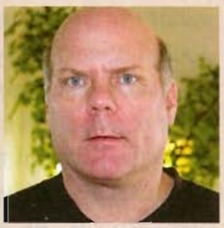
'Neuroeconomics' uses knowledge about brain mechanisms to inform economic theory, opening up the 'black box' of the brain.

by George Lowenstein, Colin Camerer and Drazen Prelec

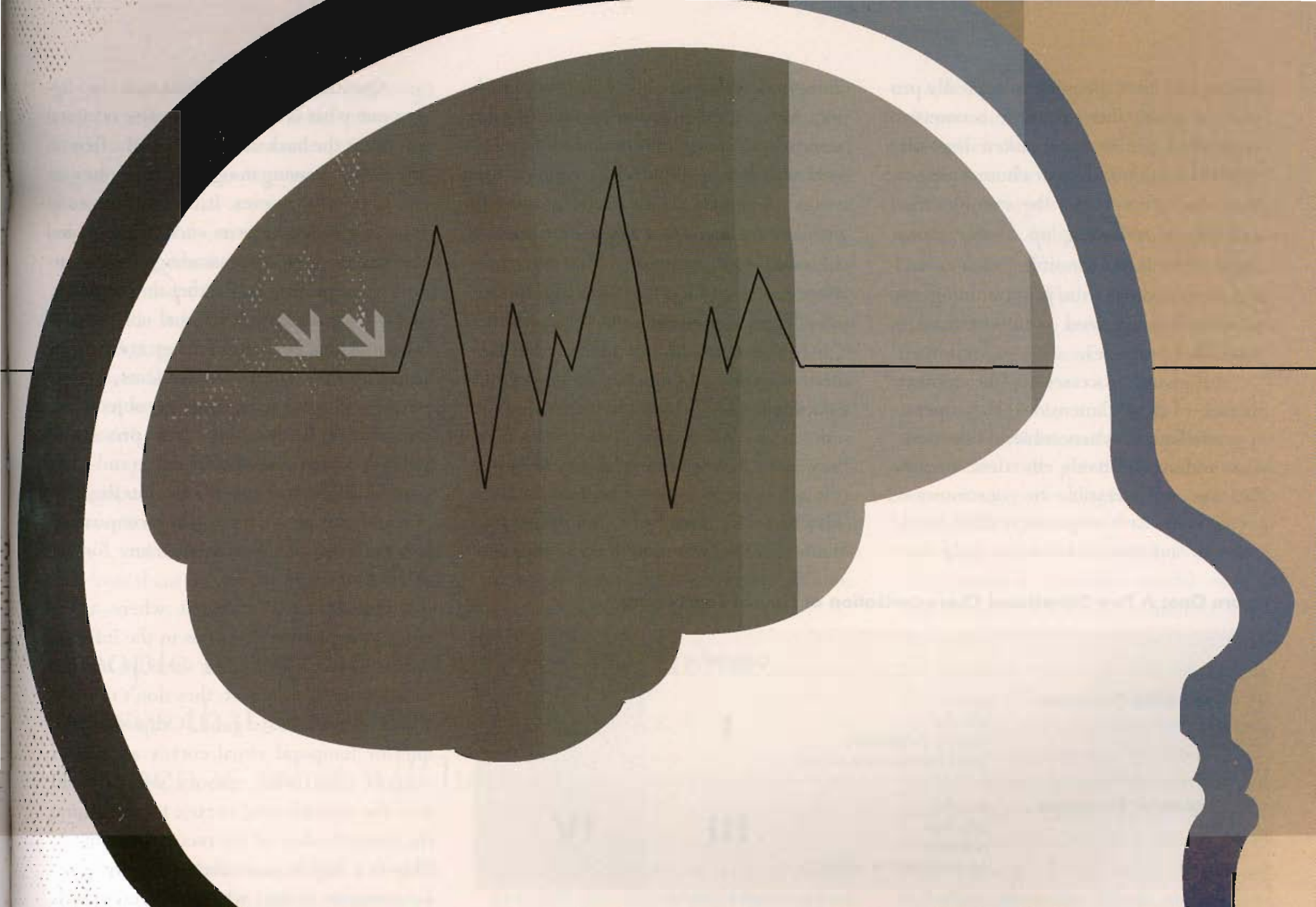

NEURO
How Neuroscience

The foundations of economic theory were constructed assuming that details about the functioning of the brain – the ultimate 'black box' – would never be known. This pessimism was expressed by economist **William Stanley Jevons** in 1871: "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 these feelings that we must estimate their comparative amounts."

In the last two decades, however, behavioural economics (the importation of ideas from psychology to economics) has become a prominent fixture on the intellectual landscape, and has, in turn spawned the new field of 'neuroeconomics'. Important insights are now emerging that prove Jevons' pessimism wrong: the study of the brain and nervous system is beginning to allow for direct measurement of thoughts and feelings; and this, in turn, is challenging our understanding of the relation between mind and action, leading to new theoretical constructs and calling old ones into question.

Automatic and Emotional Processing

The standard economic theory of 'constrained utility maximization' is often interpreted as a model of deliberation – a balancing of the costs and benefits of different options, as might characterize complex decisions like planning for retirement or buying a house. While not denying that deliberation is part of human decision making, neuroscience research points out two generic inadequacies of this approach: the important roles of *automatic* and *emotional* processing.



ECONOMICS Can Inform Economics

First, much of the brain is constructed to support 'automatic processes', which are faster than conscious deliberations and occur with little or no awareness. Because people have little or no introspective access to these processes or volitional control over them, and these processes were evolved to solve problems of evolutionary importance rather than to respect logical dicta, the behaviour they generate need not conform to normative axioms of inference and choice – and hence cannot be adequately represented by the usual maximization models.

Second, our behaviour is often strongly influenced by finely-tuned affective (i.e. emotional) systems whose basic design is common to humans and many animals. These systems are essential for daily functioning, as revealed by the fact that when affective systems are damaged or perturbed – by brain damage, stress, imbalances in neurotransmitters, or alcohol – the deliberative system generally caves in, and poor long-term decisions result.

Human behaviour emerges from the interplay between *controlled* and *automatic* systems on the one hand, and *cognitive* and *affective* systems on the other. Our goal here is to show that drawing these distinctions helps to make sense of a wide range of economic phenomena that are difficult to explain in conventional terms.

Controlled vs. Automatic Processes

Controlled processes use step-by-step logic or computations, tend to be invoked deliberately by the individual when he or she encounters a challenge or surprise, and are often associated with a subjective feeling of effort. People can typically provide a good introspective account of controlled processes; if asked how they solved a math problem or chose a new car, they can often recall the considerations and the steps leading up to the choice. Standard tools of economics, such as decision trees and dynamic programming, can be viewed as stylized representations of controlled processes.

Automatic processes are the opposite on each of these dimensions: they operate in parallel, are not accessible to consciousness, and are relatively effortless. Because they are not accessible to consciousness, people often have surprisingly little insight into why automatic choices or judgments were made. A face is perceived as 'attractive' or a verbal remark as 'sarcastic' automatically and effortlessly. It is only later that the controlled system may reflect on the judgment and attempt to substantiate it logically.

Automatic processes – whether cognitive or affective – are the default mode of brain operation: they whir along all the time, even while 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.

Cognitive vs. Affective Processes

The second distinction to be made is between *affective* and *cognitive* processes. Such a distinction is pervasive in contemporary psychology and neuroscience, and has a historical lineage going back to the ancient Greeks and earlier. The distinguishing features of affective processing are somewhat counterintuitive. Most people undoubtedly associate affect with 'feeling states', and indeed, most affect states do produce feeling states when they reach a threshold level of intensity. However, most affect probably operates below the threshold of conscious awareness.

For most researchers, the central feature of affect is not the feeling states associated with it, but its role in human motivation. All affects have 'valence' – they are either positive or negative (though some complex emotions, such as 'bittersweet', can be a combination.) Many also carry 'action tendencies'. For instance, anger motivates us to aggress, pain to take steps to ease the pain, fear to escape, and so on. Affective processes address 'go/no-go' questions that motivate approach or avoidance behaviour. Cognitive processes, in contrast, are those that answer true/false questions.

Figure One: A Two-Dimesional 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

A Fourfold Classification of Neural Processes

In combination, the two dimensions described above define four quadrants of the brain, as shown in **Figure One**. Quadrant I is in charge when you deliberate about whether to refinance your house, poring over present-value calculations; Quadrant II is the rarest in pure form; it 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'!

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. Further downstream, in the inferior temporal visual cortex, 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 extraordinarily complicated (and has proved difficult for artificial intelligence researchers to recreate in computers) because 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 'reward value' of the recognized object. This is a highly-particular representation. In economic terms, what is represented is neither pure information (i.e., that this is sushi) nor pure utility (i.e., that it is something I like) but rather a fusion of information *and* utility. It is as if certain neurons in the orbitofrontal cortex are saying, "this is sushi and I want it."

The reward value of sushi depends in turn 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. 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. Neurons in these regions fire more rapidly at the sight or taste of food when you are hungry, and fire less rapidly when you are not hungry.

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 you will 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 host, who made it herself, you'll eat it anyway (or pick it up and hide it in a napkin when she turns around.) These explicit thoughts involve anticipated feelings (your own and the host's) and draw on explicit memories from a part of the brain called the hippocampus, inputs from the affective system (sometimes referred to as the 'limbic system'), and anticipation (planning) from the prefrontal cortex.

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. For instance, we find ourselves almost compulsively eating our children's left-over Halloween candy, while obsessing about how to lose the extra 10 pounds; or gambling recklessly at the casino, even as a small voice in our head tells us to stop.

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. 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. The best-developed of these approaches is 'affect-as-information theory': at intermediate levels of intensity, people begin to become conscious of conflicts between cognitive and affective inputs. 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.

Finally, at even greater levels of intensity, affect 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 affective motivation, people often report themselves as being 'out of control' or 'acting against their own self-interest'. As **Rita Carter** writes in *Mapping the Mind*, "where thought conflicts with emotion, the latter is designed by the neural circuitry in our brains to win."

Sense-making is an important form of interaction between Quadrant I and the other quadrants. The brain's powerful drive toward sense-making leads us to strive to interpret our own behaviour. Since Quadrant I often does not have conscious access to activity in the other quadrants, it is perhaps not surprising that it tends to over-attribute behaviour to itself. 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.

People are good at persuading themselves that what they would *like* to happen is what *will* happen.

Although it is useful to distinguish between cognitive and affective processes, and between controlled and automatic processes, most judgments and behaviours result from interactions between them. Many decision-making disorders may originate in an improper division of labour between the quadrants. For instance, 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'.

Affect can also distort cognitive judgments. For example, emotions affect our perceptions of risks – anger makes people *less* threatened by risks, and sadness makes them *more* threatened. 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 hopeful 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, trading in financial markets, under-saving, and low rates of investment in education.

General Implications for Economics

Neuroscience can point out commonalities between categories that had been viewed as distinct. One example of this with important implications is the utility for money. The canonical economic model assumes that the utility for money is indirect – i.e., that money is a mere counter, only valued for the goods and services it can procure. Thus, standard economics would view, say, the pleasure from food and the 'pleasure' from obtaining money as two totally different phenomena. Neural evidence suggests, however, that the same reward circuitry of the brain in the midbrain is activated for a wide variety of different reinforcers, including attractive faces, sports cars, drugs, and money.

The idea that many rewards are processed similarly in the brain has important implications for economics, which assumes that the marginal utility of money depends on what money buys. For example, compensation of top business executives skyrocketed in the 1990s. While improvements in linking performance to pay and social comparison undoubtedly explain part of the growth, one of the puzzles is why executives need so much money. However, if abstract money rewards fire dopamine neurons, as addictive drugs do, then larger and larger money rewards become desirable even if the money won't actually be used to purchase pleasure-producing consumption.

If gaining money provides direct pleasure, then the experience of parting with it is probably painful. While there is no direct evidence of this, the assumption that paying hurts can explain many market phenomena which are otherwise puzzling. An example is the effects of payment-neutral pricing schemes on choices.

Companies often go to great lengths to disguise payments, or reduce their pain. Consumers appear to oversubscribe to flat-rate payment plans, e.g., for utilities, telephone services and health clubs. A flat-rate plan eliminates marginal costs, and allows consumers to enjoy the service without thinking about the marginal cost. Similarly, travel plans are often sold as packages, making it impossible to compute the cost of the individual components (hotel, food, transportation). Often components of the package are presented as 'free' (like Microsoft's internet browser) even though the claim is meaningless from an economic standpoint, given that the package is presented on a take-it-or-leave-it basis. One can interpret the appeal of ad-hoc currencies, such as frequent-flyer miles, or chips in casinos, as an attempt to reduce the pain-of-payment. The ad-hoc currency feels like 'play money,' and spending it does not seem to exact the same psychic cost.

Specific Applications

Both collaboration and competition between affect and cognition, and between controlled and automatic processes, can also be seen in the domain of 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 toward 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 **Figure One**, 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 behaviour.

The existence of separate affective and cognitive systems that respond differently to risks is most salient when the two systems clash. People are often 'of two minds' when it comes to risks: we drive 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 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.

Emotional reactions to risk can help to explain risk-seeking as well as risk-aversion.

Fear unleashes preprogrammed sequences of behaviour 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.

Emotional reactions to risk can help to explain risk-seeking as well as risk-aversion. 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 rein-in risk-taking. Indeed, about one per cent of the people who gamble are diagnosed as 'pathological': they report losing control and harming their personal relationships by gambling. Standard economic explanations for gambling, such as- convex utility for money, don't help explain why some gamblers binge and don't usefully inform policies to regulate the 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 that causes gamblers to seek ever-larger thrills to get modest jolts of pleasure is more likely to be present in pathological gamblers than in 'normal' people. One study shows tentatively that treatment with a drug that blocks the operation of opiate receptors in the brain reduces the urge to gamble. The same drug has been used to successfully treat compulsive shopping.

Conclusions

In many areas of economics there are basic constructs at the heart of current debates which can be usefully thought of as 'neural processes', and studied using the tools of neuroscience. Finance is a field awash in literally millions of observations of daily price movements. Despite having widespread access to extensive data, after decades of careful research there is no agreed-upon theory of why stock prices fluctuate, why people trade, and why there are so many actively-managed mutual funds, despite their poor performance. Perhaps knowing more about the basic neural mechanisms that underlie fear and greed, conformity, wishful thinking, sense-making of random series, and perceptions of expertise can help explain these puzzles.

Although we have focused here on applications of neuroscience to economics, intellectual trade could also flow in the opposite direction. Even simple ideas in economics could be useful for informing neuroscience and shifting the nature of the questions that are asked. For instance, neuroscientists do not understand how the brain allocates resources that are essentially fixed, such as 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 diverse specialized systems that make up the brain interact.

George Loewenstein is an economist teaching in the Department of Social and Decision Sciences at Carnegie-Mellon University. Colin Camerer is the Rea and Lela Axline Professor of Business Economics at the California Institute of Technology. Drazen Prelec is a professor of Management Science at MIT's Sloan School of Management.
This is an excerpt of an article that recently appeared in the *Journal of Economic Literature*. For a copy of the complete paper, e-mail christen@rotman.utoronto.ca