Neuroeconomics:

How neuroscience can inform economics

Colin Camerer
Division HSS 228-77
Caltech
Pasadena CA 91125
Camerer@hss.caltech.edu

George Loewenstein
Dept Social & Decision Sciences
Carnegie-Mellon University
Pittsburgh PA 15213
Gl20@andrew.cmu.edu

Drazen Prelec
Sloan School of Management
MIT
Cambridge, MA 02138
dprelec@mit.edu

February 3, 2003. 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 8-9, 2000, and the Arizona conference in 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 1997-98. David Laibson's presentation at the Princeton conference was particularly helpful, as were comments and suggestions from the editors and John Dickhaut, and conversations with John Allman, Greg Berns, Jonathan Cohen, Angus Deaton, Dave Grether, David Laibson, Read Montague, Charlie Plott, Matthew Rabin, Peter Shizgal, and Steve Quartz.

> Who knows what I want to do? Who knows what anyone wants to do? How can you be sure about something like that? Isn't it all a question of brain chemistry, signals going back and forth, electrical energy in the cortex? How do you know whether something is really what you want to do or just some kind of nerve impulse in the brain. Some minor little activity takes place somewhere in this unimportant place in one of the brain hemispheres and suddenly I want to go to Montana or I don't want to go to Montana. (White Noise, Don DeLillo)

Introduction

When early neoclassical economists built economic theory on a foundation of individual behavior, the psychological model they adopted was an old one. They saw human behavior as the outcome of a process of *decision-making*, weighing costs and benefits of actions to maximize utility (i.e., happiness, a la Bentham). Economists of this era had doubts about the plausibility of utility maximization. Viner (1925, 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.

Economists were also disturbed by the fact that, because utility could not be measured objectively, it could not be used to predict behavior in an independent fashion. So they gave up. As Jevons [, 1871 #3931 commented,

> 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.

Since feelings were meant to predict behavior, but could only be assessed from behavior, economists realized that without direct measurement, feelings were useless intervening constructs.

In the 1940s, the concepts of ordinal utility and revealed preference eliminated the superfluous intermediate step of positing immeasurable feelings. Revealed preference theory simply equates unobserved preferences with observed choices. Circularity is avoided by assuming that people behave consistently, which makes the theory falsifiable; once they have revealed that they prefer A to B, people should not subsequently choose B over A. Like behaviorist psychology in the 1920s, which disdained reference to unobservable psychological "mentalist" constructs, ordinal utility and revealed preferences gave economists an easy way to avoid the messy reality of the psychology underlying utility. Later extensions— discounted, expected, and subjective expected utility, and Bayesian updating — provided similar "as if" tools which sidestep cognitive detail. Economists then spent decades developing mathematical techniques to make economic predictions without having to measure thoughts or feelings directly.

But now neuroscience, the study of the brain and nervous system, is beginning to allow direct measurement of thoughts and feelings, contrary to Jevons' pessimistic prediction. These measurements are, in turn, challenging our understanding of the relation between mind and action, leading to new theoretical constructs and calling old ones into question. How can the new findings of neuroscience, and the theories they have spawned, inform an economic theory that developed so impressively in their absence?

The standard economic theory of constrained utility maximization is most naturally interpreted as a model of careful deliberation – a balancing of the costs and benefits of different options -- as might characterize complex decisions like lifetime savings planning and delicate contract design. Although economists may acknowledge that actual flesh-and-blood human beings often choose without much deliberation, the economic models as written invariably represent decisions in a 'deliberative equilibrium,' i.e., that are at a stage where further deliberation, computation, reflection, etc. would not by itself alter the agent's choice.

While not denying that deliberation is always an option for human decision making, neuroscience research points to two generic inadequacies of this approach. First, much of the brain is constructed to support 'automatic' processes (Bargh, Chaiken, Raymond and Hymes 1996; Bargh and Chartrand 1999; Schneider and Shiffrin 1977; Shiffrin and Schneider 1977), which are faster than conscious deliberations and which occur with little or no awareness or feeling of effort. Because the person has little or no introspective access to, or volitional control over them, the behavior these processes generate need not conform to normative axioms of inference and choice (and hence cannot be adequately represented by the usual maximization models).

Second, our behavior is under the pervasive and often unrecognized influence of finely tuned affective (emotion) systems that are localized in particular brain regions and whose basic design humans share with many other animals (LeDoux 1996; Panksepp 1998; Rolls 1999). These systems are absolutely essential for daily functioning. When affective systems are damaged or perturbed, by brain damage, stress, imbalances in neurotransmitters, alcohol, or the 'heat of the moment,' the deliberative system generally is not capable of getting the job done alone.

As we discuss below in section III, behavior emerges from the interplay between controlled and automatic systems on the one hand, and between cognitive and affective systems on the other. Moreover, many behaviors that are clearly established to be caused by automatic or affective systems are *interpreted* by human subjects, spuriously, as the product of cognitive deliberation (Wolford, Miller and Gazzaniga 2000). The deliberative system, which is the system that is responsible for making sense of behavior, does not have perfect access to the output of the other systems, and exaggerates the importance of processes it understands when it attempts to make sense of the body's behavior.

Neuroscience findings and methods will undoubtedly play an increasingly prominent role in economics. Indeed, a brand of neuroeconomics shaped by neuroscientist is already emerging and attracting attention, whether economists approve of it or not (e.g. Glimcher 2002; Montague and Berns 2002). Participating in the development of a shared intellectual enterprise will help us ensure that the neuroscience informs economic questions we care about. Our goal in this paper is to describe what neuroscientists do and how their discoveries and views of human behavior might inform economic analysis. In the next section (II), we describe the diversity of tools that neuroscientists use. Section III introduces a simplified account of how the four modes of thinking just discussed work separately, and interact. Section IV discusses the implications of neuroscience for the economic analysis of intertemporal choice and risk. Section V concludes.

Neuroscience Methods

Scientific technologies are not just tools scientists use to explore areas of interest. New tools *define* new scientific fields, and erase old boundaries – e.g., the telescope (slipping away from speculative cosmology) created astronomy. The same is true of economics. Its boundaries have been constantly reshaped by tools such as mathematical, econometric, and simulation methods. Likewise, the current surge of interest in neuroscience by psychologists emerged largely from new methods. This section reviews some of these methods.

Brain imaging

Brain imaging is currently the most popular neuroscientific tool. Most brain imaging involves a comparison of people performing different tasks – an "experimental" task and a "control" task. The difference between images taken while subject is performing the two tasks provides an image of the regions of the brain that are differentially activated by the experimental task.

There are three basic imaging methods. The oldest, electro-encephalogram (or EEG) uses electrodes attached to the scalp to measure electrical activity synchronized to stimulus events or behavioral responses (known as Event Related Potentials, or ERPs). Like EEG, positron emission topography (PET) scanning is an old technique in the rapidly changing time-frame of neuroscience, but is still a useful technique. PET measures blood flow in the brain, which is a reasonable proxy for neural activity, since neural activity in a region leads to increased blood flow to that region. The newest, and currently most popular, imaging method is functional magnetic resonance imaging (fMRI), which tracks activity in the brain proxied by changes in blood oxygenation.

Although fMRI is increasingly becoming the method of choice, each method has its own advantages and disadvantages. EEG has excellent temporal resolution (on the order of 1 millisecond) and is the only method used with humans that directly monitors neural activity, as opposed to, e.g., blood flow. But spatial resolution is poor, and it can only measure activity in the outer part of the brain. EEG resolution has, however, been improving through the use of ever-increasing numbers of electrodes. For economics, a major advantage of EEG is its relatively unobtrusiveness and portability, which eventually may reach the point where it will be possible to take unobtrusive measurements from people as they go about their daily affairs. PET and fMRI provide better spatial resolution than EEG, but poorer temporal resolution because blood-flow to neurally active areas occurs with a stochastic lag in the range of 2-4 seconds.

Brain imaging still provides only a crude snapshot of brain activity. Neural processes are thought to occur on a 0.1 millimeter scale in 100 milliseconds (msec), but the spatial and temporal resolution of a typical scanner is only 3 millimeters and about two seconds. Multiple trials per subject can be averaged to form composite images, but doing so constrains experimental design. However, the technology has improved rapidly and will continue to improve. Hybrid techniques that combine the strengths of different methods are particularly promising.

Single-neuron measurement

Even the finest-grained brain imaging techniques only measure 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. As we discuss below, single neuron measurement studies have produced some striking findings that, we believe, are relevant to economics. A limitation of single neuron measurement is that, because insertion of the wires damages neurons, it is largely restricted to animals.

Studying animals is informative about humans because many brain structures and functions of non-human mammals are similar to those of humans (e.g., we are more genetically similar to many species of monkeys than those species are to other species). Neuroscientists commonly divide the brain into crude regions that reflect a combination of evolutionary development, functions, and physiology. The most common, *triune* division draws a distinction between the "reptilian brain," which is responsible for basic survival functions, such as breathing, sleeping, eating, the "mammalian brain," which encompasses neural units associated with social emotions, and the "hominid" brain, which is unique to humans and includes much of our oversized cortex -- the thin, folded, layer covering the brain that is responsible for such "higher" functions as language, consciousness and long-term planning (MacLean 1990). Because single neuron measurement is largely restricted to nonhuman animals, it has so far shed far more light on the basic emotional and motivational processes that humans share with other mammals than on higher-level processes such as language and consciousness.

Electrical brain stimulation (EBS)

Electrical brain stimulation is another method that is largely restricted to animals. In 1954, psychologists James Olds and Peter Milner (Olds and Milner 1954) discovered that rats would learn and execute novel behaviors if rewarded by brief pulses of electrical brain stimulation (EBS) to certain sites in the brain. Rats (and many other vertebrates, including humans) will work hard for EBS. For a big series of EBS pulses, rats will leap over hurdles, cross electrified grids, and forego their only daily opportunities to eat, drink, or mate. Animals also trade EBS off against smaller rewards in a sensible fashion – e.g., they demand more EBS to forego food when they are hungry. Unlike more naturalistic rewards, EBS does not satiate. And electrical brain stimulation at specific sites often elicits behaviors such as eating, drinking (Mendelson 1967), or copulation (Caggiula and Hoebel 1966). Many abused drugs, such as cocaine, amphetamine, heroin, cannabis, and nicotine, lower the threshold at which animals will lever-press for EBS (Wise 1996). Despite its obvious applications to economics, we know of only one EBS study by economists (Green and Rachlin 1991).

Psychopathology and brain damage in humans

Chronic mental illnesses (e.g., schizophrenia), developmental disorders (e.g., autism), and degenerative diseases of the nervous system 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. Parkinson's Disease (PD) initially affects the basal ganglia, then spreads to the cortex. The early symptoms of PD therefore provide clues about what the basal ganglia do (Lieberman, 2000).

Localized brain damage, produced by accidents and strokes, is also a rich source of insight, especially when damage is random (e.g.Damasio 1994). When patients with known damage to an area X perform a special task more poorly than "normal" patients, and do other tasks equally well, one can infer that area X is used to do the special task. Patients who have undergone neurosurgical procedures such as lobotomy (used in the past to treat depression) or radical bisection of the brain (an extreme remedy for epilepsy, now rarely used) have also provided valuable data (see Freeman and Watts 1950, Gazzaniga and LeDoux 1978).

Finally, a relatively new method called transcranial magnetic stimulation (TMS) uses pulsed magnetic fields to *temporarily* disrupt brain function in specific regions. The difference in cognitive and behavioral functioning that results from such disruptions provides clues about which regions control which neural functions. The theoretical advantage of TMS over brain imaging is that TMS directly leads to causal inferences about brain functioning rather than the purely associational evidence provided by imaging techniques. Unfortunately, the use of TMS is currently limited to the cortex and is somewhat controversial because it can causes seizures and may have other bad long-run effects.

Basic Lessons From Neuroscience

Because most of these techniques involve localization of brain activity, this can easily foster a misperception that neuroscience is merely developing a 'geography of the brain,' a map of which brain bits do what part of the job. If that were indeed so, then there would be little reason for economists to pay attention: As long as it is clear what the brain does, why does it matter, for economics at least, where it gets done? In reality, however, neuroscience is beginning to elucidate the principles of brain organization and functioning, which in turn are radically changing our estimate of *what* the brain is trying to do. A second misperception is that neuroscience is interested only in the more basic processes of motivation, perception, and action shared by humans and nonhumans, at the expense of higher functions found only in humans. As we shall see, neuroscience today is exploring the most subtle aspects of human social perception and cognition. It is as much a *social* as a biological science (Ochsner and Lieberman 2001).

Our goal in this section is to highlight some of the findings from neuroscience that we believe will prove most relevant to economics, emphasizing those that contrast most sharply with standard rational-choice models of optimization and equilibration. Our organizing theme, depicted in Table 1, emphasizes the two distinctions mentioned in the introduction, between *controlled* and *automatic* processes (Schneider and Shiffrin 1977),[1] and between *cognition* and *affect*.

---

[1] Lieberman and his colleagues Lieberman, M.D., Reflective and reflexive judgment processes: A social cognitive neuroscience approach, in press *in* "Responding to the social world: Implicit and explicit processes in social judgments and decisions" (J. P. Forgas, K. Williams and W. von Hippel, Eds.), Philadelphia Psychology Press, pp. Lieberman, M.D., R. Gaunt, D. T. Gilbert and Y Trope, Reflection and reflexion: A social cognitive neuroscience approach to attributional inference, 2002 *in* "Advances in Experimental Social Psychology" (M. Zanna, Ed.), New York Academic Press, pp. 199-249. have recently reformulated this venerable psychological distinction by tracing it to separate brain systems: the "reflexive or X-system" and the "reflective or C-system."

Table 1: A two-dimensional characterization of neural functioning

	Cognitive	Affective
Controlled Processes ▪ serial ▪ effortful ▪ evoked deliberately ▪ good introspective access	I	II
Automatic Processes ▪ parallel ▪ effortless ▪ reflexive ▪ no introspective access	III	IV

As described by the two rows of Table 1, *controlled* processes[2] tend to be serial (they use step-by-step logic or computations), tend to be invoked deliberately by the agent when her or she encounters a challenge or surprise (Hastie 1984), are often associated with a subjective feeling of effort, and typically occur consciously. Because controlled processing is conscious, people often have reasonably good introspective access to it. Thus, if people are asked how they solved a math problem or choose a new car, they can often provide a fairly accurate account of their choice process. Standard tools of economics, such as decision trees and dynamic programming, to the extent that they are actually used by individuals, epitomize controlled processes.

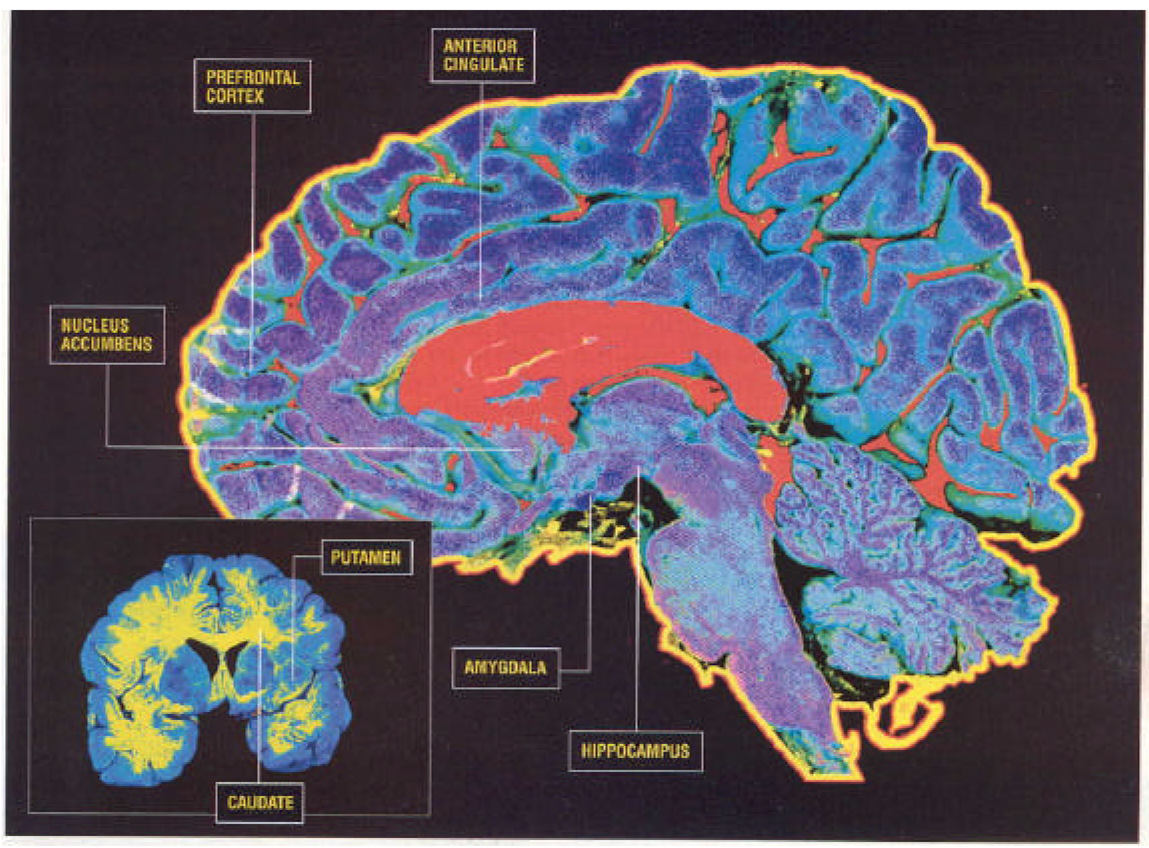
Automatic processes are the opposite of controlled processes on these dimensions; they operate in parallel, are not associated with any subjective feeling of effort, and operate outside of conscious awareness. As a result, people often have surprisingly little introspective access to why automatic choices or judgments were made. For example, a face is perceived as ‘attractive’, or a verbal remark as ‘sarcastic’, automatically and effortlessly. It’s only later that the controlled system reflects on the judgment and tries to substantiate it logically (and often does so spuriously, as discussed below (e.g., Wilson, Lindsey and Schooler 2000).

---

[2] The distinction between automatic and controlled processes was first proposed by Schneider and Shiffrin Schneider, Walter and Richard M. Shiffrin. 1977. "Controlled and automatic human information processing: I. Detection, search and attention." *Psychological Review*, 84:1, pp. 1-66.. Similar two-system models were later elaborated by many others (recently Sloman Sloman, Steven A. 1996. "The empirical case for two systems of reasoning." *Psychological Bulletin*, 119:1, pp. 3-22., and Lieberman et al. Lieberman, M.D., R. Gaunt, D. T. Gilbert and Y Trope, Reflection and reflexion: A social cognitive neuroscience approach to attributional inference, 2002 *in* "Advances in Experimental Social Psychology" (M. Zanna, Ed.), New York Academic Press, pp. 199-249.).

Automatic and controlled processes can be very roughly distinguished by where they occur in the brain (Lieberman, Gaunt, Gilbert and Trope 2002). Regions that support cognitive automatic activity are concentrated in the back (occipital), top (parietal) and side (temporal) parts of the brain (see Figure 1). The amygdala, buried below the cortex, is responsible for many important automatic affective responses, especially fear. Controlled processes occur mainly in the front (orbital and prefrontal) parts of the brain. The prefrontal cortex (pFC) is sometimes called the "executive" region, because it draws inputs from almost all other regions, integrates them to form near and long-term goals, and plans actions that take these goals into account (Shallice and Burgess, 1998). The prefrontal area is the region that has grown the most in the course of human evolution and which, therefore, most sharply differentiates us from our closest primate relatives (Manuck, Flory, Muldoon and Ferrell 2003).

Figure 1



Automatic processes — whether cognitive or affective — are the default mode of brain operation. They whir along all the time, even when we dream, constituting most of the electrochemical activity in the brain. Controlled processes occur at special moments when automatic processes become ‘interrupted’, which happens when a person encounters unexpected events, experiences strong visceral states, or is presented with some kind of explicit challenge in the form of a novel decision or other type of problem. To the degree that controlled processes are well described by economic calculation but parallel processes are not one could say that economics is about the 'interrupt' or 'override'. [3]

The second distinction, represented by the two columns of table 1, is between *cognitive* and *affective* processes. Such a distinction is pervasive in contemporary psychology (e.g., Zajonc 1998; Zajonc 1980; 1984; Zajonc and McIntosh 1992) and neuroscience (Damasio 1994; LeDoux 1996; Panksepp 1998), and has an historical lineage going back to the ancient Greeks and earlier (Plato described people as driving a chariot drawn by two horses, reason and emotion). Zajonc (1998) defines cognitive processes as those that answer true/false questions and affective processes as those that motivate approach/avoidance behavior. Affective processes include emotions such as anger, sadness, and shame, as well as "biological affects" (Buck 1999) such as hunger, pain, and the sex drive.

The quadrants in action

Quadrant I is in charge when you deliberate whether to refinance your house, poring over present-value calculations; quadrant II is used by “method actors” who imagine previous emotional experiences to fool audiences into thinking they are experiencing those emotions; quadrant III governs the movement of your hand as you return serve; and quadrant IV makes you jump when somebody says “Boo!”

Most behavior results from the interaction of all four quadrants of Table 1. Consider what happens when a party host approaches you with a plate of sushi.

Quadrant III: Your first task is to figure out what is on the plate. The occipital cortex in the back of the brain is the first on the scene, drawing in signals from your eyes via your optic nerves. It decodes the sushi into primitive patterns such as lines and corners (), then uses a “cascading process” to discern larger shapes (Kosslyn 1994). Further downstream, in the inferior temporal visual cortex (ITVC), this information becomes integrated with stored representations of objects, which permits you to recognize the objects on the plate as sushi. This latter process is

[3] As David Laibson and Andrew Caplin, respectively, aptly expressed it.

extraordinarily complicated (and has proved difficult for artificial intelligence researchers to recreate in computers) because the objects can take so many forms, orientations, and sizes.

Quadrant IV: This is where affect enters the picture. Neurons in the inferior temporal visual cortex are sensitive only to the identity of an object; they don't tell you whether it will taste good. Outputs of the inferior temporal visual cortex as well as outputs from other sensory systems feed into the orbitofrontal cortex to determine the food's reward value. Reward value depends on many factors.

First, there is your personal history with sushi. If you got sick on sushi in the past, you will have an unconscious and automatic aversion to it ("taste aversion conditioning"). The amygdala seems to play a critical role in this kind of long-term learning (LeDoux 1996). Second, the reward value of the sushi will depend on your current level of hunger; people can eat almost anything—grass, bugs, human flesh—if they are hungry enough. The orbitofrontal cortex and a subcortical region called the hypothalamus are sensitive to your level of hunger (Rolls 1999). Neurons in these regions fire more rapidly at the sight or taste of food when you are hungry, and less rapidly when you are not.

Quadrants I and II: Processing often ends before quadrants I and II go to work. If you are hungry, and like sushi, your motor cortex will guide your arm to reach for the sushi and eat it, drawing on automatic quadrant III (reaching) and IV (taste and enjoyment) processes. Under some circumstances, however, higher level processing may enter the picture. If you saw a recent documentary on the risks of eating raw fish, you may recoil; or if you dislike sushi but anticipate disappointment in the eyes of your proud host who made the sushi herself, you'll eat it anyway (or pick it up and discreetly hide it in a napkin when she turns to serve other guests).[4] These explicit thoughts involve anticipated feelings (your own and the host's) and draw on explicit memories

---

[4] Paul Romer Romer, Paul M. 2000. "Thinking and feeling." *American Economic Review*, 90:2, pp. 439-443. illustrates different neural mechanisms that lead to eschewing peanuts. One person loves the taste of peanuts, but is allergic to them and knows that the consequences of eating would be disastrous. When she is hungry, her visceral system motivates her to eat peanuts, but her deliberative system, with its ability to consider delayed consequences, inhibits her from eating them. The other person developed a "taste aversion" to peanuts many years ago, as a result of having gotten sick right after eating them. She knows at a cognitive level that the peanuts were not the cause of her sickness, but her visceral system overrules cognitive awareness. Revealed preference theory would stop at the conclusion that both women have disutility for peanuts. But the fact that the mechanisms underlying their preferences are different leads to predictable differences in other kinds of behavior. For example, the taste-averse woman will have a higher price elasticity (she'll eat peanuts if paid enough) and she will learn to enjoy peanuts after eating them a few times (her taste-aversion can be "extinguished"). The allergic woman will also like the smell of peanuts, which the taste-averter won't. Treatments also differ: Cognitive therapies for treating harmless phobias and taste-aversion train people to use

from a part of the brain called the hippocampus (see figure), inputs from the affective system (sometimes referred to as the "limbic system"), and anticipation (planning) from the prefrontal cortex.

Economics captures the parts of this process that are best described by Quadrant I. The rest of this section provides further details about automatic and affective processes in quadrants II-IV, and describes how processes in the four quadrants interact.

Automatic processes

Here, we review of some key principles of neural functioning that characterize automatic processes. Our short list includes: *parallelism*, *plasticity*, *modularity* and *specialization.* Unpacking this a bit, we would say that: (1) Much of the brain's processing involves processes that unfold in parallel and are not accessible to consciousness; (2) the brain undergoes physical changes as a result of these processes, (3) it draws upon multiple modules specialized to perform specific functions, and (4) it figures out how to use existing modules to accomplish new tasks efficiently, whatever functions they originally evolved to perform.

Parallism: Each neuron connects with thousands of other neighboring neurons in an intricate pattern. Parallelism, which is a hallmark of automatic processing, facilitates rapid response and gives the brain remarkable power when it comes to certain types of tasks, such as visual identification. "Connectionist" neural network models formulated by cognitive psychologists (Rumelhart and McClelland 1986) capture this feature of the brain and have been applied to many domains, including commercial ones. Models of this type have a very different structure than the systems of equations that economists typically work with. Unlike systems of equations, they are "black-box" — it is typically impossible to make intuitive sense of what they are doing by looking, for example, at individual parameters. By the same token, parallel processes are generally not accessible to consciousness.

Parallelism provides the brain with tremendous power because it allows for massive multitasking. It also provides redundancy that decreases the brain's vulnerability to harm. As a result of this redundancy, when neurons are progressively destroyed in a region, the consequences are typically gradual rather than sudden ("graceful degradation").

---

their conscious quadrant I processing to overrule visceral quadrant IV impulses, whereas treatment of the allergic peanut-avoider will concentrate on a medical cure.

Plasticity: Cognitive operations operate through electrochemical interactions between neurons. In a process first identified by the early neuroscientist Donald Hebb (1949), when signals are repeatedly conveyed from one neuron to another, the connections between those neurons strengthen. In much the same way that the muscles and organs of the body change as a function of the body's activity (and inactivity), the brain changes physiologically as a result of the operations that it performs. Thus, violinists who finger violin strings with their left hand show enlarged development of cortical regions which correspond to fingers on the left hand (Elbert, Pantev, Wienbruch, rockstroh and Taub 1995), and the brain regions responsible for navigation and spatial memory (the hippocampus) of London taxi drivers are larger than comparable areas in non-taxi drivers (Maguire, Gadian, Johnsrude, Good, Ashburner, Frackowiak and Frith 2000).

Plasticity, like parallelism, lessens the brain's vulnerability to environmental damage by allowing it to recover from damage like strokes and "rewire" itself, particularly early in life. In one study that illustrates the power of plasticity, the optic nerves of ferrets were disconnected at birth and reconnected to the auditory cortex (the portion of the brain that processes sound). The ferrets learned to "see" using auditory cortex, and some neurons in their auditory cortex actually took on the physical characteristics of neurons in the visual cortex (von Melchner, Pallas and Sur 2000).[5]

An important implication of plasticity is that information processing is unlikely to be reversible because the physiological processes that produce learning are themselves not reversible. Standard theories of information processing assume that people can ignore the effect of useless information or can undo the effect of information that is redundant or discredited. However, there are numerous demonstrations of violations of these principles. People tend to believe messages that are repeated, even if, on each repetition they recognize that the message is false (Gilbert and Gill 2000). When people form beliefs based on evidence that is later discredited definitively, the belief founded on the discredited evidence persists (Ross, Lepper and Hubbard 1975). And, people are subject to a ‘curse of knowledge’ such that, once they know something to be true (or false) they exaggerate the degree to which others must know it (Camerer, Loewenstein and Weber 1989).

Modularity: Neurons in different parts of the brain have different shapes and structures, different functional properties, and assemble themselves into relatively discrete modules that are functionally specialized. Progress in neuroscience often involves tracing well-known

[5] Ferrets were selected for this experiment because they, like humans, are born when they are still at a relatively immature state of development when the brain is still highly plastic.

psychological functions to circumscribed brain areas. For example, Broca's and Wernicke's areas are involved, respectively, in the production and comprehension of language. Thus, patients with Wernicke damage can speak sentences of correctly articulated words strung together into ungrammatical gibberish.

Beyond uncovering the modular structure of the brain, neuroscience has led to the discovery of new functional modules. Beyond the standard modules, such as face recognition, language and so on, research hints at the existence of some modules that are quite surprising. Thus, surgeons conducting brain surgery on an epileptic patient discovered a small region of her brain which, when stimulated, caused her to laugh (Fried 1998), hinting at the existence of a 'humor module'. Perhaps more surprisingly, neuroscientists have located an area in the temporal lobe that, when stimulated electrically, produces intense religious feelings -- e.g., the sense of a holy presence or even explicit visions of god or Christ, even in otherwise unreligious people (Persinger and Healey 2002). Of particular interest to economics, many neuroscientists believe there is a specialized 'mentalizing' (or 'theory of mind') module, which controls a person's inferences about what other people believe, feel, or might do. The first clues for such a dedicated module came from tests conducted by developmental psychologists in which two children are shown an object in the process of being concealed (seeFrith 2001a). One child then leaves, and the other child observes as the object is moved to a new location. The child who remains in the room is then asked to predict where the child who left will look for the object when she returns. Normal children are typically able to solve this problem around age three or four. Autistic children as a rule master this distinction much later (8-12 years), and with great difficulty, although some (especially those with 'Asperger syndrome') have normal or superior intelligence. At the same time, the autistic child will have no difficulty with general inferences having a similar logical form (e.g., if a photo is taken of the object's location, and the object is then moved, they will correctly infer that the photo will register the object in the old place, before it was moved).

Autistic adult individuals may compensate in many ways and eventually pass such basic tests of mentalizing. However, they have difficulty appreciating more subtle social meanings (e.g., irony), and will sometimes wonder at the "uncanny ability" of non-autistic persons to "read minds" (Frith 2001b). Individuals with Asperger's syndrome, presented with problems that involving mentalizing show lower activation in the medial prefrontal regions as compared with normals, but greater activation in an adjacent, but more ventral (lower) region of the prefrontal cortex, which is normally responsible for general reasoning (Happe, Ehlers, Fletcher, Frith, Johansson, Gillberg, Dolan, Frackowiak and Frith 1996). A natural interpretation is that Asperger individuals eventually deduce the answer through a complex process of reasoning instead of grasping it directly by a specialized module. In the medical literature, one can likewise find patients with brain lesions who exhibit difficulties with mentalizing tasks but not with other types of cognition (Rowe et al., 2001; Blair and Cipoletti, 2000). This, too, is consistent with the hypothesis of a separable mentalizing module.

The possibility of a mentalizing module has gained credibility and substance through converging neuroscientific evidence. fMRI studies have shown that when normal adults are given pairs of closely matching judgment problems, differing only in whether they do or do not require mentalizing, the mentalizing problems lead to greater activation in the left medial prefrontal cortex (Fletcher, Happe, Frith, Baker, Dolan, Frackowiak and Frith 1995). As ultimate proof, one would like to identify neuronal populations that are specifically turned on by mentalizing activity. Neuroscience is not there yet, but recent single-cell recordings in monkeys have identified an intriguing class of "mirror neurons" in the prefrontal cortex, which fire either when an experimenter performs a physical action (e.g., grasping a peanut) or when the monkey performs (“mirrors”) the same action. Having such neurons makes learning by imitation easy and supports mind reading by, for example, internally simulating the facial expressions of others.

What is the relevance of this to economics? Economic theory plainly assumes that agents can ‘mentalize’, i.e., make inferences from the actions of others to their underlying preferences and beliefs. Such mentalizing inferences sustain a Bayesian equilibrium, and are not normally regarded any more or less difficult than any other types of inferences. The first, and most elementary implication of the neuroscience evidence would be to put a “flashing yellow” light above this assumption. Mentalizing is indeed a special ability, and, as we have seen, general logical-deductive reasoning can only partially compensate for its absence. Furthermore, it would not be surprising to find normal individuals differing in their mentalizing skills. Indeed, in one of the first imaging studies conducted by economists, McCabe et al. (2001) theorized that mentalizing was important in games involving trust and cooperation. They found that players who were more trusting and cooperative showed more brain activity in Brodmann area 10 (thought to be the locus of mentalizing) and more activity in the limbic system which processes emotions.

Specialization: In a process that is not well understood, the brain figures out how to do the tasks it is assigned, efficiently, using the modules it has at its disposal. When the brain is confronted with a new problem it initially draws heavily on diverse modules, including, often, the prefrontal cortex. But over time, activity becomes more streamlined, concentrating in modules that specialized in processing relevant to the task. In one study, subjects' brains were imaged as they played the computer game Tetris, which requires rapid hand-eye coordination and spatial reasoning. When subjects began playing, they were highly aroused and many parts of the brain were active. As they got better at the game, overall blood-flow to the brain decreased markedly, and activity became localized in only a few brain regions. Much as an economy ideally adjusts to the introduction of a new product by gradually shifting production to the firms that can produce the best goods most cheaply, with experience at a task or problem, the brain seems to gradually shift toward modules that can solve problems automatically and efficiently with low-effort.

It naturally follows, that, for a wide range of problems and tasks, people will rely on cognitive capabilities — modules — that are relatively well developed, such as visual perception and object recognition rather than operations that we are not very good at, like decomposing and then summing up costs and benefits. In one now famous study, Gobet and Simon (1996) tested memory for configurations of chess pieces positioned on a chess board. They found that expert chess players were able to store the positions of players almost instantly – but only if they were in positions corresponding to a plausible game. For randomly arranged chess pieces, the experts were not much better than novices. More recent research in decision making suggests that this is a far more general phenomenon; much decision making takes the form of pattern matching rather than of an explicit weighing of costs and benefits (e.g., Leboeuf 2002; Medin and Bazerman 1999)

Automaticity and human capabilities: Economics implicitly assumes that people have general cognitive capabilities that can be applied to any type of problem, and hence that they will perform equivalently on problems that have similar structure. Automaticity, in contrast, suggests that performance will depend critically on whether a particular problem can be, and is in fact, processed by a module that is well adapted to that form of processing.

When a specialized module exists and is applied to a particular task, processing is rapid and the task feels relatively effortless. Automatic processes involved in vision, for example, are lightning fast and occur with no feeling of mental effort, so people aren't aware how powerful and sophisticated are the processes that allow it to happen. Even the most powerful computers still don't hold a candle to humans when it comes to visual perception. When we lack such tailored modules, however, we are likely to seem extraordinarily flat-footed because we will be forced to 'muscle it out' with quadrant 1 processing, much as autistic individuals seem to solve theory of mind problems. As a general rule, we should expect people to be geniuses when presented with problems that can be, and are, processed by dedicated modules, but relatively obtuse when they are forced to rely on controlled processes.[6]

A neat illustration of this is provided by the "Wason four-card problem" in logic. Subjects are shown four cards, each with a letter on one side and a number on the other. The exposed card faces read "X," "Y," "1", and "2." Subjects are asked which cards would need to be turned over to test the rule: "If there is an X on one side there is a 2 on the other." Few subjects give the right answers, which are X and 1 (if there's an X on the opposite side of the "1" the rule is broken). However, most subjects give the right answer when the logically equivalent problem is put in a cheating-detection frame. For example, if there are four children from two different towns and two school districts and the rule is "If a child lives in Concord he goes to Concord High", most subjects realize that the home address of the student who goes to Concord High must be checked to see if she is cheating (Cosmides 1989).

Automaticity and economic constructs: Knowing how the brain solves problems – and what specialized modules it has at its disposal to do so – will change our understanding of how people differ from one-another when it comes to economic behavior. Economists currently approach human behavior with constructs such as 'time preference', 'risk preference', and 'altruism' that are seen as stable over time and consistent across activities. But empirical evidence shows that risk-taking, time discounting and altruism are very weakly correlated (often zero) across situations. This inconsistency may result from the fact that we have the wrong set of constructs for understanding how people differ from one-another. Thus, as we discuss below, while there is no module that directly corresponds to time preference, there may be modules that are responsible for different dimensions of time preference – e.g., for inhibiting emotion-driven behavior and for thinking about future consequences.

---

[6] Thus, even high-functioning adults do only one or two steps of strategic thinking, in a wide variety of games Camerer, Colin F. 2003. *Behavioral game theory: Experiments on strategic interaction*. Princeton: Princeton University Press.
Johnson, Eric J., Colin F. Camerer, Sen Sankar and Talia Tymon. 2002. "Detecting failures of backward induction: Monitoring information search in sequential bargaining." *Journal of Economic Theory,* 104:1, pp. 16-47.. This

Automaticity and context and framing effects: The combination of automatic sense-making and specialized modules provides a neural language for understanding the effects of framing and context. Framing and context effects occur when contextual or descriptive factors affect choices, even when the outcomes of those choices are the same (c.f. Camerer and Loewenstein in press). Hershey, Kunreuther, and Schoemaker (1982), for example, found that choice between the same pair of certain and risky prospects was substantially affected by whether the decision was expressed as a "gamble," in which case people were risk-seeking, or as "insurance," in which case they were risk averse. In ultimatum bargaining games, players receiving an unequal division of a fixed “pie” are more likely to accept it if it was described as a high price offered by a monopoly seller (Camerer 2003, chapter 2; Hoffman, McCabe, Shachat and Smith 1994). Expert physicians’decisions about whether to treat cancer with radiation and surgery varied depending on whether the outcomes of those treatments were expressed as rates of death or rates of survival (McNeil, Pauker, Sox and Tversky 1982).

If decisions are made through pattern-matching and specialized modules, then subtle contextual features can lead to sharp changes in behavior that are not justified by a change in economic incentives. For example, parents at a day care center pick up their children *later* after a small fine is instituted for late pickups (seeGneezy and Rustichini in press; cfTitmuss 1970, on blood donation). When there is no fine, picking up your children promptly is construed as honoring a moral obligation to the daycare center. But when a fine is charged, parents construe their decision as whether it is worth paying a small price to be late— and many decide that it is.

Affective processes

The fashion in which the brain evolved is critical to understanding human behavior. In many domains, such as eating, drinking, sex, and even drug use, human behavior resembles that of our close mammalian relatives, which is not surprising because we share many of the neural mechanisms that are largely responsible for these behaviors. Many of the processes that occur in these systems are affective rather than cognitive; they are directly concerned with motivation. This might not matter for economics were it not for the fact that the principles that guide the affective

---

evidence suggests that there is no all-purpose module for iterating strategic reasoning endlessly, contrary to the implicit prediction in the idea that game-theoretic equilibria arise from introspection regardless of a game’s complexity.

system – the way that it operates – is so much at variance with the standard economic account of behavior.

Homeostasis: To understand how the affective system operates, one needs to recall that humans did not evolve to be happy, but to survive and reproduce. An important process by which the body attempts to achieve these goals is called *homeostasis*. Homeostasis involves detectors that monitor when a system departs from a 'set-point',[7] and mechanisms that restore equilibrium when such departures are detected. Some – indeed most – of these mechanisms do not involve deliberate action. Thus, when the core body temperature falls below the 98.6F set-point, blood tends to be withdrawn from extremities, and when it rises above the set-point one begins to sweat. But other processes do involve deliberate action – e.g., putting on one's jacket when cold or turning on the air conditioner when hot. The brain motivates one to take such actions using both a “carrot” and a “stick.” The stick reflects the fact that departing from a set-point usually feels bad – e.g., it feels bad to be either excessively hot or cold – and this negative feeling motivates one to take actions that move one back toward the set-point. The carrot is a process called "alliesthesia" (Cabanac 1979) whereby actions that move one toward the set-point tend to feel pleasurable. Thus, when one’s core body temperature falls below 98.6F, almost anything that raises body temperature (such as placing one's hand in warm water) feels good, and conversely when the body temperature is elevated almost anything that lowers body temperature feels good.

The role of homeostasis in human behavior poses a fundamental challenge to the economic account of behavior. As economists, we are used to thinking of preferences as the starting point for human behavior and behavior as the ending point. A neuroscience perspective, in contrast, views explicit behavior as only one of many mechanisms that the brain uses to maintain homeostasis, and preferences as transient state variables that ensure survival and reproduction. The traditional economic account of behavior, which assumes that humans act so as to maximally satisfy their preferences, starts in the middle (or perhaps even towards the end) of the neuroscience account.

---

[7] The mechanisms that monitor the body’s state can be exquisitely complex, and can include both internal and external cues. In the case of nutiritional regulation, for example, internal cues include gastric distension (Gibbs et al., 1981), receptors sensitive to the chemical composition of the food draining from the stomach Greenberg, D., G.P. Smith and J. Gibbs. 1990. "Intraduodenal infusions of fats elicit satiety in sham-feeding rats." *The American Journal of Physiology*, 259:1, pp. R110-R118. and glucose levels in the blood Campfield, L.A. and F.J. Smith, Systemic factors in the control of food intake, 1990 *in* "Handbook of behavioural neurobiology, Vol. 10, Neurobiology of food and fluid intake" (E. M. Stricker, Ed.), New York Plenum, pp. 183-206.. External cues include time of day, estimated time till the next feeding, and the sight or smell of food.

Rather than viewing pleasure as the goal of human behavior, a more realistic account would view pleasure as a homeostatic cue — an informational signal.[8,9]

An important feature of many homeostatic systems is that they are highly attuned to *changes* in stimuli rather than their *levels*. A dramatic demonstration of such sensitivity to change came from single-neuron studies of monkeys responding to juice rewards (see Schulz and Dickinson 2000). They measured the firing of dopamine neurons in the animal's ventral striatum, which is known to play a powerful role in motivation and action. In their paradigm, a tone was sounded, and two seconds later a juice reward was squirted into the monkey's mouth. Initially, the neurons did not fire until the juice was delivered. Once the animal learned that the tone forecasted the arrival of juice two seconds later, however, the same neurons fired at the sound of the tone, but *did not* fire when the juice reward arrived. These neurons are responding, not to reward or its absence, but to deviations from expectations. (When the juice was expected from the tone, but was not delivered, the neurons fired at a very low rate, as if expressing disappointment.) The same pattern can be observed at a behavioral level in animals, who will work harder (temporarily) when a reinforcement is suddenly increased and go `on strike' when reinforcement falls.[10] Neural sensitivity to change is probably important in explaining why the evaluation of risky gambles depends on a reference point which encodes whether an outcome is a gain or a loss (see section IV), why self-reported happiness (and behavioral indicators like suicide) depend on changes in income and wealth, rather than levels (Oswald 1997), and why violations of expectations trigger powerful emotional responses (Mandler 1982).

---

[8] Even deliberative behavior generated by quadrant I is typically organized in a fashion that resembles homeostasis (Loewenstein, mountain). As proposed in Miller, Gallanter and Pribram's Miller, George A., Eugene Galanter and Karl H. Pribram. 1960. *Plans and ther structure of behavior*. New York: Holt. classic Plans and the Structure of Behavior, most human volitional behavior takes the same form of goal setting and goal seeking. The only difference between such goal-driven behavior and homeostatically driven behaviors such as eating is that the former goal state is, in some sense, self-chosen.

[9] Of course, even the neuroscience account begins, in some sense, in the middle. A more complete understanding of behavior would also ask how these different mechanisms evolved over time. Since evolution selects for genes that survive and reproduce, the result of evolution is unlikely to maximize pleasure or minimize pain.

[10] As Rolls Rolls, E.T. 1999. *The brain and emotion*. New York: Oxford University Press. writes, "We are sensitive to some extent not just to the absolute level of reinforcement being received, but also to the change in the rate or magnitude of reinforcers being received. This is well shown by the phenomena of positive and negative contrast effects with rewards. Positive contrast occurs when the magnitude of a reward is increased. An animal will work very much harder for a period (perhaps lasting for minutes or longer) in this situation, before gradually reverting to a rate close to that at which the animal was working for the small reinforcement. A comparable contrast effect is seen when the reward magnitude (or rate at which rewards are obtained) is reduced - there is a negative overshoot in the rate of working for a time".

Raw motivation?: Economists usually view behavior as a search for pleasure (or, equivalently, escape from pain). But, there is considerable evidence from neuroscience and other areas of psychology that the motivation to take an action is not always closely tied to hedonic consequences.

Ken Berridge (1996), a neuroscientist at the University of Michigan, argues that decision making involves the interaction of two separate, though overlapping systems, one responsible for pleasure and pain (the "liking" system), and the other for motivation (the "wanting" system). This challenges the fundamental supposition in economics that one only strives to obtain what one likes. Berridge finds that certain lesions and pharmacological interventions can selectively enhance a rat's willingness to work for a food, without changing the pleasure of eating the food or the desire to work in general. How can you know how much pleasure food provides, independently of how much the rat is willing to labor for it? Fortunately, animal facial expressions reveal with some precision whether something tastes good, bad or indifferent. In economic language, the experiments create a situation where the utility of food and disutility of work remain the same, but the amount of work-for-reward goes up. This implies that it is possible to be motivated to take actions that bring no pleasure.

Berridge believes that the later stages of many drug addictions presents prototypical examples of situations of what he terms "wanting" without "liking;" drug addicts often report a complete absence of pleasure from taking the drug they are addicted to, coupled with an irresistible motivation to do so. In fact, it's hard to imagine the modern psychotherapeutic service industry being what it is if human beings never experienced the problem of 'wanting without liking.' Other examples of situations in which there often seems to be a disconnect between one's motivation to obtain something and the pleasure one is likely to derive from it are sex and curiosity (seeLoewenstein 1994).

Economics proceeds on the assumption that satisfying people's wants is a good thing. While this is probably generally a safe assumption, if wanting and liking are two separate processes, then it cannot be assumed that satisfying someone's desires necessarily makes them better off. Welfare economics, then, would need to be augmented by an analysis of when and why wanting and liking do, sometimes, diverge.

Interactions between the systems

Behavior emerges from a continuous interplay between neural systems supporting activity within each of the four quadrants. Three aspects of this interaction bear special emphasis, which we labeled "collaboration," "competition" and "sense-making." Collaboration captures the insight that decision-making, which is to say “rationality” in the broad, nontechnical sense of the word, is not a matter of shifting decision-making authority from Quadrants II, III, and IV toward the deliberative, non-affective Quadrant I, but more a matter of maintaining proper collaboration in activity across all four quadrants. If Quadrant I tries to do the job alone, it will fail, often spectacularly. Competition reflects the fact that, in fact, different processes – most notably affective and cognitive – often drive behavior in conflicting directions and compete for control of behavior. Sense-making refers to how we make sense of such collaboration and competition – how we make sense of our behavior. While behavior is, in fact, determined by the interaction of all four quadrants, conscious reflection on our behavior, and articulating reasons for it, is basically quadrant I trying to make sense of that interaction. And, not surprisingly, quadrant I has a tendency to explain behavior in terms it can understand – in terms of quadrant I processes.

Collaboration: Although it is heuristically useful to distinguish between cognitive and affective processes, and between controlled and automatic processes, most judgments and behaviors result from interactions between them. Collaboration, delegation of activity, and proper balance across the quadrants are essential to normal decision making. For example, since quadrant I processes are neurally expensive, thinking too hard when automatic processes work well is inefficient.

Affect is, and should be, influenced by cognition. The way you feel while waiting impatiently for a friend who is late, for example, will depend critically on whether you think they had an accident en route or simply forgot about you. Even affective states such as pain, which are commonly viewed as physiological rather than psychological, have an important cognitive component: Holding tissue damage constant, how much misery one experiences depends powerfully on factors such as what caused the pain (less misery if it was for a good cause), what the pain means (greater misery if it signals something significant, like cancer), and how long one expects the pain to last.

More interesting, and less well known, is the fact that affect serves as an essential input into decision making. The affective system provides inputs into deliberative decision making in the form of crude affective evaluations of behavioral options – what Damasio (1994) refers to as "somatic markers." Damasio and his colleagues show that individuals with minimal cognitive, but major emotional, deficits have difficulty making decisions, and often make poor decisions when they do (Bechara, Damasio, Damasio and Anderson 1994; Bechara, Damasio, Damasio and Lee 1999; Damasio 1994). Affect is so useful that sometimes thinking too hard can lead to inferior choices. Timothy Wilson and his colleagues show that introspecting about one's reasons for preferring a particular choice object blocks access to one's emotional reactions and reduces the quality of decision making (Wilson, Lisle, Schooler, Hodges, Klaaren and LaFleur 1993; Wilson and Schooler 1991). In one study (Wilson, Lisle, Schooler, Hodges, Klaaren and LaFleur 1993) college students selected their favorite poster from a set of posters. Those who were instructed to think of reasons why they liked or disliked the posters before they chose one ended up, on average, less happy with their choice of posters (and less likely to put it on their dorm room wall) than subjects who were not asked to provide reasons.

Needless to say, affect can also distort cognitive judgments. For example, emotions have powerful effects on memory– e.g., when people become sad, they tend to recall sad memories (which often increases their sadness). Emotions also affect perceptions of risks – anger makes people less threatened by risks, and sadness makes them more threatened (Lerner and Keltner 2001). Emotions also create "motivated cognition"-- people are good at persuading themselves that what they would *like* to happen is what *will* happen. Quack remedies for desperate sick people, and get-rich-quick scams are undoubtedly aided by the human propensity for wishful thinking. Wishful thinking may also explain high rates of new business failure (Camerer and Lovallo 1999), trading in financial markets, undersaving, and low rates of investment in education (foregoing large economic returns). As LeDoux (1996, p. 19) writes, "While conscious control over emotions is weak, emotions can flood consciousness. This is so because the wiring of the brain at this point in our evolutionary history is such that connections from the emotional systems to the cognitive systems are stronger than connections from the cognitive systems to the emotional systems."

Many decision-making disorders may originate in an improper division of labor between automatic and controlled process, or through an excess or a deficit of affective input. Psychiatry recognizes a decision making continuum defined by the impulsive, "light" decision-making style at one end and the compulsive, "heavy" style at the other. The decisions of an impulsive individual are excessively influenced by external stimuli, pressures, and demands. Such a person may not be able to give a more satisfying explanation of an action except that 'he felt like it' or because 'someone asked' him to do it (Shapiro, 1965). By contrast, an obsessive-compulsive person will subject even the most trivial decisions to extensive deliberation and calculation.

Competition: When it comes to spending money or delaying gratification, taking or avoiding risks and behaving kindly or nastily toward other people, people often find themselves of "two minds"; our affective systems drive us in one direction and cognitive deliberations in another. We find ourselves almost compulsively eating our children's left-over Halloween candy, while obsessing about how to lose the extra 10 pounds; spending furiously at the casino, even as a small voice in our head tell us that we are behaving self-destructively; longing to buy collision insurance even though we know that it's a rip-off; trying to build up courage to step up to the podium; or powerfully tempted to give to the street-corner beggar even though we know that our dollar will go further if donated to United Way.

All of these deviations occur because our affective system responds to different cues, and differently to the same cues, as our cognitive system does. As Rolls (1999, 282, 282) writes,

> emotions often seem very intense in humans, indeed sometimes so intense that they produce behaviour which does not seem to be adaptive, such as fainting instead of producing an active escape response, or freezing instead of avoiding, or vacillating endlessly about emotional situations and decisions, or falling hopelessly in love even when it can be predicted to be without hope or to bring ruin. The puzzle is not only that the emotion is so intense, but also that even with our rational, reasoning, capacities, humans still find themselves in these situations, and may find it difficult to produce reasonable and effective behaviour for resolving the situation.

Such divergences between emotional reactions and cognitive evaluations arise, Rolls argues, because "in humans, the reward and punishment systems may operate implicitly in comparable ways to those in other animals. But in addition to this, humans have the explicit system (closely related to consciousness) which enables us consciously to look and predict many steps ahead" (page 282). Thus, for example, the sight or smell of a cookie might initiate motivation to consume on the part of the affective system, but might remind the cognitive system that one is on a diet.

Exactly how cognitive and affective systems compete for control of behavior is not well understood. At a neurological level, it appears that an organ in the reptilian brain called the striatum (part of a larger system called the basal ganglia) plays a critical role. The striatum receives inputs from all parts of the cerebral cortex, including the motor cortex, as well as from affective systems such as the amygdala. Lesions of pathways that supply dopamine to the striatum leads, in animals, to a failure to orient to stimuli, a failure to initiate movements, and a failure to eat or drink (Marshall et al. 1974). In humans, depletion of dopamine in the striatum is found in Parkinson's disease, the most dramatic symptom of which is a lack of voluntary movement. The striatum seems to be involved in the selection of behaviors from competition between different cognitive and affective systems – in producing one coherent stream of behavioral output, which can be interrupted if a signal of higher priority is received.

The extent of collaboration and competition between cognitive and affective systems, and the outcome of conflict when it occurs, depends critically on the intensity of affect (Loewenstein 1996; Loewenstein and Lerner 2003). At low levels of intensity, affect appears to play a largely "advisory" role. A number of theories posit that emotions carry information that people use as an input into the decisions they face (e.g. Damasio 1994; Peters and Slovic 2000). Affect-as-information theory represents the most well-developed of these approaches (Clore 1992; Schwarz 1990; Schwarz and Clore 1983).

At intermediate levels of intensity, people begin to become conscious of conflicts between cognitive and affective. It is at such intermediate levels of intensity that one observes the types of efforts at self-control that have received so much attention in the literature (Elster 1977; Mischel, Ebbesen and Zeiss 1972; Schelling 1978; 1984).

Finally, at even greater levels of intensity, visceral factors can be so powerful as to virtually preclude decision making. No one “decides” to fall asleep at the wheel, but many people do. Under the influence of intense emotions, people often report themselves as being "out of control" or "acting against their own self-interest" (Baumeister, Heatherton and Tice 1994; Bazerman, Tenbrunsel and Wade-Benzoni 1998; Hoch and Loewenstein 1991; Loewenstein 1996). As Rita Carter writes, in her superb introduction to neuroscience, Mapping the Mind, "where thought conflicts with emotion, the latter is designed by the neural circuitry in our brains to win" (1999: 54).

Erroneous sense-making: The brain's powerful drive toward sense making leads us to strive to interpret our own behavior. Such interpretations use quadrant I mechanisms to make sense of behavior which is caused by all four quadrants and their interaction. Since quadrant I often does not have conscious access to behavior in the other quadrants, it is perhaps no surprising that it tends to overattribute behavior to itself – i.e. to a deliberative decision process.

Research with EEG recordings has shown (Libet 1985) that the precise moment at which we become aware of an intention to perform an action trails the initial wave of brain activity associated with that action (the EEG "readiness potential") by about 300 msec. The overt behavioral response itself then follows the sensation of intention by another 200 msec. Hence, what is registered in consciousness is a regular pairing of the sensation of intention followed by the overt behavior. Because the neural activity antecedent to the intention is inaccessible to consciousness, we experience 'free will' (i.e., we cannot identify anything that is causing the feeling of intention). Because the behavior reliably follows the intention, we feel that this 'freely willed' intention is causing the action—but in fact, both the sensation of intention and the overt action are caused by prior neural events which are inaccessible to consciousness. [11]

Quadrant I tends to explain behavior in the only terms it knows—as the result of deliberative decision making (see Nisbett and Wilson 1977; Wegner and Wheatley 1999). A dramatic study demonstrating this phenomenon was conducted with a "split-brain" patient (who had an operation separating the connection between the two hemispheres of his brain). The patient's right hemisphere could interpret language but not speak, and the left hemisphere could speak (LeDoux 1996). The patient's right hemisphere was instructed to wave his hand (by showing the word "wave" on the left part of a visual screen, which only the right hemisphere processed). The left hemisphere saw the right hand waving but was unaware of the instructions that had been given to the right hemisphere (because the cross-hemisphere connections were severed). When the patient was asked why he waved, the left hemisphere (acting as spokesperson for the entire body) invariably came up with a plausible explanation, like "I saw somebody I knew and waved at them."

Yet a third illustration of sense-making comes from the patient, mentioned earlier, who laughed when a certain part of her brain was stimulated electrically. Such surgery is conducted with the patient conscious, so the surgeons were able to ask her what she found so funny. She replied "You guys are just so funny – standing around." Later, when the same area was stimulated again she reported that her reaction was triggered by a picture of a horse that happened to be on the wall of the operating room. In addition to pointing to the possible existence of a 'humor module', these findings highlight the cognitive brain's remarkable tendency to make sense of emotional

---

[11] "… the brain contains a specific cognitive module that binds intentional actions to their effects to construct a coherent experience of our own agency. Haggard, P., S. Clark and J. Kalogeras. 2002. "Voluntary action and conscious awareness." *Nature Neuroscience*, 5:4, pp. 382-385.." [

reactions – in this case to find a plausible reason for laughing, even when the real reason was the placement of the surgeon's probe.

Given that people are capable of rational deliberation, why doesn't quadrant I thinking simply correct initial impressions formed from automatic quadrant III and IV processes when they are wrong? Sometimes it does— pilots learn to trust their instrument panels, even when they conflict with strong sensory intuitions about where their plane is headed— but cognitive override is difficult. Quadrant I has to (a) recognize that an initial impression is wrong (which requires self-awareness about behavior in the other quadrants) then (b) deliberately correct that impression. But when sense making works outside of consciousness it will not generate alarm bells to trigger the recognition required in (a). Even when the external influence is obvious and inappropriate, or the subject is warned ahead of time, the deliberation required to correct the first impression is not automatic or effortless (Gilbert 2002) it is real mental work, competing for mental resources and attention with all the other work that needs to be done at the same moment. The struggle between rapid unconscious pattern-detection processes and their slow, effortful adjustment by deliberation is not a fair fight; so automatic impressions will influence behavior much of the time.

This is important for phenomena such as labor market discrimination. Economic models assume that labor market discrimination against minorities is either a *taste* (a dislike of minorities or a distaste for working with them), or a *belief* (that minority status is a proxy for unobservable differences in skill, or "statistical discrimination"). Neuroscience suggests a different answer. Automaticity contributes to discrimination because neural networks rapidly spread activation through associated concepts and stereotypes. Affect contributes to discrimination because automatic affective reactions have such a powerful effect on cognitive judgments. Discrimination in this view involves rapid, automatic, associations between social categories, stereotypes, and affect.

Such an account receives support from remarkable experiments which demonstrated subtle 'implicit associations' between demographic categories and good or bad adjectives (to try it out on yourself, go to http://buster.cs.yale.edu/implicit/). Subjects in these computer-administered studies are shown a mixed up series of stereotypically black or white names (Tyrone or David) and positive or negative adjectives (mother or devil). They are asked to tap one key when they see one type of name or adjective, and a different key if they see the other type of name or adjective, at which point the computer moves on to the next name or adjective. The dependent variable is how long it takes the subject to work their way through the complete list of names and adjectives. White subjects work their way through the list much more quickly when one key is linked to the pair (black or negative) and the other to (white or positive) than they are when one key is linked to [black or positive] and the other to [white or negative].

What's going on? The brain encodes associations in neural networks which spread activation to related concepts.[12] For white students, black names are instantly associated with negative concepts, whether they realize it or not, because the association is automatic (rapid and unconscious). This implicit association test (IAT) hints that the cognitive roots of labor market discrimination may be mental associations which are neither tastes nor beliefs, per se. The same idea might explain the fact that taller and more attractive people are more likely to earn higher wages (e.g.Persico, Andrew and Silverman 2002) and other rewards (e.g., the Presidency). Because people lack introspective access to the processes that produce such biases, they are unable to correct for them even when they are motivated to make impartial judgments and decisions.

A major implication of the brain's proclivity for erroneous sense-making is that one needs to be suspicious of intuitive, introspective, accounts of behavior, and specifically of the tendency to view one's own behavior as the outcome of deliberations. In other words, our ‘lay theory’ of action has much affinity with the economic model. Even though much of the brain’s activity is ‘cognitively inaccessible,’ we have the illusion that we are able to make sense of it, and we tend to make sense of it in terms of Quadrant I processes. This is exactly why we need to rely on neuroscience rather than intuitive sense making to figure out what causes behavior.

Implications of neuroscience for economics

To add value to economics, neuroscience needs — at a minimum — to provoke thought, and suggest interesting, fresh perspectives on old problems. This section illustrates how the concepts and findings just discussed might affect the way that two traditional topics in economics – intertemporal choice and decision making under risk and uncertainty -- are approached.

---

[12] There is also evidence that discriminatory associations might be flexible because Asian-American students who are more immersed in American culture exhibit IAT patterns which are less pro-Asian.

Intertemporal Choice and self-control

The standard perspective in economics views intertemporal choice as a trade-off of utility at different points in time. Individual differences in the way that people make this tradeoff are captured by the notion of a discount rate — a rate at which people discount future utilities as a function of when they occur. The notion of discounting, however, gained currency not because of any supportive evidence, but based only on its convenient similarity to financial net present value calculations (Loewenstein 1992). In fact, empirical research on time discounting challenges the idea that people discount all future utilities at a constant rate. (For a review of the evidence, see Frederick, Loewenstein and O'Donoghue 2002). The notion of time discounting, it appears, neither describes the behavior of individuals nor helps us to classify individuals in a useful fashion.

Humans appear to be unique among animals in terms of caring about, and thus making immediate sacrifices for, the sake of desired future consequences. To understand intertemporal choice in humans, therefore, we need to take account not only of the processes we share with other animals (especially those represented by quadrant IV, of table 1), but also those that are uniquely human (mainly quadrant I from table 1).[13] Intertemporal choice illustrates nicely how these processes both collaborate and compete with one-another.

Collaboration is illustrated by the fact that decisions to delay gratification involve an admixture of affect and cognition. They require a cognitive awareness of the delayed benefits in delaying gratification – e.g., that desisting from eating cake today will mean a more pleasing body type in the future. But, as many researchers have observed, cognitive awareness, alone, is insufficient to motivate delay of gratification; emotions play a critical role in forward-looking decision making. "The capacity to experience anxiety and the capacity to plan," David Barlow (1988:12) notes, are "two sides of the same coin." Cottle and Klineberg (1974) argue that people only care about the delayed consequences of their decisions to the degree that contemplating such consequences evokes immediate affect. In support of this view, they cite the effects of frontal lobotomies which, they argue, create a deficiency in areas of the brain that underlie the capacity for images of absent events to generate experiences of pleasure or discomfort (p.15). The neurosurgeons who performed these operations wrote of their frontal-lobotomy patients that: "the

[13] Given this discontinuity, we believe it is a mistake to draw a connection between hyperbolic time discounting in animals and hyperbolic time discounting. The similarity, we believe, is purely coincidental and is not all that

capacity for imagination is still present, and certainly not sufficiently reduced to render the patients helpless, and affective responses are often quite lively, [but there is] a separation of one from the other" (Freeman and Watts 1942:303). The work of Damasio and colleagues discussed earlier (Bechara, Damasio, Tranel and Damasio 1997; Damasio 1994) lends further credence to this perspective, as does research on psychopaths, who are characterized by both emotional deficits when it comes to imagining the future and by insensitivity to the future consequences (as well as consequences to others) of their behavior (Cleckley 1941; Hare 1965; 1966; Lykken 1957).

Competition is illustrated by the ubiquity of self-control problems in which one's cognitive judgment of the best course of behavior departs from the actions one is affectively motivated to take. As discussed above, the affective system is designed to ensure that certain survival and reproduction functions are met and it achieves this function in part by motivating individuals to take certain actions. In most animals, emotions and drives motivate behaviors that have short-term goals, such as eating, drinking, copulating, so the automatic system is inherently myopic. Humans are different from other animals in that we worry about or derive immediate pleasure from thinking about delayed consequences, so our affective system can also motivate behaviors that have long-term beneficial consequences. Indeed, a number of human pathologies, such as anxiety disorders, workaholism, and self-destructive miserliness, seem to be driven by an *excess* of future-mindedness. In humans, therefore, one cannot equate affect with myopia, even though so many myopic behaviors, such as overeating, impulse-buying and sexual risk-taking are associated with powerful immediate affect.

An intriguing aspect of self-control is that it is often associated with a subjective feeling of mental effort. It is tempting to attribute this to the fact that self-control involves the same part of the brain – the executive prefrontal cortex – that is itself associated with feelings of mental effort. Perhaps this is why exercising willpower feels so difficult, and why exercising self-control in one domain can undermine its exercise in another, as demonstrated by a series of clever experiments conducted by Roy Baumeister and colleagues (see, e.g., Baumeister and Vohs 2003). In a typical study, subjects on diets who resisted temptation (by foregoing the chance to grab snacks from a nearby basket) later ate more ice cream in an ice-cream taste test and also quit earlier when confronted with an intellectual problem they couldn't solve. They acted as if their ability to resist

---

surprising given that many perceptional and judgmental processes other than time discounting exhibit a hyperbolic form.

temptation was temporarily "used up" by resisting the snacks (or, alternatively, that they had "earned" a reward of ice cream by skipping the tempting snacks). In this viewpoint, advertising which tempts struggling dieters is a theft (or negative externality) that robs the dieters of something valuable.

How might one model intertemporal choice differently as a result of the insights from neuroscience? First, the neuroscience research points to ways to "unpack" the concept of time preference. Clearly, ability to think about future consequences is important, which is probably why time preference is correlated with measured intelligence (Mischel and Metzner, 1962). Second, because people are likely to make myopic choices when under the influence of powerful drives or emotions (Loewenstein 1996), this suggests that a key to understanding impulsivity in individuals might be to understand what types of situations get them 'hot'. Third, we might be tempted to look for individual differences in what could be called 'willpower' – i.e., the availability of the scarce internal resource that allows people to inhibit viscerally-driven behaviors.

Such a model could help to explain not only impulsivity, but also why many people have self-control problems of the opposite type of those typically examined in the literature – e.g., tightwads who can't get themselves to spend enough, workaholics who can't take a break, and people who, far from losing control in the bedroom, find themselves frustratingly unable to do so. All of these patterns of behavior can easily be explained by the, possibly uniquely human, propensity to experience emotions, such as fear, as a result of thinking about the future. Indeed, it is likely that one of the main tools that the prefrontal cortex uses to impose self-control when affective forces would otherwise favor short-sighted self-destructive behavior is to create 'deliberative affect' via directed imagery and thought (Giner-Sorolla 2001).

Such a framework might also help to explain why people appear so inconsistent when their behavior is viewed through the lens of discounted utility. The ability to think about future consequences may not be strongly correlated with the degree to which different experiences produce visceral reactions, and these in turn might not be correlated with an individual's level of willpower. Indeed, Frederick et al. (2002) found close to zero correlations between numerous behaviors that all had an important intertemporal component, but much higher correlations between behaviors that seemed to draw on the same dimension of intertemporal choice – e.g., which required suppression of specific emotions such as anger.

Decision-making under risk and uncertainty

The expected utility model views decision making under uncertainty as a tradeoff of utility under different states of nature -- i.e., different possible scenarios. But, much as they do to delayed outcomes, people react to risks at two different levels. On the one hand, as posited by traditional economic theories and consistent with quadrant I of table 1, people do attempt to evaluate the objective level of risk that different hazards could pose. On the other hand, and consistent with quadrant IV, people also react to risks at an emotional level, and these emotional reactions can powerfully influence their behavior (Loewenstein, Weber, Hsee and Welch 2001).

A lot is known about the neural processes underlying affective responses to risks. Much risk averse behavior is driven by immediate fear responses to risks, and fear, in turn, seems to be largely traceable to a single small area of the brain called the amygdala. The amygdala constantly scans incoming stimuli for indications of potential threat, and responds to inputs both from automatic and controlled processes in the brain. Vuilleumier et al. (2001) observed equivalent amygdala activation in response to fearful faces that were visually attended to or in the peripheral region which falls outside of conscious perception (cf. LeDoux 1996; Morris, Buchel and Dolan 2001; Whalen, Rauch, Etcoff, McInerney, Lee and Jenike 1998). But the amygdala also receives cortical inputs, which can moderate or even override its automatic quadrant IV response. In a paradigmatic experiment that illustrates cortical overriding of amygdala activation, an animal such as a rat is "fear-conditioned" – by repeatedly administering a signal such as a tone followed by administration of a painful electric shock. Once the tone becomes associated in the animal's mind with the shock, the animal responds to the tone by jumping or showing other over signs of fear. In the next phase of the experiment, the tone is played repeatedly without administering the shock, until the fear response becomes gradually "extinguished." At this point, one might think that the animal has 'unlearned' the connection between the tone and the shock, but the reality is more complicated and interesting. If the neural connections between the cortex and the amygdala are then severed, the original fear response to the tone reappears, which shows that fear conditioning is not erased in 'extinction' but is suppressed by the cortex and remains latent in the amygdala. At the deepest level, then, far learning may be permanent, which could be evolutionarily useful in allowing a very rapid response if the cause of the fear reappears.

Decision making under risk and uncertainty, like intertemporal choice, nicely illustrates both collaboration and competition between systems. When it comes to collaboration, risk taking (or avoiding) behavior involves an exquisite interplay of cognitive and affective processes. In a well-known study that illustrates such collaboration (Bechara, Damasio, Tranel and Damasio 1997), patients suffering prefrontal damage (which, as discussed above, produces a disconnect between cognitive and affective systems) and normal subjects chose a sequence of cards from four decks whose payoffs the subjects only learned from experience. Two decks had more cards with extreme wins and losses (and negative expected value); two decks had less extreme outcomes but positive expected value. Both groups exhibited similar skin conductance (sweating – an indication of fear) after large-loss cards were encountered, but, compared to normals, prefrontal subjects rapidly returned to the high-paying risky decks after suffering a loss and, as a result went "bankrupt" more often. Although the immediate emotional reaction of the prefrontal patients to losses was the same as the reaction of normals (measured by skin conductance), the damage patients apparently do not store the pain of remembered losses as well as normals, so their skin conductance rose much less than normals when they resampled the high risk decks. Subsequent research found a similar difference between normal subjects who were either high or low in terms of emotional reactivity to negative events. Those who were more reactive were more prone to sample from the lower-paying, safer decks of cards (Peters and Slovic 2000).

Damasio et al.'s research shows that insufficient fear can produce nonmaximizing behavior when risky options have negative value. But, it is well established that fear can also discourage people from taking advantageous gambles (see, e.g., Gneezy and Potters 1997). Indeed, Loewenstein, Shiv, Bechara, Damasio and Damasio (2002) found showed that frontal patients actually make more money on a task in which negative emotions cause normal subjects to be extremely risk averse: a series of take-it-or-leave-it choices to play a gamble with a 50% chance of losing $1.00 or gaining $1.50. Normal subjects and frontal subjects were about equally likely to play the gamble on the first round, but normals rapidly stopped playing when they experienced losses while frontal patients' play was unresponsive to losses. Clearly, having frontal damage undermines the overall quality of decision making; but there are situations in which frontal damage can result in superior decisions.

At a more macro level, emotional reactions to risk can help to explain risk-seeking as well as risk-aversion (Caplin and Leahy 1997). Thus, when gambling is pleasurable a model that incorporates affect naturally predicts that people will be risk-seeking and that self-control will be required to reign in risk-taking. Indeed, about 1% of the people who gamble are diagnosed as “pathological”—they report losing control, “chasing losses”, and harming their personal and work relationships by gambling (Pathological Gambling, 1999). The standard economic explanations for gambling-- convex utility for money or a special taste for the act of gambling-- don't help explain why some gamblers binge and don’t usefully inform policies to regulate availability of gambling. Neuroscience may help. Pathological gamblers tend to be overwhelmingly male, and tend to also drink, smoke, and use drugs much more frequently than average. Genetic evidence shows that a certain gene allele (D_2A1), which causes gamblers to seek larger and larger thrills to get modest jolts of pleasure, is more likely to be present in pathological gamblers than in normal people (Comings 1998). One study shows tentatively that treatment with naltrexone, a drug that blocks the operation of opiate receptors in the brain, reduces the urge to gamble (e.g. Moreyra, AIbanez, Saiz-Ruiz, Nissenson and Blanco 2000). The same drug has been used to successfully treat “compulsive shopping” (McElroy, Satlin, Pope, Keck and al. 1991).

Understanding the affective and cognitive components of reactions to risk is especially important when the two diverge and hence compete for control of behavior (see Loewenstein, Weber, Hsee and Welch 2001). People are often 'of two minds' when it comes to risks; we drive (or wish we were driving as we sit white-knuckled in our airplane seat) when we know at a cognitive level that it is safer to fly. We fear terrorism, when red meat poses a much greater risk of mortality. And, when it comes to asking someone out on a date, getting up to speak at the podium, or taking an important exam, our deliberative self uses diverse tactics to get us to take risks, or to perform in the face of risks, that our visceral self would much prefer to avoid. Perhaps the most dramatic illustrations of the separation of visceral reactions and cognitive evaluations, however, comes from the phobias that so many people suffer from; the very hallmark of a phobia is to be unable to face a risk that one recognizes, objectively, to be harmless. Moreover, fear unleashes preprogrammed sequences of behavior that aren't always beneficial. Thus, when fear becomes too intense it can produce counter-productive responses such as freezing, panicking, or 'dry-mouth' when speaking in public. The fact that people pay for therapy to deal with their fears, and take drugs (including alcohol) to overcome them, can be viewed as further “evidence” that people, or more accurately, people's deliberative selves, are not at peace with their visceral reactions to risks.

The divergence between different systems' evaluations of risk can also be seen when it comes to judgments of probability. Numerous studies by psychologists have observed systematic divergences between explicit judgments of probability in different settings (presumably the product of controlled processing) and implicit judgments or judgments derived from choice. For example, Kirkpatrick and Epstein (see also Denes-Raj, Epstein and Cole 1995; 1992; Windschitl and Wells 1998) found that people prefer to draw a bean from a bowl containing 10 winning beans and 90 losing beans than from a bowl containing 1 winning bean and 9 losing beans. Subjects say that they know the explicit probabilities of winning are the same, but they still have an automatic quadrant III preference for the bowl with more winning beans. An important feature of good probability judgment is logical coherence: Probabilities of mutually exclusive and exhaustive events should add to one, and conditional probabilities like P(recession|downturn) should equal P(recession *and* downturn) /P(downturn). Logical coherence is violated in at least two interesting ways. One is `conjunction fallacy'- the tendency to judge events with two components A and B as more likely than A or B alone. While most subjects (even statistically sophisticated ones) make conjunction errors on some problems, when those errors are pointed out quadrant I wakes up, and the subjects sheepishly recognize the error and correct it (Kahneman and Frederick 2002). Another violation is that subjects often report probabilities which are logically incoherent. fMRI evidence suggests an explanation for why probability judgments are incoherent, but can be corrected upon reflection: When guessing probabilities, the left hemisphere of the brain is more active; but when answering logic questions, the right hemisphere is more active (Parsons and Osherson 2001). Enforcing logical coherence requires the right hemisphere to `check the work' of the left hemisphere.

Neural evidence also substantiates the distinction between risk (known probability) and "Knightian" uncertainty, or ambiguity. Subjects facing ambiguous gambles—knowing they lack information they would like to have about the odds—often report a feeling of discomfort or mild fear. Brain imaging shows that different degrees of risk and uncertainty activate different areas of the brain (McCabe, Houser, Ryan, Smith and Trouard 2001; Rustichini, Dickhaut, Ghirardato, Smith and Pardo 2002) which corroborates the subjects' self-reports.

Conclusions

Economics parted company from psychology in the early 20th century as economists became skeptical that basic psychological forces could be measured without inferring them from behavior, which led to adoption of the useful tautology between unobserved utilities and observed (revealed) preferences. But remarkable advances in neuroscience now make direct measurement of thoughts and feelings possible for the first time, opening the "black box" which is the building block of any economic interaction and system — the human mind.

Most economists are both curious about neuroscience, but instinctively skeptical that it can tell us how to do better economics. The tradition of ignoring psychological regularity in making assumptions in economic theory is so deeply ingrained — and has proved relatively successful — that knowing more about the brain seems unnecessary. Economic theory could chug along successfully for the next few years paying no attention at all to cognitive neuroscience (and surely will). But it is hard to believe that *some* neuroscientific regularities will not help explain *some* extant anomalies, particularly those anomalies that have been debated for decades.

For example, in ultimatum bargaining games players usually offer close to half of a fixed sum to another player, and that responder often reject small offers, though if players are self-interested game theory predicts they should offer very little and accept anything. Neuroscience is helping resolve why players reject small offers. Hill and Sally (2002) found that autistic adults were more likely than normal adults to offer nothing to the other player. The autists seem incapable of understanding what another player might believe and do; as a result, ironically, they play like self-interested game theorists! Sanfey et al (2002) did fMRI imaging when players received either even offers ($5 out of $10) or unfair offers ($3). Their findings illustrate the collaboration of affect and cognition. When players received unfair offers there was more activation in insular cortex (an area associated with feelings of disgust) and more activation in anterior cingulate cortex (ACC, an area active in difficult decisions requiring cognitive control). The brain activity shows that players faced with unfair offers react with disgust, while the ACC struggles to decide which is worse— disgust or poverty. These direct measures of brain activity can also be used to predict what will happen-- players who have more insular cortical activity after unfair offers are more likely to reject those offers.

These studies of ultimatum games are early successes which provide important clues. In many other domains much less is known. For example, in finance, after decades of careful research there is no definitive theory of why stock prices fluctuate, why there is so much trading, and why there are so many actively-managed mutual funds (more than there are NYSE stocks) despite the funds' poor overall performance. Perhaps knowing more about basic neural mechanisms that underlie conformity, attention paid to large price changes, wishful thinking, sense making of random series, and perceptions of expertise can help explain these puzzles.

There are many anomalies in intertemporal choice. In the US, credit card debt is substantial ($5000 per household) and a million personal bankruptcies have been declared in each of several years (Laibson, Repetto and Tobacman 1998). Healthier food is cheaper and more widely available than ever before, but spending on dieting and obesity are both on the rise. Surely understanding how brain mechanisms process reward, and curb or produce compulsion, might help explain these facts and shape sensible policy and regulation.

Prevailing models of advertising assume that ads convey information or signal a product's quality or, for "network" or "status" goods, a product's likely popularity. Many of these models seem like strained attempts to explain effects of advertising without incorporating the obvious intuition that advertising taps neural circuitry of reward and desire.

Finally, the current models do not provide a satisfying theory of how individuals differ. We characterize people as impulsive or deliberate, stable or neurotic, decisive or indecisive, mature or immature, foolish or wise, depressed or optimistic, scatterbrained or compulsively organized. The consumers who spend countless dollars on self-help, "organize your life" manuals, and who sustain the huge, and infinitely varied psychological counseling industry, are typically unhappy where they stand on some of these dimensions, and are looking for ways to change. Comparative economic development, entrepreneurial initiative and innovation, business cycle sensitivity, and other important macroeconomic behaviors are probably sensitive to the distribution of these and other psychological "assets." Yet there is really no intelligent way to discuss them with the language of beliefs and desires, which is the only language operating in Quadrant I.

Neuroscientific studies of simple reward circuitry in rats and primates vindicate some of the simplest ideas of economics (e.g., substitution and existence of neurons which encode expected reward; see Glimcher 2002). However, much human judgment and choice is of a very different character. Behavior is dependent on evolutionarily older systems that humans share with animals. The controlled, cognitive, system needs to get through to the animal to control behavior. In many situations people report feeling out of control; consciously they know what is best, but they are unable to implement that course of action. Most cognitive and affective processing does not occur in the rule-like systematic fashion envisioned by standard economics, but involves automatic, massively parallel systems to which we have little or no introspective access. Most human behavior, like that of other mammals, is the product of processes that fall outside of quadrant I.

In the short-run, an "incremental" approach in which psychological evidence suggests functional forms will help enhance the realism of existing models. For example, the two-parameter "ß-δ" hyperbolic discounting approach (where ß expresses preference for immediacy, and is equal to one in the standard model) is an example that has proven productive theoretically (e.g. Laibson, Repetto and Tobacman 1998, O'Donoghue and Rabin 1999). Laibson's (2001) model of homeostatic response to environmental cues in addiction is an incremental model well grounded in recent neuroscience.

However, we believe that in the long run a more "radical" departure from current theory will become necessary, radical in the sense that it goes beyond the Quadrant I beliefs-desires model, and attempts to explicitly represent multiple brain mechanisms. Standard economic theories rely on an implicit assumption that controlled, cognitive processes (quadrant I of figure 1) are the key to economic decision making. Our theme in this paper is that radical models should respect the fact that brain mechanisms combine controlled and automatic processes, operating using cognition and affect. The Platonic metaphor of the mind as a charioteer driving twin horses of reason and emotion is on the right track—except that cognition is a smart pony, and emotion a big elephant.

Of course, the challenge in radical-style theorizing is to develop models of how multiple mechanisms interact which are precise. Can this be done**?** The answer is Yes. Bernheim and Rangel (2002) and Benhabib and Bisin (2002) have recent models with interacting mechanisms much like those in our Figure 1. Furthermore, while interactions of multiple brain mechanisms might appear to be too radical a change from equilibrium with utility maximization, we think many familiar tools can be used to do radical neuroeconomics. Interactions of cognition and affect might resemble systems like supply and demand, or feedback loops which exhibit multiple equilibria. The interaction of controlled and automatic processes might be like an inventory policy or agency model in which a controller only steps in when an extreme state of the system (or unusual event) requires controlled processes to override automatic ones. The influence of affect on choices is a very general type of state-dependence (where the "state" is affective, and is influenced by external cues and also by internal deliberation and restraint). Instead of solving for equilibria in these interacting-mechanism models, solve for steady states or cyclic fluctuations. Instead of summarizing responses to changes by comparative statics, study impulse-response functions.

Finally, although we focused solely on applications of neuroscience to economics, intellectual trade could also flow in the opposite direction. Neuroscience is shot through with familiar economic language — delegation, division of labor, constraint, coordination, executive function — but these concepts are not formalized in neuroscience as they are in economics. Neuroscientists also do not understand how the brain allocates resources that are essentially fixed (e.g., blood flow and attention). An "economic model of the brain" could help neuroscientists comprehend how various brain systems interact and allocate scarce brain resources. Simple concepts in economics, like mechanisms for rationing under scarcity, and general versus partial equilibrium responses to shocks, could help neuroscientists understand how the entire brain interacts.

References

Bargh, John A., Shelly Chaiken, Paula Raymond and Charles Hymes. 1996. "The Automatic Evaluation Effect: Unconditional Automatic Attitude Activation with a Pronunciation Task." *Journal of Experimental Social Psychology*, 32, pp. 104-128.

Bargh, John A. and Tanya L. Chartrand. 1999. "The unbearable automaticity of being." *American Psychologist*, 54:7, pp. 462-479.

Barlow, David H. 1988. *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York, NY: Guilford Press.

Baumeister, Roy F., Todd F. Heatherton and Dianne M. Tice. 1994. *Losing Control: How and Why People Fail at Self-Regulation*. San Diego: Academic Press.

Baumeister, Roy F. and Kathleen D. Vohs, Willpower, choice, and self-control, 2003 *in* "Time and Decision: Economic and psychological perspectives on intertemporal choice" (G. F. Loewenstein, D. Read and R. F. Baumeister, Eds.), New York Russell Sage, pp.

Bazerman, Max H., Ann E. Tenbrunsel and Kimberly A. Wade-Benzoni. 1998. "Negotiating With Yourself and Losing: Making Decisions With Competing Internal Preferences." *Academy of Management Review*, 23:2, pp. 225-241.

Bechara, Antoine, Antonio R. Damasio, Hanna Damasio and Steven W. Anderson. 1994. "Insensitivity to future consequences following damage to human prefrontal cortex." *Cognition*, 50:1-3, pp. 7-15.

Bechara, Antoine, Hanna Damasio, Antonio R. Damasio and G. P. Lee. 1999. "Different contributions of the human amgdala and ventromedial prefrontal cortex to decision-making." *Journal of Neuroscience*, 19:13, pp. 5473-5481.

Bechara, Antoine, Hanna Damasio, Daniel Tranel and Antonio R. Damasio. 1997. "Deciding advantageously before knowing the advantageous strategy." *Science*, 275:5304, pp. 1293-1294.

Benhabib, Jess and Alberto Bisin, Self-control and consumption-savings decisions: Cognitive perspectives, 2002, New York University, pp.

Bernheim, Douglas and Antonio Rangel, Addiction and cue-conditioned cognitive processes, 2002, Stanford, Stanford University, pp.

Buck, Ross. 1999. "The biological affects: A typology." *Psychological Review*, 106:2, pp. 301-336.

Cabanac, Michel. 1979. "Sensory pleasure." *Quarterly Review of Biology*, 54, pp. 1-29.

Caggiula, Anthony R. and Bartley G. Hoebel. 1966. ""Copulation-reward site" in the posterior hypothalamus." *Science*, 153:3741, pp. 1284-1285.

Camerer, Colin F. 2003. *Behavioral game theory: Experiments on strategic interaction*. Princeton: Princeton University Press.

Camerer, Colin F. and George Loewenstein, Behavioral economics: Past, present, and future, in press *in* "Readings in Behavioral Economics" (C. F. Camerer, G. Loewenstein and M. Rabin, Eds.), New York Russell Sage, pp.

Camerer, Colin, George Loewenstein and Martin Weber. 1989. "The curse of knowledge in economic settings: An experimental analysis." *Journal of Political Economy*, 97:5 October, pp. 1232-1254.

Camerer, Colin and Dan Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 89:1 March, pp. 306- 318.

Campfield, L.A. and F.J. Smith, Systemic factors in the control of food intake, 1990 *in* "Handbook of behavioural neurobiology, Vol. 10, Neurobiology of food and fluid intake" (E. M. Stricker, Ed.), New York Plenum, pp. 183-206.

Caplin, A. and J. Leahy, Psychological expected theory and anticipatory feelings, 1997 NYU Working Paper 4766, pp.

Carter, Rita. 1999. *Mapping the mind*. Berkeley, CA: University of California Press.

Cleckley, Hervey. 1941. *The mask of sanity*. St. Louis: C.V. Mosby.

Clore, Gerald L., Cognitive phenomenology: Feelings and the construction of judgment, 1992 *in* "The construction of social judgments" (L. L. Martin and A. Tesser, Eds.), Hillsdale, NJ Erlbaum, pp. 133-163.

Comings, David E. 1998. "the molecular genetics of pathological gambling." *CNS Spectrums*, 3:6, pp. 20-37.

Cosmides, Leda. 1989. "The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task." *Cognition*, 31:3, pp. 187-276.

Cottle, Thomas J. and Stephen L. Klineberg. 1974. *The present of things future: Explorations of time in human experience*. New York: Free Press.

Damasio, Antonio R. 1994. *Descartes' error : Emotion, reason, and the human brain*. New York: G.P. Putnam.

Denes-Raj, Veronika, Seymour Epstein and Jonathan Cole. 1995. "The generality of the ratio-bias phenomenon." *Personality and Social Psychological Bulletin*, 21:10, pp. 1083-1092.

Elbert, Thomas, Christo Pantev, Christian Wienbruch, Brigitte rockstroh and Edward Taub. 1995. "Increased cortical representation of the fingers of the left hand in string players." *Science*, 270, pp. 305-307.

Elster, Jon. 1977. *Ulysses and the sirens*. Cambridge, England: Cambridge University Press.

Fletcher, Paul, Francesca Happe, Uta Frith, C. Baker, Ray Dolan, Richard Frackowiak and Chris Frith. 1995. "Other minds in the brain: A functional imagining study of "theory of mind" in story comprehension." *Cognition*, 57, pp. 109-128.

Frederick, Shane, George Loewenstein and Ted O'Donoghue. 2002. "Time discounting and time preference: A critical review." *Journal Of Economic Literature*, 40, pp. 351-401.

Freeman, Walter J. and JamesW Watts. 1950. *Psychosurgery in the treatment of mental disorders and intractable pain*. Springfield, Illinois: Thomas.

Fried, Izhak. 1998. "Technical Comment: The hippocampus and human navigation." *Science*, 282, pp. 2151.

Frith, Uta. 2001a. "Mind blindness and the brain in autism." *Neuron*, 32:6, pp. 969-979.

Frith, Uta. 2001b. "What framework should we use for understanding developmental disorders?" *Developmental Neuropsychology*, 20:2, pp. 555-563.

Gazzaniga, M.S. and J.E. LeDoux. 1978. *The Integrated Mind*. New York: Plenum.

Gilbert, Daniel T., Inferential correction, 2002 *in* "Heuristics and biases: The psychology of intuitive judgment" (T. Gilovich, D. Griffin and D. Kahneman, Eds.), Cambridge Cambridge University Press, pp. 167-184.

Gilbert, Daniel T. and Michael Gill. 2000. "The momentary realist." *Psychological Science*, 11, pp. 394-398.

Giner-Sorolla, Roger. 2001. "Guilty pleasures and grim necessities: Affective attitudes in dilemmas of self-control." *Journal of Personality and Social Psychology*, 80:2, pp. 206-221.

Glimcher, Paul. 2002. *Decisions, Uncertainty and the Brain: The Science of Neuroeconomics*. Cambridge: MIT Press.

Gneezy, Uri and Jan Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *The Quarterly Journal of Economics*,:, pp. 631-645.

Gneezy, Uri and Aldo Rustichini, Incentives, punishment and behavior, in press *in* "Readings in behavioral economics" (C. F. Camerer, G. F. Loewenstein and M. Rabin, Eds.), New York, NY Russell Sage Foundation Press, pp.

Gobet, Fernand and Herbert A. Simon. 1996. "Recall of random and distorted chess positions: Implications for the theory of expertise." *Memory & Cognition*, 24:4, pp. 493-503.

Green, Leonard J. and Howard Rachlin. 1991. "Economic substitutabilit of electrical brain stimulation, food, water." *Journal of the Experimental Analysis of Behavior*, 55, pp. 133-143.

Greenberg, D., G.P. Smith and J. Gibbs. 1990. "Intraduodenal infusions of fats elicit satiety in sham-feeding rats." *The American Journal of Physiology*, 259:1, pp. R110-R118.

Haggard, P., S. Clark and J. Kalogeras. 2002. "Voluntary action and conscious awareness." *Nature Neuroscience*, 5:4, pp. 382-385.

Happe, Francesca, Stefan Ehlers, Paul Fletcher, Uta Frith, Maria Johansson, Christopher Gillberg, Ray Dolan, Richard Frackowiak and Chris Frith. 1996. ""Theory of mind" in the brain. Evidence from a PET scan study of Asperger syndrome." *NeuroReport*, 8:1, pp. 197-201.

Hare, R. D. 1965. "Psychopathy, fear arousal and anticipated pain." *Psychological Reports*, 16, pp. 499-502.

Hare, R. D. 1966. "Temporal gradient of fear arousal in psychopaths." *Journal of Abnormal and Social Psychology*, 70, pp. 422-445.

Hastie, Reid. 1984. "Causes and effects of causal attribution." *Journal of Personality & Social Psychology*, 46:1, pp. 44-56.

Hebb, Donald. 1949. *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hershey, J. C. , H.C. Kunreuther and P.J. Schoemaker. 1982. "Sources of Bias in Assessment Procedures for Utility Functions." *Management Science*, 28:8, pp. 936-945.

Hill, Elisabeth and David Sally, Dilemmas and bargains: Theory-of-mind, cooperation and fairness, 2002, London, University College, pp.

Hoch, Stephen J. and George F. Loewenstein. 1991. "Time-inconsistent preferences and consumer self-control." *Journal of Consumer Research*, 17:4, pp. 492-507.

Hoffman, Elizabeth, Kevin McCabe, Keith Shachat and Vernon L. Smith. 1994. "Preferences, property rights and anonymity in bargaining games." *Games and Economic Behavior*, 7, pp. 346-380.

Johnson, Eric J., Colin F. Camerer, Sen Sankar and Talia Tymon. 2002. "Detecting failures of backward induction: Monitoring information search in sequential bargaining." *Journal of Economic Theory*, 104:1, pp. 16-47.

Kahneman, Daniel and Shane Frederick, Representativeness revisited: Attribution substitution in intuitive judgment, 2002 *in* "Heuristics of intuitive judgment: Extensions and applications" (T. Gilovich, D. Griffin and D. Kahneman, Eds.), New York Cambridge University Press, pp.

Kirkpatrick, Lee A. and Seymour Epstein. 1992. "Cognitive-experiential self-theory and subjective probability: Further evidence for two coneptual systems." *Journal of Personality and Social Psychology*, 63:4, pp. 534-544.

Kosslyn, Stephen M. 1994. *Image and Brain*. Cambridge, MA: MIT Press.

Laibson, David. 2001. "A cue-theory of consumption." *Quarterly Journal of Economics*, 116:1, pp. 81-120.

Laibson, David, Andrea Repetto and Jeremy Tobacman. 1998. "Self-control and saving for retirement." *Brookings Papers on Economic Activity*, 1, pp. 91-196.

Leboeuf, Robyn Aimee. 2002. "Alternating selves and conflicting choices: Identity salience and preference inconsistency." *Dissertation Abstracts International*, 63:2-B, pp. 1088.

LeDoux, Joseph E. 1996. *The emotional brain: The mysterious underpinnings of emotional life*. New York, NY: Simon & Schuster, Inc.

Lerner, Jennifer and Dacher Keltner. 2001. "Feer, anger, and risk." *Journal of Personality and Social Psychology*, 81, pp. 146-159.

Libet, Benjamin. 1985. "Unconsciious cerebral initiative and the role of conscious will in voluntary action." *Behavior and Brain Sciences*, 8, pp. 529-566.

Lieberman, M.D., Reflective and reflexive judgment processes: A social cognitive neuroscience approach, in press *in* "Responding to the social world: Implicit and explicit processes in social judgments and decisions" (J. P. Forgas, K. Williams and W. von Hippel, Eds.), Philadelphia Psychology Press, pp.

Lieberman, M.D., R. Gaunt, D. T. Gilbert and Y Trope, Reflection and reflexion: A social cognitive neuroscience approach to attributional inference, 2002 *in* "Advances in Experimental Social Psychology" (M. Zanna, Ed.), New York Academic Press, pp. 199-249.

Loewenstein, George, The fall and rise of psychological explanations in the economics of intertemporal choice, 1992 *in* "Choice over time." (G. Loewenstein and J. Elster, Eds.), New York Russell Sage Foundation, pp. 3-34.

Loewenstein, George. 1994. "The psychology of curiosity: A review and reinterpretation." *Psychological Bulletin*, 116:1, pp. 75-98.

Loewenstein, George. 1996. "Out of control: Visceral influences on behavior." *Organizational Behavior and Human Decision Processes*, 65, pp. 272-292.

Loewenstein, George F. and Jennifer Lerner, The role of affect in decision making, 2003 *in* "Handbook of Affective Sciences" (R. J. Davidson, K. R. Scherer and H. H. Goldsmith, Eds.), Oxford Oxford University Press, pp. 619-642.

Loewenstein, George F., Elke U. Weber, Christopher K. Hsee and Ned Welch. 2001. "Risk as feelings." *Psychological Bulletin*, 127:2, pp. 267-286.

Loewenstein, George, Baba Shiv, Antoine Bechara, Hanna Damasio and Antonio R. Damasio, The dark side of emotions in decision making: When a dysfunctional ventromedial prefrontal cortex can lead to more optimal decisions, 2002, Iowa City, University of Iowa, pp.

Lykken, David T. 1957. "A study of anxiety in the sociopathic personality." *Journal of Abnormal and Social Psychology*, 55, pp. 6-10.

MacLean, Paul D. 1990. *The Triune Brain in Evolution: Role in Paleocerebral Function*. New York: Plenum.

Maguire, Eleanor A., David G. Gadian, Ingrid S. Johnsrude, Atriona D. Good, John Ashburner, Richard S.J. Frackowiak and Christopher D. Frith. 2000. "Navigation-related structural change in the hippocampi of taxi drivers." *Proceedings of the National Academy of Sciences*, 97:8, pp. 4398-4403.

Mandler, George, The structure of value: Accounting for taste, 1982 *in* "Affect and cognition" (M. S. Clark and S. T. Fiske, Eds.), Hillsdale Lawrence Erlbaum, pp. 3-36.

Manuck, Stephen B., Janine Flory, Matthew Muldoon and Robert E. Ferrell, Is there a neurobiology of intertemporal choice, 2003 *in* "Time and decision: Economic and psychological perspectives on intertemporal choice" (G. F. Loewenstein, D. Read and R. Baumeister, Eds.), New York Russell Sage, pp.

McCabe, Kevin, Daniel Houser, Lee Ryan, Vernon Smith and Theodore Trouard. 2001. "A funtional imaging study of cooperation in two-person reciprocal exchange." *Proceedings of the National Academy of Sciences*, 98:20, pp. 11832-11835.

McElroy, Susan L., Andrew Satlin, Harrison Pope, Paul E. Keck and et al. 1991. "Treatment of compulsive shoping with antidepressants: A report of three cases." *Annuals of Clinical Psychiatry*, 3:3, pp. 199-204.

McNeil, Barbara, Stephen G. Pauker, Harold C. Sox and Amos Tversky. 1982. "On the elicitation of preferences for alternative therapies." *New England Journal of Medicine*, 306, pp. 1259-1262.

Medin, Douglas L. and Max H. Bazerman. 1999. "Broadening behavioral decision research: Multiple levels of cognitive processing." *Psychonomic Bulletin & Review*, 6:4, pp. 533-547.

Mendelson, Joseph. 1967. "Lateral hypothalamic stimulation in satiated rats: The rewarding effects of self-induced drinking." *Science*, 157:3792, pp. 1077-1079.

Miller, George A., Eugene Galanter and Karl H. Pribram. 1960. *Plans and ther structure of behavior*. New York: Holt.

Mischel, Walter, Ebbe B. Ebbesen and Antonette Zeiss. 1972. "Cognitive and Attentional Mechanisms in Delay of Gratification." *Journal of Personality and Social Pychology*, 21:2, pp. 204-218.

Montague, P Read and Gregory S. Berns. 2002. "Neural economics and the biological substrates of valuation." *Neuron*, 36:October, pp. 265-284.

Moreyra, Paula, Angela AIbanez, Jeronimo Saiz-Ruiz, Kore Nissenson and Carlos Blanco. 2000. "Review of the phenomenology, etiology and treatment of pathological gambling." *German Journal of Psychiatry*, 3, pp. 37-52.

Morris, J.S., C. Buchel and R.J. Dolan. 2001. "Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning." *NeuroImage*, 13:6, pp. 1044-1052.

Nisbett, Richard E. and Timothy D. Wilson. 1977. "Telling more than we can know: Verbal reports on mental processes." *Psychological Bulletin*, 84:3, pp. 231-259.

O'Donoghue, Ted and Matthew Rabin. 1999. "Doing it now or later." *American Economic Review*, 89:1, pp. 103-124.

Ochsner, Kevin, N. and Matthew D. Lieberman. 2001. "The emergence of social cognitive neuroscience." *American Psychologist*, 56:9, pp. 717-734.

Olds, James and Peter Milner. 1954. "Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain." *Journal of Comparative & Physiological Psychology*, 47, pp. 419-427.

Oswald, Andrew. 1997. "Happiness and economic performance." *Economic Journal*, 107, pp. 1815-1831.

Panksepp, J. 1998. *Affective Neuroscience*. Oxford: Oxford University Press.

Parsons, Lawrence M. and Daniel Osherson. 2001. "New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning." *Cerebral Cortex*, 11, pp. 954-965.

Persico, Nicola, Postlewaite Andrew and Dan Silverman, The efect of adolescent experience on labor market outcomes: the case of height, 2002, Philadelphia University of Pennsylvania, pp.

Persinger, Michael and Faye Healey. 2002. "Experimental facilitation of the sensed presence: Possible intercalation between the hemispheres induced by complex magnetic fields." *Journal of Nervous and Mental Disease*, 190:8, pp. 533-541.

Peters, Ellen and Paul Slovic. 2000. "The springs of action: Affective and analytical information processing in choice." *Personality and Social Psychological Bulletin*, 26:12, pp. 1465-1475.

Rolls, E.T. 1999. *The brain and emotion*. New York: Oxford University Press.

Romer, Paul M. 2000. "Thinking and feeling." *American Economic Review*, 90:2, pp. 439-443.

Ross, Lee, Mark R. Lepper and Michael Hubbard. 1975. "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm." *Journal of Personality and Social Pychology*, 32:5, pp. 880-892.

Rumelhart, David E. and James L McClelland. 1986. *Parallel distributed processing: Explorations in the microsctructure of cognition. Volume 1*: MIT Press.

Rustichini, Aldo, John Dickhaut, Paolo Ghirardato, Kip Smith and Jose V. Pardo, A brain imaging study of procedural choice, 2002, University of Minnesota, pp.

Sanfey, A.G., J.K. Rilling, J.A. Aaronson, Leigh E. Nystrom and J.D. Cohen, Probing the neural basis of economic decision making: An fMRI investigation of the ultimatum game, 2002, pp.

Schelling, Thomas C. 1978. "Egonomics, or the Art of Self-Management." *American Economic Review*, 68:2 May, pp. 290- 294.

Schelling, Thomas C., The Intimate Contest for Self-Command, 1984 *in* "Choice and Consequence", Cambridge, MA Harvard University Press, pp. 57-81.

Schneider, Walter and Richard M. Shiffrin. 1977. "Controlled and automatic human information processing: I. Detection, search and attention." *Psychological Review*, 84:1, pp. 1-66.

Schulz, Wolfram and A. Dickinson. 2000. "Neuronal coding of prediction errors." *Annual Review of Neuroscience*, 23, pp. 473-500.

Schwarz, Norbert, Feelings as information: Informational and motivational functions of affective states, 1990 *in* "Handbook of motivation and cognition: Foundations of social behavior" (E. T. Higgins and R. M. Sorrentino, Eds.), New York, NY Guilford Press, pp. 527-561.

Schwarz, Norbert and Gerald L. Clore. 1983. "Mood, misattribution and judgments of well-being: Informative and directive functions of affective states." *Journal of Personality and Social Psychology*, 45:3, pp. 513-523.

Sciences, National Academy of. 1999. *Pathological gambling*. Washington, DC: National Academy Press.

Shiffrin, Richard M. and Walter Schneider. 1977. "Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory." *Psychological Review*, 84:2, pp. 127-190.

Sloman, Steven A. 1996. "The empirical case for two systems of reasoning." *Psychological Bulletin*, 119:1, pp. 3-22.

Titmuss, R.M. 1970. *The gift relationship*. London, UK: Allen and Unwin.

Viner, J. 1925. "The utility concept in value theory and its critics." *Journal of Political Economy*, 33, pp. 369-387.

von Melchner, Laurie, Sarah L. Pallas and Mriganka Sur. 2000. "Visual behavious mediated by retinal projections directed to the auditory pathway." *Nature*, 404:April, pp. 871-876.

Vuilleumier, P., J.L. Armony, J. Driver and Ray Dolan. 2001. "Effects of attention and emotion on face processing in the human brian: An event-related fMRI study." *Neuron*, 30, pp. 829-841.

Wegner, Daniel M. and Thalia Wheatley. 1999. "Apparent mental causation: Sources of the experience of will." *American Psychologist*, 54:7, pp. 480-492.

Whalen, P.J., S.L. Rauch, Nancy Etcoff, S.C. McInerney, M. Lee and M.A. Jenike. 1998. "Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge." *Journal of Neuroscience*, 18, pp. 411-418.

Wilson, T. D., D. J. Lisle, J. W. Schooler, S. D. Hodges, K. J. Klaaren and S. J. LaFleur. 1993. "Introspecting about reasons can reduce post-choice satisfaction." *Personality and Social Psychology Bulletin*, 19, pp. 331-339.

Wilson, Timothy D., Samuel Lindsey and Tonya Y. Schooler. 2000. "A model of dual attitudes." *Psychological Review*, 107:1, pp. 101-126.

Wilson, Timothy D. and Jonathan W. Schooler. 1991. "Thinking too much: Introspection can reduce the quality of preferences and decisions." *Journal of Personality and Social Psychology*, 60:2, pp. 181-192.

Windschitl, Paul and Gary Wells. 1998. "The alternative-outcomes effect." *Journal of Personality and Social Psychology*, 75, pp. 1411-1423.

Wise, Roy A. 1996. "Addictive drugs and brain stimulation reward." *Annual Review of Neuroscience*, 19, pp. 319-340.

Wolford, George, Michael B. Miller and Michael Gazzaniga. 2000. "The left hemisphere's role in hypothesis formation." *Journal of Neuroscience,* 20, pp. 1-4.

Zajonc, R., Emotions, 1998 *in* "The Handbook of Social Psychology" (D. Gilbert, S. Fiske and G. Lindzey, Eds.), New York Oxford University Press, pp. 591-632.

Zajonc, R. B. 1980. "Feeling and thinking: Preferences need no inferences." *American Psychologist*, 35, pp. 151-175.

Zajonc, R. B. 1984. "On the primacy of affect." *American Psychologist*, 39:2, pp. 117-123.

Zajonc, Robert B. and Daniel N. McIntosh. 1992. "Emotions research: Some promising questions and some questionable promises." *Psychological Science*, 3:1, pp. 70-74.