use, showing that there are multiple equilibria in which cues either trigger use or are ignored. The more elaborate model of Bernheim and Rangel (in press), is a paradigmatic example of how economic theory can be deeply rooted in neuroscientific details. They assume that when a person is in a hot state they use drugs; in a cold state, whether they use is a rational choice. A variable S, from 0 to N, summarizes the person's history of drug use. When he uses, S goes up; when he abstains S goes down. They characterize destructively addictive drugs and prove that the value function is declining in the drug-use history variable S. By assuming the cold state reflects the person's true welfare, they can also do welfare analysis and compare the efficiency effects of policies like *laissez-faire*, drug bans, sin taxes and regulated dispensation.

Decision-making under Risk and Uncertainty

Perhaps the most rapid progress in neuroeconomics will be made in the study of risky decision-making. We focus on three topics: risk judgments, risky choice and probability.

Risk and ambiguity: In most economic analyses risk is equated with variation of outcomes. But for most people, risk has more dimensions (particularly emotional ones). Studies have long shown that potential outcomes which are catastrophic and difficult to control are perceived as more risky (controlling for statistical likelihood); see Peters and Slovic (2000). Business executives say risk is the chance of loss, especially a large loss, often approximated by semivariance (the variance of the loss portion of an outcome distribution); see Luce and Weber (1986), MacCrimmon and Wehrung (1986) and recent interest in "value-at-risk" measures in finance.

Fig. 2. Opening the brain at the Sylvian fissure (between temporal and frontal lobes) shows the insula cortex (frontal pole is on the right). *Illustration courtesy of Ralph Adolphs*

These properties are exemplified by the fear of flying (which is statistically much safer than driving) phobias and public outcry to dangers which are horrifying, but rare (like kidnappings of children and terrorist bombings). Since economic transactions are inherently interpersonal, emotions which are activated by social risks, like shame and fear of public speaking could also influence economic activity in interesting ways.

A lot is known about the neural processes underlying affective responses to risks; see Loewenstein, Hsee, Welch and Weber (2001). Much aversion to risks is driven by immediate fear responses, which are largely traceable to a small area of the brain called the amygdala; cf. LeDoux (1996). The amygdala is an "internal 'hypochondriac' " which provides "quick and dirty" emotional signals in response to potential fears. But the amygdala also receives cortical inputs which can moderate or override its responses.[6]

An interesting experiment illustrating cortical override begins with fear-conditioning—repeatedly administering a tone cue followed by a painful electric shock. Once the tone becomes associated in the animal's mind with the shock, the animal shows signs of fear after the tone is played, but before

[6] For example, people exhibit fear reactions to films of torture, but are less afraid when they are told the people portrayed are actors and asked to judge some unemotional properties of the films.

the shock arrives (the tone is called a "conditioned stimulus" *à la* Pavlov's famous salivating dogs). When the tone is played repeatedly but not followed by a shock, the animal's fear response is gradually "extinguished".

At this point, a Bayesian might conclude that the animal has simply "unlearned" the connection between the tone and the shock (the posterior probability $P(\text{shock}|\text{tone})$ has fallen). But the neural reality is more nuanced than that. If the shock is then readministered following the tone, after a long period of extinction, the animal immediately relearns the tone–shock relation and feels fear very rapidly.[7]

Furthermore, if the connections between the cortex and the amygdala are severed, the animal's original fear response to the tone immediately reappears. This means the fear response to the tone has not *disappeared* in the amygdala, it is simply being *suppressed* by the cortex.

Another dimension of risky choice is "ambiguity"—missing information about probabilities people would like to know but don't (e.g., the Ellsberg paradox). Using fMRI, Hsu and Camerer (2004) found that the insula cortex was differentially activated when people chose certain money amounts rather than ambiguous gambles. The insula (shown in Figure 2) is a region that processes information from the nervous system about bodily states—such as physical pain, hunger, the pain of social exclusion, disgusting odors and choking. This tentative evidence suggests a neural basis for pessimism or "fear of the unknown" influencing choices.

Risky choice: Like risk judgments, choices among risky gambles involve an interplay of cognitive and affective processes. A well-known study reported in Bechara, Damasio, Tranel and Damasio (1997) illustrates such collaboration. Patients suffering prefrontal damage (which, as discussed above, produces a disconnect between cognitive and affective systems) and normal subjects chose cards from one of four decks. Two decks had more cards with extreme wins and losses (and negative expected value); two decks had less extreme outcomes but positive expected value (EV), and subjects had to learn these deck compositions by trial-and-error. They compared behavior of normal subjects with patients who had damage to prefrontal cortex (PFC; which limits the ability to receive emotional "somatic markers" and creates indecision). Both groups exhibited similar skin conductance reactions (an indication of fear) immediately after large-loss cards were encountered.

[7] This is hard to reconcile with a standard Bayesian analysis because the same "likelihood evidence" (i.e., frequency of shock following a tone) which takes many trials to condition fear in the first part of the experiment raises the posterior rapidly in just one or two trials in the later part of the experiment. If the animal had a low prior belief that tones might be followed by shocks, this could explain slow updating in the first part. But since the animal's revealed posterior belief after the extinction is also low, there is no simple way to explain why updating is so rapid after the fear is reinstalled.

However, normal subjects learned to avoid those risky "bad decks" but the prefrontal-damage patients rapidly returned to the bad decks shortly after suffering a loss. In fact, even among normal subjects, those who were lowest in emotional reactivity acted more like the prefrontal patients; see Peters and Slovic (2000).

Homeostasis in the body implies that people will adapt to changes and, consequently, are more sensitive to changes than to absolute levels. Kahneman and Tversky (1979) suggest the same principle applies to gains and losses of money from a point of reference and, furthermore, that the pain of loss is stronger than the pleasure of equal-sized gains. Imaging studies show that gains and losses are fundamentally different because losses produce more overall activation and slower response times, and there are differences in which areas are active during gain and loss; see Camerer, Johnson, Rymon and Sen (1993) and Smith and Dickhaut (2002). Dickhaut, McCabe, Nagode, Rustichini and Pardo (2003) found more activity in the orbito-frontal cortex when thinking about gains compared to losses, and more activity in inferior parietal and cerebellar areas when thinking about losses. O'Doherty, Kringelbach, Rolls, Jornak and Andrews (2001) found that losses differentially activated lateral OFC and gains activated medial OFC. Knutson, Westdorp, Kaiser and Hommer (2000) found strong activation in mesial PFC on both gain and loss trials, and additional activation in anterior cingulate and thalamus during loss trials.

Single-neuron measurement by Schultz and colleagues, as reported in Schultz and Dickinson (2000), and Glimcher (2002) in monkeys has isolated specific neurons which correspond remarkably closely to familiar economic ideas of utility and belief. Schulz isolates dopaminergic neurons in the ventral tegmental "midbrain" and Glimcher studies the lateral inferior parietal (LIP) area. The midbrain neurons fire at rates which are monotonic in reward amount and probability (i.e., they "encode" reward and probability). The LIP neurons seem to encode expected value in games with mixed-strategy equilibria that monkeys play against computerized opponents.

An interesting fact for neuroeconomics is that *all* the violations of standard utility theories exhibited in human choice experiments over money have been replicated with animals. For example, in "Allais paradox" choices people appear to overweight low probabilities, give a quantum jump in weight to certain outcomes, and do not distinguish sharply enough between intermediate probabilities; see e.g. Prelec (1998). Rats show this pattern too, and also show other expected utility violations; see e.g. Battalio, Kagel and Green (1995). People also exhibit "context-dependence": whether A is chosen more often than B can depend on the presence of an irrelevant third choice C (which is dominated and never chosen). Context-dependence means people compare choices within a set rather than assigning separate numerical utilities. Honey-bees exhibit the same pattern; see Shafir, Waite and Smith (2002). The striking parallelism of choices across species suggests that the human neural circuitry for these decisions is "old", and perhaps specially adapted to the challenges all species face—foraging, reproduction and survival—but not necessarily consistent with rationality axioms.

Gambling: Economics has never provided a satisfactory theory of why people both insure *and* gamble. Including emotions and other neuroscientific constructs might help. Like drug addiction, the study of pathological gambling is a useful test case where simple theories of rationality take us only so far. About 1% of the people who gamble are "pathological"—they report losing control, "chasing losses", and harming their personal and work relationships; cf. National Research Council (1999). Pathological gamblers are overwhelmingly male. They drink, smoke and use drugs much more frequently than average. Many have a favorite game or sport they gamble on. Gambling incidence is correlated among twins, and genetic evidence shows that pathologicals are more likely to have a certain gene allele (D_2Al), which means that larger thrills are needed to get modest jolts of pleasure; see Comings (1998). One study shows that treatment with naltrexone, a drug that blocks the operation of opiate receptors in the brain, reduces the urge to gamble; see e.g. Moreyra, Aibanez, Saiz-Ruiz, Nissenson and Blanco (2000).[8]

Game Theory and Social Preferences

In strategic interactions (games), knowing how another person thinks is critical to predicting that person's behavior. Many neuroscientists believe there is a specialized "mind-reading" (or "theory of mind") area which controls reasoning about what others believe and might do.

Social preferences: McCabe, Houser, Ryan, Smith and Trouard (2001) used fMRI to measure brain activity when subjects played games involving trust, cooperation and punishment. They found that players who cooperated more often with others showed increased activation in Broadmann area 10 (thought to be one part of the mind-reading circuitry) and in the thalamus (part of the emotional "limbic" system). Their finding is nicely corroborated by Hill and Sally (2002), who compared normal and autistic subjects playing ultimatum games, in which a proposer offers a take-it-or-leave-it division of a sum of money to a responder. Autists often have trouble figuring out what other people think and believe, and are thought to have deficits in area 10. About a quarter of their autistic adults offered nothing in the ultimatum game, which is consistent with an inability to imagine why others would regard an offer of zero as unfair and reject it.

[8] The same drug has been used to successfully treat "compulsive shopping"; see McElroy, Satlin, Pope, Keck and Hudson (1991).

One of the most telling neuroscientific findings comes from Sanfey, Rilling, Aaronson, Nystrom, Leigh and Cohen's (2003) fMRI study of ultimatum bargaining. By imaging the brains of subjects responding to offers, they found that very unfair offers ($1 or $2 out of $10) differentially activated prefrontal cortex (PFC), anterior cingulate (ACC) and insula cortex. The insula cortex is known to be activated during the experience of negative emotions like pain and disgust. ACC is an "executive function" area which often receives inputs from many areas and resolves conflicts among them.[9]

After an unfair offer, the brain (ACC) struggles to resolve the conflict between wanting money (PFC) and disliking the "disgust" of being treated unfairly (insula). Whether players reject unfair offers or not can be predicted rather reliably (a correlation of 0.45) by the level of their insula activity. It is natural to speculate that the insula is a neural locus of the distaste for inequality or unfair treatment posited by recent models of social utility, which have been successfully used to explain robust ultimatum rejections, public goods contributions, and trust and gift-exchange results in experiments; see Fehr and Gächter (2000) and Camerer (2003, Ch. 2).[10]

In a similar vein, de Quervain, Fischbacher, Treyer, Schellhammer, Schynyder, Buck and Fehr (2004) used PET imaging to explore the nature of costly third-party punishment by players A, after B played a trust game with player C and C decided how much to repay. When C repaid too little, the players A often punished C at a cost to themselves. They found that when players A inflicted an economic punishment, a reward region in the striatum (the nucleus accumbens) was activated—"revenge tastes sweet". When punishment was costly, regions in prefrontal cortex and orbitofrontal cortex were differentially active, which indicates that players are responding to the cost of punishment.

Zak, Matzner and Kurzban (2003) explored the role of hormones in trust games. In a canonical trust game, one player can invest up to $10 which is tripled. A second "trustee" player can keep or repay as much of the tripled investment as they want. Zak *et al.* measured eight hormones at different points in the trust game. They find an increase in oxytocin—a hormone

[9] The ACC also contains "spindle cells"—large neurons shaped like spindles, which are almost unique to human brains; see Allman, Hakeem, Erwin, Nimchinsky and Hof (2001). These cells are probably important for the activities which distinguish humans from our primate cousins, such as language, cognitive control and complex decision-making.

[10] The fact that the insula is activated when unfair/offers are rejected shows how neuroeconomics can deliver fresh predictions: it predicts that low offers are less likely to be rejected by patients with insula damage, and more likely to be rejected if the insula is stimulated indirectly (e.g., by exposure to disgusting odors). We don't know if these predictions are true, but no current model would have made them.

which rises during social bonding (such as breast-feeding)—in the trustee if the first player "trusts" her by investing a lot.

Interesting evidence of social preferences comes from studies with monkeys. Brosnan and de Waal (2003) find that monkeys will reject small rewards (cucumbers) when they see other animals getting better rewards (grapes, which they like more). Hauser, Chen, Chen and Chuang (2003) also find that tamarins act altruistically toward other tamarins who have benefited them in the past. These studies imply that we may share many properties of social preference with monkey cousins.

Iterated thinking: Another area of game theory where neuroscience should prove useful is iterated strategic thinking. A central concept in game theory is that players think about what others will do, and about what others think *they* will do, and this reasoning (or some other process, like learning, evolution or imitation) results in a mutually consistent equilibrium in which each player guesses correctly what others will do (and chooses their own best response given those beliefs). From a neural view, iterated thinking consumes scarce working memory and also requires one player to put herself in another player's "mind". There may be no generic human capacity to do this beyond a couple of steps. Studies of experimental choices, and payoff information subjects look up on a computer screen, suggest 1–2 steps of reasoning are typical in most populations; cf. e.g. Costa-Gomes, Crawford and Broseta (2001), Johnson, Camerer, Sen and Tymon (2002), and see Camerer, Ho and Chong (2004).[11] Bhatt and Camerer (2004) find differential activation in the insula in players who are poor strategic thinkers, which they interpret as reflecting self-focus that harms strategizing.

V. Conclusions

Economics parted company from psychology in the early twentieth century after economists became skeptical that basic psychological forces could be measured without inferring them from behavior (and then, circularly, using those inferred forces to predict behavior). Neuroscience makes this measurement possible for the first time. It gives a new way to open the "black box" which is the building block of economic systems—the human mind.

More ambitiously, students are often bewildered that the models of human nature offered in different social sciences are so different, and often contradictory. Economists emphasize rationality; psychologists

[11] It is important to note, however, that principles like backward induction and computation of equilibrium can be easily taught in these experiments. That means these principles are not computationally difficult, *per se*, they are simply *unnatural*. In terms of neural economizing, this means these principles should be treated like efficient tools which the brain is not readily-equipped with, but which have low "marginal costs" once they are acquired.

emphasize cognitive limits and sensitivity of choices to contexts; anthropologists emphasize acculturation; and sociologists emphasize norms and social constraint. An identical question on a final exam in each of the fields about trust, for example, would have different "correct" answers in each of the fields. It is possible that a biological basis for behavior in neuroscience, perhaps combined with all-purpose tools like learning models or game theory, could provide some unification across the social sciences; cf. Gintis (2003).

Most economists we talk to are curious about neuroscience but skeptical of whether we need it to do economics. The tradition of ignoring the inside of the "black box" is so deeply ingrained that learning about the brain seems like a luxury we can live without. But it is inevitable that neuroscience will have *some* impact on economics, eventually. If nothing else, brain fMRI imaging will alter what psychologists believe, leading to a ripple effect which will eventually inform economic theories that are increasingly responsive to psychological evidence. Furthermore, since some neuroscientists are already thinking about economics, a field called neuroeconomics will arise whether we like it or not. So it makes sense to initiate a dialogue with the neuroscientists right away.

Economics *could* continue to chug along, paying no attention to cognitive neuroscience. But, to ignore a major new stream of relevant data is always a dangerous strategy scientifically. It is not as if economic theory has given us the final word on, e.g., advertising effectiveness, dysfunctional consumption (alcoholism, teenage pregnancy, crime), and business cycle and stock market fluctuations. It is hard to believe that a growing familiarity with brain functioning will not lead to better theories for these and other economic domains, perhaps surprisingly soon.

In what way might neuroscience contribute to economics? First, in the applied domain, neuroscience measurements have a comparative advantage when other sources of data are unreliable or biased, as is often the case with surveys and self-reports. Since neuroscientists are "asking the brain, not the person", it is possible that direct measurements will generate more reliable indices of *some* variables which are important to economics (e.g., consumer confidence, and perhaps even welfare).

Second, basic neuroeconomics research will ideally be able to link hypotheses about specific brain mechanisms (location, and activation) with unobservable intermediate variables (utilities, beliefs, planning ahead), and with observable behavior (such as choices). One class of fruitful tasks is those where some theories assume choice A and choice B are made by a common mechanism, but a closer neural look might suggest otherwise. For example, a standard assumption in utility theory is that marginal rates of substitution exist across very different bundles of goods (and, as a corollary, that all goods can be priced in money terms). But some tradeoffs are simply too difficult or morally repulsive (e.g., selling a body part). Elicited preferences often vary substantially with descriptions and procedures; e.g. Ariely, Loewenstein and Prelec (2003). Neuroscience might tell us precisely what a "difficult" choice or a "sacred preference" is, and why descriptions and procedures matter.[12]

A third payoff from neuroscience is to suggest that economic choices which are considered different in theory are using similar brain circuitry. For example, studies cited above found that insula cortex is active when players in ultimatum games receive low offers, when people choose ambiguous gambles or money, when people see faces of others who have cooperated with them, and in players who are poor strategic thinkers. This suggests a possible link between these types of games and choices which would never have been suggested by current theory.

A fourth potential payoff from neuroscience is to add precision to functions and parameters in standard economic models. For instance, which substances are cross-addictive is an empirical question which can guide theorizing about dynamic substitution and complementarity. A "priming dose" of cocaine enhances craving for heroin, for example; cf. Gardner and Lowinson (1991). Work on brain structure could add details to theories of human capital and labor market discrimination.[13] The point is that knowing which neural mechanisms are involved tell us something about the nature of the behavior. For example, if the oxytocin hormone is released when you are trusted, and being trusted sparks reciprocation, then raising oxytocin exogeneously could increase trustworthy behavior (if the brain doesn't adjust for the exogeneity and "undo" its effect). In another example, Lerner, Small and Loewenstein (in press) show that changing moods exogeneously changes buying and selling prices for goods. The basic point is that understanding the effects of biological and emotional processes like hormone

[12] Grether, Plott, Rowe, Sereno and Allman (2004) study a related problem—what happens in second-price Vickrey auctions when people learn to bid their valuations (a dominant strategy). They find that the anterior cingulate is more active before people learn to bid their values, which is a neural way of saying that bidding valuations is not transparent.

[13] It has been known for some time that brains rapidly and unconsciously ("implicitly") associate same-race names with good words ("Chip-sunshine" for a white person) and opposite-race names with bad words ("Malik-evil"); see e.g. McConnell and Leibold (2001). This fact provides a neural source discrimination which is neither a taste nor a judgment of skill based on race (as economic models usually assume). Opposite-race faces also activate the amygdala, an area which processes fear; cf. Phelps, O'Connor, Cunningham, Funayama, Gatenby, Gore and Banaji (2000). Importantly, implicit racial associations can be disabled by first showing people pictures of faces of familiar other-race members (e.g., showing Caucasians a picture of actor Denzel Washington). This shows that the implicit racial association is not a "taste" in the conventional economic sense (e.g. it may not respond to prices). It is a cognitive impulse which interacts with other aspects of cognition.

release and moods will lead to new types of predictions about how variations in these processes affect economic behavior.

In the empirical contracts literature there is, surprisingly, no adverse selection and moral hazard in the market for automobile insurance; cf. Chiappori, Abbring, Heckman and Pinquet (2001). But there is plenty of moral hazard in healthcare use and worker behavior. A neural explanation is that driving performance is both optimistic (everyone thinks they are an above-average driver, so poor drivers do not purchase fuller coverage) and automatic (and is therefore unaffected by whether drivers are insured) but healthcare purchases and labor effort are deliberative. This suggests that "degree of automaticity" is a variable that can be usefully included in contracting models.

Will it ever be possible to create formal models of how these brain features interact? The answer is definitely "Yes", because models already exist; cf. e.g. Benhabib and Bisin (2004), Loewenstein and O'Donoghue (2004) and Bernheim and Rangel (in press). A key step is to think of behavior as resulting from the interaction of a small number of neural systems—such as automatic and controlled processes, or "hot" affect and "cold" cognition, or a module that chooses and a module that interprets whether the choice signals something good or bad about underlying traits; see Bodner and Prelec (2003). While this might seem complex, keep in mind that economics is *already* full of multiple-system approaches. Think of supply and demand, or the interaction of a principal and an agent she hires. The ability to study these complex systems came only after decades of careful thought (and false modeling starts) and sharp honing by many smart economists. Could creating a general multiple-system model of the brain really be that much harder than doing general equilibrium theory?

References

Allman, J., Hakeem, A., Erwin, J., Nimchinsky, E. and Hof, P. (2001), The Anterior Cingulate Cortex: The Evolution of an Interface between Emotion and Cognition, *Annals of the New York Academy of Sciences 935*, 107–117.

Ariely, D., Loewenstein, G. and Prelec, D. (2003), Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences, *Quarterly Journal of Economics 118*, 73–106.

Battalio, R., Kagel, J. and Green, L. (1995), *Economic Choice Theory: An Experimental Analysis of Animal Behavior*, Cambridge University Press, New York.

Bechara, A., Damasio, H., Tranel, D. and Damasio. A. (1997), Deciding Advantageously Before Knowing the Advantageous Strategy, *Science 275*, 1293–1294.

Becker, G. and Murphy, K. (1988), A Theory of Rational Addiction, *Journal of Political Economy 104*, 675–700.

Benhabib, J. and Bisin, A. (2004), Modelling Internal Commitment Mechanisms and Self-control: A Neuroeconomics Approach to Consumption–Savings, working paper, New York University, February. http://www.econ.nyu.edu/user/benhabib/hypegeb41.pdf.

Bernheim, B. D. and Rangel, A. (in press), Addiction and Cue-conditioned Cognitive Processes, *American Economic Review*.
Berridge, K. C. and Robinson, T. E. (1998), The Role of Dopamine in Reward: Hedonics, Learning, or Incentive Salience?, *Brain Research Reviews 28* (3), 308–367.
Bhatt, M. and Camerer, C. (2004), Strategic Thinking in the Brain, working paper.
Bodner, R. and Prelec, D. (2003), The Diagnostic Value of Actions in a Self-signaling Model, in I. Brocas and J. Carillo, (eds.), *The Psychology of Economic Decisions*, Vol. 1, Oxford University Press, Oxford.
Brosnan, S. and de Waal, F. (2003), Monkeys Reject Unequal Pay, *Nature 425*, 297–299.
Cabanac, M. (1979), Sensory Pleasure, *Quarterly Review of Biology 54*, 1–29.
Camerer, C. F. (2003), *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton University Press, Princeton, NJ.
Camerer, C., Ho, T.-H. and Chong, J.-K. (2004), A Cognitive Hierarchy Theory of One-shot Games, *Quarterly Journal of Economics 119*, 861–898.
Camerer, C. F., Johnson, E., Rymon, T. and Sen, S. (1993), Cognition and Framing in Sequential Bargaining for Gains and Losses, in K. Binmore, A. Kirman and P. Tani (eds.), *Frontiers of Game Theory*, MIT Press, Boston, MA, 27–48.
Camerer, C., Loewenstein, G. and Prelec, D. (2004), Neuroeconomics: How Neuroscience Can Inform Economics. http://www.hss.caltech.edu/~camerer/neurojel22final0504.pdf.
Chiappori, P., Abbring, J., Heckman, J. and Pinquet, J. (2001), Testing for Adverse Selection versus Moral Hazard from Dynamic Data, unpublished, University of Chicago.
Comings, D. E. (1998), The Molecular Genetics of Pathological Gambling, *CNS Spectrums 3* (6), 20–37.
Conlin, M., Lynn, M. and O'Donaghue, T. (2003), The Norm of Restaurant Tipping, *Journal of Economic Behavior and Organization 52*, 297–321.
Costa-Gomes, M., Crawford, V. P. and Broseta, B. (2001), Cognition and Behavior in Normal-form Games: An Experimental Study, *Econometrica 69*, 1193–1235.
Damasio, A. R. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, G. P. Putnam & Sons, New York.
de Quervain, J.-F., Fischbacher, U., Treyer, Y., Schellhammer, M., Schynyder, U., Buck, A. and Fehr, E. (2004), The Neural Basis of Altruistic Punishment, working paper, University of Zurich.
Dickhaut, J., McCabe, K., Nagode, J. C., Rustichini, A. and Pardo, J. V. (2003), The Impact of the Certainty Context on the Process of Choice, *Proceedings of the National Academy of Sciences 100*, 3536–3541.
Ekman, P. (1992), *Telling Lies: Clues to Deceit in the Marketplace, Marriage, and Politics*, W. W. Norton, New York.
Fehr, E. and Gächter, S. (2000), Fairness and Retaliation: The Economics of Reciprocity, *Journal of Economic Perspectives 14* (3), 159–181.
Fletcher, P., Happe, F., Frith, U., Baker, C., Dolan, R., Frackowiak, R. and Frith, C. (1995), Other Minds in the Brain: A Functional Imaging Study of "Theory of Mind" in Story Comprehension, *Cognition 57*, 109–128.
Gardner, E. L. and Lowinson, J. (1991), Marijuana's Interaction with Brain Reward Systems: Update 1991, *Pharmacology Biochemistry and Behavior 40*, 571–580.
Gazzaniga, M. S. (1988), *Mind Matters*, Houghton Mifflin, Boston, MA.
Gilovich, T. (1991), *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*, Free Press, New York.
Gilovich, T., Vallone, R. and Tversky, A. (1985), The Hot Hand in Basketball: On the Misperception of Random Sequences, *Cognitive Psychology 17*, 295–314.
Gintis, H. (2003), Towards the Unity of the Behavioral Sciences. http://www-unix.oit.umass.edu/~gintis/.

Glimcher, P. (2002), *Decisions, Uncertainty and the Brain: The Science of Neuroeconomics*, MIT Press, Cambridge, MA.

Grether, D. M., Plott, C. R., Rowe, D. B., Sereno, M. and Allman, J. M. (2004), An fMRI Study of Selling Strategy in Second Price Auctions, Working Paper no. 1189, California Institute of Technology. http://www.hss.caltech.edu/SSPapers/wp1189.pdf.

Gruber, J. and Koszegi, B. (2001), Is Addiction "Rational"? Theory and Evidence, *Quarterly Journal of Economics 116* (4), 1261–1303.

Hauser, M., Chen, K., Chen, F. and Chuang, E. (2003), Give Unto Others: Genetically Unrelated Cotton-top Tamarin Monkeys Preferentially Give Food to Those Who Altruistically Give Food Back, *Proceedings of the Royal Society of London B 270*, 2363–2370.

Hill, E. and Sally, D. (2002), Dilemmas and Bargains: Theory-of-mind, Cooperation and Fairness, working paper, University College, London.

Hsu, M. and Camerer, C. F. (2004), Ambiguity-aversion in the Brain, working paper, California Institute of Technology.

Hung, A. (2001), A Theoretical and Empirical Study of Addiction, working paper, Carnegie-Mellon University.

Jevons, W. (1871), *The Theory of Political Economy*, 1st edn., Macmillan and Co., London and New York.

Johnson, E. J., Camerer, C. F., Sen, S. and Tymon, T. (2002), Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining, *Journal of Economic Theory 104* (1), 16–47.

Kahneman, D. (1994), New Challenges to the Rationality Assumption, *Journal of Institutional and Theoretical Economics 150*, 18–36.

Kahneman, D. and Peavler, W. S. (1969), Incentive Effects and Pupillary Changes in Association Learning, *Journal of Experimental Psychology 79*, 312–318.

Kahneman, D. and Tversky, A. (1979), Prospect Theory: An Analysis of Decision under Risk, *Econometrica 47*, 263–291.

Knutson, B., Westdorp, A., Kaiser, E. and Hommer, D. (2000), fMRI Visualization of Brain Activity during a Monetary Incentive Delay Task, *NeuroImage 12*, 20–27.

Laibson, D. (2001), A Cue-theory of Consumption, *Quarterly Journal of Economics 116* (1), 81–120.

LeDoux, J. E. (1996), *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, Simon & Schuster, New York.

Lerner, J. S., Small, D. A. and Loewenstein, G. (in press), Heart Strings and Purse Strings: Carry-over Effects of Emotions on Economic Transactions, *Psychological Science*.

Levenson, R. W. (1988), Emotion and the Autonomic Nervous System: A Prospectus for Research on Autonomic Specificity, in L. W. Hugh (ed.), *Social Psychophysiology and Emotion: Theory and Clinical Applications*, John Wiley, Chichester, 17–42.

Lieberman, M. D. (2000), Introversion and Working Memory: Central Executive Differences, *Personality and Individual Differences 28*, 479–486.

Lo, A. and Repin, D. (2002), The Psychophysiology of Real-time Financial Risk Processing, *Journal of Cognitive Neuroscience 14*, 323–339.

Loewenstein, G. (1996), Out of Control: Visceral Influences on Behavior, *Organizational Behavior and Human Decision Processes 65*, 272–292.

Loewenstein, G. and Issacharoff, S. (1994), Source-dependence in the Valuation of Objects, *Journal of Behavioral Decision Making 7*, 157–168.

Loewenstein, G. and O'Donoghue, T. (2004), Animal Spirits: Affective and Deliberative Influences on Economic Behavior, working paper, Department of Social and Decision Sciences, Carnegie-Mellon University.

Loewenstein, G., Hsee, C., Welch, N. and Weber, E. (2001), Risk As Feelings, *Psychological Bulletin 127* (2), 267–286.

Loewenstein, G., O'Donoghue, T. and Rabin, M. (in press), Projection Bias in Predicting Future Utility, *Quarterly Journal of Economics*.
Luce, R. D. and Weber, E. U. (1986), An Axiomatic Theory of Conjoint, Expected Risk, *Journal of Mathematical Psychology 30*, 188–205.
MacCrimmon, K. R. and Wehrung, D. A. (1986), *Taking Risks: The Management of Uncertainty*, Free Press, New York.
McCabe, K., Houser, D., Ryan, L., Smith, V. and Trouard, T. (2001), A Functional Imaging Study of Cooperation in Two-person Reciprocal Exchange, *Proceedings of the National Academy of Sciences 98* (20), 11832–11835.
McConnell, A. and Leibold, J. (2001), Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes, *Journal of Experimental Social Psychology 37*, 435–442.
McElroy, S. L., Satlin, A., Pope, H., Keck, P. E. and Hudson, J. (1991), Treatment of Compulsive Shopping with Antidepressants: A Report of Three Cases, *Annals of Clinical Psychiatry 3* (3), 199–204.
Moreyra, P., Aibanez, A., Saiz-Ruiz, J., Nissenson, K. and Blanco, C. (2000), Review of the Phenomenology, Etiology and Treatment of Pathological Gambling, *German Journal of Psychiatry 3*, 37–52.
National Research Council (1999), *Pathological Gambling*, National Academy Press, Washington, DC.
Nisbett, R. E. and Wilson, T. D. (1977), Telling More Than We Can Know: Verbal Reports on Mental Processes, *Psychological Bulletin 84* (3), 231–259.
O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Jornak, J. and Andrews, C. (2001), Abstract Reward and Punishment Representation in the Human Orbitofrontal Cortex, *Nature Neuroscience 4*, 95–102.
Peters, E. and Slovic, P. (2000), The Springs of Action: Affective and Analytical Information Processing in Choice, *Personality and Social Psychological Bulletin, 26* (12), 1465–1475.
Phelps, E., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C. and Banaji, M. R. (2000), Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activity, *Journal of Cognitive Neuroscience 12*, 1–10.
Prelec, D. (1998), The Probability Weighting Function, *Econometrica 66* (3), 497–527.
Prelec, D. and Loewenstein, G. (1998), The Red and the Black: Mental Accounting of Savings and Debt, *Marketing Science 17*, 4–28.
Romer, P. M. (2000), Thinking and Feeling, *American Economic Review 90* (2), 439–443.
Sanfey, A. G., Rilling, J. K., Aaronson, J. A., Nystrom, L. E. and Cohen, J. D. (2003), Probing the Neural Basis of Economic Decision Making: An fMRI Investigation of the Ultimatum Game, *Science 300* (5626), 1755–1758.
Schulz, W. and Dickinson, A. (2000), Neuronal Coding of Prediction Errors, *Annual Review of Neuroscience 23*, 473–500.
Shafir, S., Waite, T. A. and Smith, B. H. (2002), Context-dependent Violations of Rational Choice in Honeybees (Apis Mellifera) and Gray Jays (Perisoreus Canadensis), *Behavioral Ecology and Sociobiology 51*, 180–187.
Simons, D. J. and Chabris, C. F. (1999), Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events, *Perception 28*, 1059–1074.
Smith, K. and Dickhaut, J. (2002), Auction Structure, Heart Rate, and Neuroeconomic Behavior, working paper.
Viner, J. (1925), The Utility Concept in Value Theory and Its Critics, *Journal of Political Economy 33*, 369–387.
Wegner, D. M. and Wheatley, T. (1999), Apparent Mental Causation: Sources of the Experience of Will, *American Psychologist 54* (7), 480–492.

Wertenbroch, K. (1998), Consumption Self-control by Rationing Purchase Quantities of Virtue and Vice, *Marketing Science 17* (4), 317–337.
Whalen, P. J., Rauch, S. L., Etcoff, N., McInerney, S. C., Lee, M. and Jenike, M. A. (1998), Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity Without Explicit Knowledge, *Journal of Neuroscience 18*, 411–418.
Zak, P. (2004), Neuroeconomics. working paper, Claremont Graduate School.
Zak, P., Matzner, W. and Kurzban, R. (2003), The Neurobiology of Trust, working paper.
Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C. and Berns, G. S. (2004), Human Striatal Response to Monetary Reward Depends on Saliency, *Neuron 42*, 509–517.