

Gruter Institute Project on Values and Free Enterprise

With support from:
John Templeton Foundation
UCLA – Sloan Research Program
Ann & Gordon Getty Foundation

Values and Value: Moral Economics

Paul J. Zak
Claremont Graduate University
Neuroeconomics Studies and Department of Economics

Loma Linda University
Department of Neurology

This chapter is part of a collection posted on the SSRN website in the Economics Research Network section located at -
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalBrowse&journal_id=926595

Values and Value: Moral Economics

By

Paul J. Zak*

Center for Neuroeconomics Studies and Department of Economics
Claremont Graduate University

Department of Neurology
Loma Linda University

and

Gruter Institute for Law and Behavioral Research

Forthcoming in:

Moral Markets: The Critical Role of Values in the Economy

Paul J. Zak, Editor

Princeton University Press

* Center for Neuroeconomics Studies, 160 E. 10th St., Claremont Graduate University, Claremont, CA 91711-6165, paul@pauljak.com. Support from the John Templeton Foundation and the Gruter Institute for Law and Behavioral Research are gratefully acknowledged.

We need ...an evolutionary theory of morals...and (its) essential feature will be that morals are not a creation of reason, but a second tradition independent from the tradition of reason, which helps us adapt to problems which exceed by far the limits of our capacity of rational perception.

-Friedrich von Hayek, in a speech at Oesterreichisches College, Wien, 1985

I. Introduction

J. Clifford Baxter graduated from New York University and then served in the Air Force, obtaining the rank of Captain in 1985. After leaving the military, he completed an MBA at Columbia University and began a business career. In 1991, he joined a small oil and gas pipeline company in Houston, Texas called Enron. Cliff Baxter steadily climbed the corporate ladder, becoming the Vice-Chairman of Enron in October, 2000. During Cliff Baxter's tenure, Enron began trading energy contracts and grew to become the nation's seventh largest corporation. Toward the end of the 1990s, Baxter began complaining to Enron's CEO Jeff Skilling about their business practices. In a letter to Chairman Kenneth Lay, former Enron Vice-President Sherron Watkins said that "...Cliff Baxter complained mightily to Skilling and all who would listen about the inappropriateness of our transactions..." (CNN.com, 2002). Baxter resigned from Enron in May, 2001, but continued to work for the company as a consultant for the next half-year. Between October, 1998 and his resignation, he exercised \$22 million in stock options.

By all accounts, Cliff Baxter lived by a code of high moral values. He served with distinction in the military, he had a happy family life with his wife and two children. He was among the few employees at Enron who left a paper trail establishing criticisms of the company's ethical transgressions and legal abuses. Early in the morning of January 26, 2002, Baxter wrote the following note to his wife.

Carol,

I am so sorry for this. I feel I just can't go on. I have always tried to do the right thing but where there was once great pride now its gone. I love you and the children so much. I just can't be any good to you or myself. The pain is overwhelming. Please try to forgive me.

Cliff.

This note was found in his car, parked near his house, with a self-inflicted fatal gun-shot wound to the head.

Cliff Baxter was driven to succeed and worked incredibly hard to achieve success at Enron. But even resigning from Enron did not sufficiently erase the pain he appeared to have felt for violating a moral code, and he chose to end his life rather than endure the pain of this violation.¹ There must have been many Enron employees like Cliff Baxter who were fine citizens, driving the speed limit, paying their taxes, and who wouldn't think stealing from their local grocery store. For example, Enron Treasurer Jeff McMahon wrote to the CEO that "My integrity forces me to negotiate the way I believe is correct" when discussing questionable shell corporations that were set up apparently

¹ At the time of his death, Cliff Baxter was taking a number of psychotropic medications, including anti-depressants.

to directly benefit senior management (McLean and Elkin, 2003, p.210). Yet most employees uncomfortable with Enron's business practices did not resign from Enron, and no others are known to have killed themselves.²

There are three leading explanations for the unethical and likely illegal behavior of a large number of Enron employees. The first is that the process of economic exchange values greed and self-serving behaviors and inadvertently produces a society of rapacious and perhaps evil people. An extension of this line of argument is that all those living in modern economies are dehumanized, a view popularized by socialists like Karl Marx. The amount of anonymous charitable giving each year by individuals living in developed economies belies this claim. As I write this, Hurricane Katrina had recently hit the gulf coast of the United States, devastating New Orleans and many other cities. Donations to the Red Cross and other relief organizations topped \$1 billion within weeks of this disaster. Similar aid was provided worldwide for victims of the Asian Pacific Tsunami of December 26, 2004. If humans in modern economies are so greedy, why all the charity?

A second explanation for the behavior of Enron employees is that there could be a selection bias in which amoral greedy people were hired in key posts and this behavior filtered down to other employees. An implication of this view is that government regulation is critical to keep these "bad eggs" in check. After the accounting scandals in the first years of the 21st century, the United States passed the Sarbanes-Oxley Act to tighten accounting standards and hold corporate board members personally liable for corporate misdeeds. This additional regulation is costly to both the federal government and to firms, potentially reducing aggregate economic productivity. Is such regulation necessary if the proportion of violators is small? For example, the fraction of those who cheat on their taxes in the US is small, even though the chance of being prosecuted for tax fraud is nearly zero (Slemrod, 1992). Indeed, why don't all or most employees steal, cheat and lie to get ahead? Neither the chance of being caught nor the Marxist view of the world adequately explain the mostly decent behavior by employees, managers, and citizens in most societies.

A third explanation is that there was a particular institutional environment at Enron that, for example, drove employees to cheer forest fires that were driving up energy prices while destroying people's homes and lives. This explanation nests the two previous ones—senior managers at Enron devised the company's procedures and compensation that pitted employees against each other for survival and provided incentives to violate accounting standards. One way this was done was to break up decisions into parts so most individuals were only responsible for moving the decision forward and could not claim ultimate responsibility for an action. Given this environment, many employees acquiesced to the incentives that resulted in unethical and possibly illegal behavior. This explanation is more subtle than the first two, as its base assumption is that most people behave ethically most of the time. Nevertheless, in the right circumstances many people can be induced to violate what seems to be an internal representation of values that holds unethical behavior in check.

So which explanation is right? And perhaps more importantly, is there a mechanism at work that sustains economic order in a highly decentralized moderately regulated economy like in the US? This chapter will survey neuroscientific research and

² Daniel Watkins, a consultant to Enron and other companies who worked for Arthur Andersen, killed himself in December, 2002.

discuss recent experiments from my lab on the physiologic basis for interpersonal decision-making to support the follow thesis: most people, most of the time, behave ethically, and that a set of shared values is essential to the functioning of modern economies. I call this “moral economics.” Next, I will draw implications for law, institutional design, and public policy that follow from moral economics. What I would like to demonstrate is that values are a critical ingredient in producing the historically high living standards in many nations, and are a prerequisite to economic growth in the developing world.

2. Economics and Values

2.1 What are values?

Values are guides to action. They are enduring beliefs more basic than heuristics, but can be seen as building blocks for heuristics. For example, Aristotle identified eight values: courage, temperance, liberality, magnificence, proper pride, good temper, modesty, and friendliness. Two thousand years later, Benjamin Franklin recognized thirteen values: temperance, silence, order, resolution, frugality, industry, sincerity, justice, moderation, cleanliness, tranquility, chastity, and humility. Values can be thought of as the constituents of a person’s character traits. Some values appear to be universal (for more on this, see Robert Solomon’s and Lynn Stout’s chapters in this volume). Schwartz (1994) identifies a set of broad values that appear to be held in nearly all societies from his a survey of 44 countries. A value to respect other’s lives, if generally held, offers one protection (in most circumstances) for one’s own life. Other values similarly support social probity. Following one’s values may generate internal rewards as well. Psychologically, one may consider oneself a “good” person when one conforms to shared values.

Rokeach (1973) distinguishes between instrumental and terminal values. Instrumental values are the means to achieve a goal. Examples include being ambitious, courageous, honest and loving. Terminal values are end goals. Examples of terminal value include freedom, security, pleasure, and prosperity. Values—especially instrumental values—can further be broken down into those that are primarily personal and those that are interpersonal or social. The latter are often called moral values, where “morality” comes from the broad acceptance or prohibition of a behavior within or even across societies. As I will discuss below, engaging in, or observing violations of, moral values typically produces a strongly felt physiologic response. Values are therefore a motivation for behavior.

Values, especially in young people but also in adults, evolve with experience. Nevertheless, values are typically stable within individuals, though variation across individuals is common. A defining characteristic of values is that they constrain our choices. That is, values are exercised at a cost. An example is that most people, absent psychopathology, have a deeply held prohibition against killing others. This means that the person talking on the cell phone who is holding up traffic cannot be killed (even though we might fleetingly desire this; David Buss (2005) reports that 91% of men and 84% of women have had at least one vivid homicidal fantasy). In most societies, the prohibition against killing is lifted for soldiers fighting other soldiers, police officers chasing dangerous criminals, and anyone in imminent mortal danger from another. Values, then, proscribe behavior and are followed at a cost. Values may be limited to

those in one's social group; "enemies", "slaves", and "savages" have, at certain times in history, been considered subhuman and therefore values can be violated when interacting with them. In any situation, several values may interact or even conflict when deciding how to behave. Further, the weight put on instrumental vs. terminal values varies across individuals (Schwartz, 1994), but less so across settings.

This leads one to ask how values might have evolved. One hypothesis for the rapid and extraordinary growth of the human brain is that this occurred to support increasingly complex social behaviors (Reader & Laland, 2002). In particular, cooperation with nonkin is the hallmark of modern civilizations. Cooperation between unrelated individuals enabled the specialization of labor and the generation of surplus in societies, fueling technological advances and increasing living standards (Diamond, 1997). If this supposition is correct, we must ask how humans evolved the ability to behave cooperatively with nonkin. Neuroscientist Cort Pedersen (2004) argues that the same mechanism that in mammals facilitates care and attachment for offspring, permitted early humans to "attach" and cooperate with extended family and eventually nonkin. Pedersen hypothesizes that cooperation with nonkin in humans is facilitated by the neuroactive hormone oxytocin, as this molecule promotes attachment to offspring and, in monogamous mammals, attachment to reproductive partners. Most humans are serially monogamous, so extending the attachment role of oxytocin to nonkin in human is reasonable. Further, humans have an unusually long period of adolescence, requiring a strong and long-term mechanism of attachment (chimpanzees, our closest genetic relatives, are sexually mature at about eight years of age, while humans need at least 12 years before they can reproduce, and in modern societies typically much longer).

As a result of this powerful physiologic attachment in humans, we attach to friends, coworkers, pets, and even ideas (Zak, 2006). An example may clarify this idea. Most of us, if we were to see a stranger trip and break an ankle, would have an immediate physiologic reaction in which we would internally represent the other's pain. This can be called empathy, and for many people will motivate an offer to help this stranger. Indeed, it is difficult to suppress this response, but not impossible. A paramedic or physician observing the broken ankle would quickly suppress the feeling of empathy and get to work assisting the victim medically.

The shared representation of another's pain appears to be the causal mechanism through which we cooperate with strangers. Bowles and Gintis (2004) have called this the evolution of strong reciprocity—the ability to cooperate with others with the expectation that they will reciprocate. This impulse to reciprocate is based on an internal guidance system, most likely utilizing oxytocin. The associated activation of brain regions is the neural representation of values. The neural mechanisms supporting values are detailed in Sections 2.4 and 2.5 below.

2.2 *Values are costly but also have benefits*

Values can be construed to impose a cost on individuals. If one of my values is concern for the environment, then I may choose to purchase a hybrid gas-electric vehicle for a premium price, or perhaps to buy other "green" products that may be more expensive than standard products (at least initially). As an example, I once served as a consultant to the electric power industry on the market prospects for emerging renewable energy appliances such as solar air conditioning. Even if the cost of electricity

went up five-fold, most of these appliances would have to run for 30 to 50 years before they would be more cost-efficient than existing technologies. This is partially due to cost to build each unit (which is high because the demand for units is low)—a vicious circle. This could be altered by, for example, government subsidies to adopt these technologies, but the subsidies would be very expensive for many years. As a society, we may still choose to subsidize renewable energy appliances, even though this is economically inefficient, if that is an overriding value of the majority (in a democracy). In other words, societies may be willing to pay to follow their values. A simpler example is that in the United States we are willing to spend large sums of money to save single individuals who are in immediate danger, for instance the stranded hiker trapped on a snowy mountain, or the child who falls down an abandoned well. And yet at the same time, as a society we have chosen not to offer explicit health insurance to a large segment of the population, undoubtedly resulting in accelerated mortality for these people. This is probably due in part to the American value of self-reliance, but it has costs.

Because values are costly, there must be associated benefits if values are to be followed. Individuals may follow their values for a variety of reasons—to obtain the internal reward associated with following values; to avoid the internal pain associated with violating values; to acquire external benefits of exhibiting values, such as increased social status within one's peer group; and to be able to influence others to acquire or follow a certain set of values. There is some evidence that choosing to follow one's values may generate activation in regions of the human brain associated with reward (O'Doherty, et al., 2001). As a result of these factors, in many instances, individuals appear willing to follow what appears to be an internal mechanism in which the cost of values is compensated by an internal or external benefit (more on this point in Section 2.4 below).

Values may also have social benefits. In general, if most employees follow generally accepted social values (don't steal, don't lie, etc.), then monitoring costs by firms and the government are reduced and substantial deadweight losses are avoided. As a society, we seek to balance the costs of monitoring or not monitoring individuals or businesses. For example, the audit rules for the Internal Revenue Service are implicitly based on the understanding that in most cases direct monitoring is not necessary. What is needed is some likelihood that random audits and other methods to discover tax cheats are able to capture some proportion of violators.

This insight into human behavior—that most people will follow the law most of the time—was only formalized by the US federal government in the last 25 years. In the U.S., a cost-benefit analysis of new regulation and legislation prior to enactment was implemented by Executive Order 12291, signed by President Ronald W. Reagan on February 17, 1981. This order was strengthened by further executive orders over the next decade by Presidents Reagan and George H.W. Bush. The requirement of cost-benefit analysis was extended to a wider set of legislation by President William J. Clinton in 1993 (Executive Order 12866). One goal of these directives is use the optimal amount of regulatory oversight since this is costly and in some cases unnecessary. An analogy to understand the cost and benefits of regulatory monitoring is manufacturing quality control. Typically only a small subset of manufactured objects are monitored for quality, often using sophisticated algorithms to minimize sampling costs while maintaining a high average quality.

While cost-benefit analyses of monitoring are a move towards moral economics, they typically ignore how laws might counter-intuitively actually encourage violations. An example is the rule that one must pick up one's child from day care no later than a certain time of day. If this rule is violated, a teacher must stay late to watch the child, and the child may be stressed by having a late parent. In a recent experiment run at two daycare centers in Israel, both of which had a rule that parents must pick up their children no later than 4pm, one imposed a fine of \$3 for each time the child was picked up late. The other daycare center simply stayed with the stated policy that depended on the parents' following the rule absent a sanction. At the daycare with the fine, parents' mindset seemed to have changed; the fine seemed to remove the implicit social sanction associated with being late because now one just paid a penalty. Over the course of three weeks, the daycare with the fine saw a doubling of the parents who came late and proportion of latecomers stayed steady thereafter, even after the fines were terminated! (Gneezy & Rustichini, 2000). The lesson here is that oversight and penalties may crowd out the good behaviors that most people, most of the time, follow.

2.3 Moral Sentiments

Of Adam Smith's two great books, *The Theory of Moral Sentiments* (TMS) is typically considered much less important than *The Wealth of Nations*, though this view is starting to change. In TMS, Smith identified the social nature of human beings and sought to understand the nature of, and motivation for, morality. Smith used the notion of sympathy as a primary psychological feature that guided moral behaviors. He argued that we "feel bad" when we violate a moral value and this produces a desire to avoid this feeling in the future. The source of sympathy according to Smith, is the ability to identify with the emotions of others.

Man, say they, conscious of his own weakness, and of the need which he has for the assistance of others, rejoices whenever he observes that they adopt his own passions, because he is then assured of that assistance; and grieves whenever he observes the contrary, because he is then assured of their opposition. But both the pleasure and the pain are always felt so instantaneously, and often upon such frivolous occasions, that it seems evident that neither of them can be derived from any such self-interested consideration.

-Adam Smith, *The Theory of Moral Sentiments* 1759

Smith suggested that this emotional correspondence was innate and impossible to suppress. The quote from Hayek with which I started this article identifies a similar noncognitive mechanism for moral values.

Smith's view, has two primary parts. First, the mechanism that supports virtuous behaviors is primarily affective, not cognitive. Second, there appear to be universally-shared moral behaviors that serve as the backbone for a well-functioning society. It is worth noting that Smith had these exceptional insights 250 years ago. These ideas were essentially ignored until the last 50 years, and rigorous tests of Smith's proposed affective mechanism are less than 10 years old.

2.4 Neural Mechanisms of Moral Sentiments

There are several brain mechanisms that together seem to function as a moral compass, guiding us on appropriate modes of behavior in our daily interactions with other humans. The first mechanism occurs when we observe other's motor movements. Observing others' movements involuntarily activates in the observer's brain regions associated with the planning and execution of such movements, even when the observer is not moving at all. These neurons have been called "mirror neurons" (Rizzolatti et al., 1996) as they appear to produce an internal simulation in the observer of the occurring action. Mirror neurons have been found in primates, including humans (Rizzolatti et al., 2001; Grezes & Decety, 2001) and in some birds.

A second mechanism helps us infer the cognitive state of others. It has been called "mentalizing" or having a "theory of mind" (ToM). ToM permits us to forecast the beliefs and intentions of others by putting ourselves into the other's place and asking what we would do in such a situation. This is an extraordinarily useful ability when interacting with others, especially strangers. Children under five years old are unable to mentalize and impairments in ToM are common in those with autism. ToM has been localized to the medial prefrontal cortex (Frith & Frith, 2003), and has been shown by neuroeconomists to affect economic decision-making (McCabe et al., 2001; Zak, 2004; Camerer & Bhatt, 2005).

A third way the brain interprets what others are doing is through affective representation. This can occur when we observe an action occurring (a photo of a nail in a hand), by observing a facial expression (fear, disgust), or simply by the knowledge that another is being harmed in some way (Singer et al., 2004a; Wicker et al., 2003; Canli & Amin, 2002; Decety & Chaminade, 2003). The areas activated vary by study, but all include limbic regions associated with emotional responses. This internal representation of others' emotions can be called empathy.

An example of a representations of others' emotional states is an informal experiment I did with the members of my lab. A week after Hurricane Katrina, I asked who had donated money to the relief effort. Several students raised their hands, and then I asked them why they had given money. Each person could offer an explicit reason, often with an accompanying image of human suffering when recounting their motivation. And each clearly displayed emotion on the telling. Even with spatial and temporal distance, others' emotions are felt in ourselves and influence behavior.

Interestingly, emotional responses can be provoked absent direct viewing of an individual. For example, when a person makes an intentional monetary sacrifice signifying that he or she trusts a stranger, the brain of the person being trusted produces a surge in the neuroactive hormone oxytocin (Zak, Kurzban & Matzner, 2004; 2005). In numerous studies of rodents, oxytocin has been shown to facilitate attachment to offspring and in some monogamous mammals prosocial behaviors toward unrelated conspecifics. Oxytocin is a physiologic signature of empathy, and appears to induce a temporary attachment to others (Zak, 2006; Carter, et al., 2006). Infusing the human brain with a moderate dose of exogenous oxytocin can induce people to trust strangers with one's money to a much greater degree relative to those receiving a placebo (Kosfeld et al., 2005). High densities of oxytocin receptors are located in regions of the brain associated with emotions (Zak, Park, Ween, & Graham, 2006).

Oxytocin appears to facilitate a representation of what another is feeling. These mechanisms are automatic and by-and-large beyond our conscious control. That is not to say that these mechanisms are not modulated by the external and internal

environment. For example, during episodes of extreme stress, other-regarding behavior is suppressed as survival of the individual becomes paramount. This is analogous to starving Hurricane Katrina victims breaking into convenience stores to obtain food—there is no moral (or legal) prohibition against this in times of crisis. This may be one explanation for the lack of moral behavior by those at Enron—the enduring stress of “making the numbers” caused them, at some point, to make up the numbers. This moral violation appears to be modulated when others nearby are doing the same thing. A similar explanation is given for “ordinary” German citizens who tortured and killed Jews under the Nazi regime—their neighbors were doing the same thing so it became acceptable. The same argument holds for the genocides in Cambodia, Rwanda, Kosovo, Armenia, etc. Social psychologist James Waller calls this “moral disengagement” (Waller, 2002).

There is great heterogeneity across people in the ability to empathize with others and in the associated neural activation during emotionally engaging activities (Canli et al, 2001; Singer et al, 2004a). For example, I studied 212 subjects making trusting decisions and showed that their brains released oxytocin approximately proportional to the intentional monetary signal of trust received from a stranger: the stronger the signal of trust, the more oxytocin is released. Approximately 98% of these subjects also had proportional behavioral responses: the higher their oxytocin levels, the more they shared money with the person who initially demonstrated trust in them. But the other 2% of subjects, though their brains produced a surge of oxytocin, were untrustworthy, keeping all or nearly all of the pot of money they controlled (Zak, 2005). Two percent is roughly the proportion of sociopaths in the population. A discussion of the neural mechanisms that produced this behavior are beyond the scope of this chapter, but is worth noting that the relative contributions of nature and nurture are not well understood (see Zak, 2006). In addition, there are interesting gender differences in physiologic responses associated with distrustful behaviors that in men are driven by testosterone (Zak, Borja et al, 2005; Dabbs & Dabbs, 2000; Toussaint & Webb, 2005). The point here is that affective responses to social stimuli vary across individuals and quite likely within an individual as circumstances change.

2.5 Tests of the Neural Mechanisms of Moral Sentiments

Explicit tests of the brain regions active during moral decisions have been carried out recently by a number of researchers. I present a brief summary of this work to emphasize that there is direct evidence for the neural substrates of moral emotions. The relevance of this section to the paper’s thesis is that many values draw upon moral emotions.

The first neuroimaging study of moral decisions was led by Joshua Greene and colleagues (2001). They asked subjects to answer questions about how they would behave in a series of personal and impersonal moral dilemmas drawn from philosophy. Personal moral dilemmas take the form “Would you do something directly to harm another to save a group of people from certain harm.” Impersonal dilemmas are similar, but the proposed action occurs at a distance; i.e. would you put into motion actions that would harm one person to save others. Contrasting neural activation in the personal vs. impersonal dilemmas, these researchers find that prefrontal brain regions associated with theory of mind, as well as evolutionarily older subcortical regions related to emotions have greater activation, while areas of the brain associated with working

memory have reduced activations. These authors conclude that making decisions in personal moral dilemmas differentially draws on affective representations of outcomes.

In a related study, Moll et al., (2002) had seven subjects view pictures of emotionally charged scenes and asked them to rate their moral content. Comparing neural activity in moral violations (e.g. a man with a knife to the throat of a woman) versus those that were simply unpleasant (e.g. a mutilated body), Moll and colleagues find the strongest activation in cortical regions associated with emotional processing. Several additional contrasts (e.g. moral vs. neutral pictures) also show subcortical processing of moral scenes, especially in the amygdala. These neural signals often have a peripheral somatic basis as measured by skin conductance or heart rate (Scheman, 1996).

Sanfey et al. (2003) asked subjects in an MRI scanner to engage in a simple strategic interaction called the ultimatum game (UG). In the UG there is a proposer and a responder. The proposer is given sum of money, say \$10 and is instructed to offer a split of this to the responder. If the responder accepts the split, the money is paid. If the responder rejects the split, both parties earn nothing. Behaviorally, offers by the proposer less than 30% of the total are almost always rejected in industrialized societies (this regularity does not hold for “small scale” nomadic, agrarian, and pastoral societies, see Camerer, et al. 2004). The question Sanfey and colleagues asked was why people reject good money to punish another person for being stingy. They demonstrated that neural activity when a responder received a low offer of the split, contrasted with an equal or hyperfair offer, was greater in regions of the brain associated with emotions, especially a region known to activate with visceral disgust (the insular cortex). Subjects who received stingy offers appeared to be disgusted by them and were motivated to punish the transgressor even at a cost to themselves.

Similarly, de Quervain et al. (2004) asked subjects to play a game that admitted cooperation and defection, with the ability to directly punish those who defected (either at a cost or symbolically without cost). The design of the experiment, though, was that subjects knew they interacted only once, so that costly punishment would not benefit the punisher's earnings from the game (though punishment might be viewed as benefiting others). Costly punishment (as compared to symbolic punishment) produced strong activation in mid-brain limbic regions associated with rewarding behaviors. Subjects in this study punished because it felt good to do so.

Nonhuman social primates also act in ways that suggest that appropriate social behaviors have a physiologic basis. As discussed in Sarah Brosan's chapter in this volume, monkeys appear to have innate values of fairness and equity. They are willing to forgo food to punish another monkey who has violated these expected norms. Since monkeys have quite small prefrontal cortices relative to humans, this behavioral evidence suggests that these values are at least partially felt, not thought out. It is also worth noting that some monkeys understand symbolic exchange, i.e. the use of money. So the evolutionary basis for exchange is ancient, as are, apparently, the value of a fair division from exchange.

To summarize this section, a large number of researchers have demonstrated that the neural representation of moral values are automatic and difficult to suppress, and often utilize affective representations in the brain. I propose that values utilize similar neural mechanisms. In addition, values can be learned. That is, values may use potentiated pathways in the brain, biasing choices a certain way for a given environment. Long-term potentiation requires experience with this choice, while initial choice may

have a genetic basis. We can conclude from this that values have a neural representation, but need to be reinforced. Rick Shreve of Dartmouth College, who has contributed to the conferences that this volume reports, related an example to me of a violation of moral behavior and its instantiation. During a medical supply relief effort to Vilnius, Lithuania, that Rick participated in 1991, Latvian medical school faculty and students were asked to unload the supplies because of the fear of stealing. Shreve observed a Lithuanian professor take some toothbrushes from a partially open box. Later when he discussed this with colleagues, he was told that he had observed “Homo Sovieticus,” who, due years under the depersonalized and capricious rule of the Soviets, lacked the ability to make moral decisions. This suggests that the internal representation of values can be muted under extreme or chronic duress. Conversely, those of us with children typically spend a substantial amount of time inculcating what we see as proper values.

3. The Mechanics of Moral Economics

3.1 Modifying the Standard Economic Model

The argument given so far is that a subset of values that are universally held, and that there are consistent neural representations of such shared values. Some values, especially personal moral values, draw on affective responses to stimuli. The emotions produce coarse but quick reactions to stimuli. Emotions also impact impersonal decisions, but are more important when making decisions directly involving oneself.

In a standard economic model of constrained optimization, we can consider values to place two types of constraints on utility maximization and associated utility flows associated with the behavior. The first is a constraint on achievable outcomes. This would function like any other constraint (budget, time, etc.) that limits the goods one can consume and may use a terminal value. For example, suppose you are given a box of chocolates. You might decide not to consume all of these yourself and instead distribute them to friends or family because you value sharing. This limits your own consumption, and increases others’ consumption. It may also produce internal and external rewards (utility), for example, when a box of chocolates is given to a romantic partner on Valentine’s Day.

A second constraint is placed on economic maximization through instrumental values. Frey and colleagues (Frey et al., 2004; Frey & Stutzer, 2005) have recently introduced the notion of “procedural utility”—the notion that how something is done provides its own utility flow. A straightforward example is the UG that we discussed earlier. Most subjects prefer more money to less, but if offered a 10% split of a sum of money, will reject it out of the obvious inequality and stinginess of the proposer. (A typical subject response is “Who the hell does he think he is!”). The utility gain from receiving a small sum of money is less valuable than the process of accepting it. It is worth re-emphasizing that following one’s values is often costly, even though it “feels right.” In my experiments, non-reciprocators appear to have normal neural activity guiding them towards reciprocation, but appear to ignore or suppress these signals. Such subjects have aspects of antisocial personalities, but also make the most money in cooperation experiments (Zak, 2005).

The model I am proposing is similar to the mathematical model of prosocial emotions by Bowles & Gintis (2003). In this model, agents engage in a public goods-type game and are both self-interested and receive utility flows associated with pride,

guilt, empathy, shame and regret. In equilibrium, agents display a higher degree of cooperation than in the standard model absent social emotions. Prosocial emotions are the levers through which values are brought to decision-making. The Bowles and Gintis conception of emotions producing utility flows is similar to Frey's and his colleagues' conception of procedural utility. Overall, the Bowles and Gintis model captures nicely the notion that affect-laden values guide economic decisions.

3.2 Are Values Necessary in Economics? A Discussion of Generalized Trust

I would like to demonstrate the predictive importance of shared values by critiquing a well-known paper that I coauthored. Figure 1 shows survey data on the proportion of people in 42 countries who answered “yes” to the question “Are people in your country very trustworthy” or “Are people in your country mostly trustworthy.” The data vary from 3% answering yes in Brazil to 65% in Norway (!). Zak & Knack (2001) build a fairly standard mathematical model of purely self-interested economic principals who seek to make investments over time using agents as intermediaries, where agents have asymmetric information about market returns. Principals can monitor agents at a cost, and the principals' degree of trust is measured by their decision how much not to monitor. The formal model predicts that environments in which contract enforcement is high, social ties are tight, and people are similar (e.g. in income, language, ethnicity, etc.), and incomes are higher would have elevated levels of trust. Indeed, in empirical tests of the theory, these four factors explain 76% of the order-of-magnitude variation in the data. Great model, right?

The shortcoming of this model is that the formal and informal enforcement of investment contracts is specified only *on average*. Why an individual in the model does not cheat on a specific contract is not known. Nor does the model specify how agents that cheat are penalized, either formally or informally. Or if they feel something when they cheat or are cheated. These shortcuts were used primarily because they were analytically convenient. But the omission is not trivial: in this paper we also showed that generalized trust is among the strongest predictors that economists have ever found of whether a country would have increasing or decreasing living standards. Low trust countries are by-and-large low growth countries, while high trust countries see per capita incomes grow steadily. So, we better figure out what trust is and what supports it.

While the model of Zak & Knack (2001) does provide some important insights into the social and institutional foundations for economic growth, it says nothing about how quotidian transactions raise living standards. The inclusion of shared values can get us out of this inability to understand human behaviors that occur in high and low trust countries. A parsimonious microeconomic model of trust would include the instantiation of the value of trustworthiness in societies where generalized trust is high. This provides the novel prediction that high trust countries will have moderate, rather than high, degrees of formal monitoring of contracts. Because of the strong expectation, based on people's experience, that most people in Norway fulfill their contracts, there need not be a policeman (or lawyer) involved in each and every contract. People have a sense of what is expected and appropriate in transactions, and most of time they follow this. This may be why countries whose citizens are more similar have higher trust—the cognitive and affective mechanisms that induce the understanding of another's intentions may simply be easier to read when those around us are similar. Zak & Knack

(2001) do identify that countries with more dissimilarities, like the United States, need stronger formal enforcement of contracts. But the model does not explain why with only sporadic enforcement, most of the time most contracts are fulfilled—even in one-shot transactions. Only an appeal to some other mechanism can do this. I would like to suggest that values are one such device—stochastic punishment is insufficient since that would predict more cheating than is observed even in high trust countries. Further, not only do values do the job, they are, as I argued above, consistent with a large neuroscience literature that shows that values have a physiologic basis.

4. Applications of Moral Economics

Moral economics is not free of formal institutions to enforce contracts and mediate disputes. It does require that formal institutions reasonably impose impartial and fair outcomes when asked to do so or the institutions will be inconsistent with shared values and will not be followed. The seminal neurojurisprudence thinker Margaret Gruter repeatedly emphasized that laws were formalizations of social behaviors that typically had long histories (Gruter, 1991). The promulgation of laws that run counter to these informal norms—which are based on shared values—are doomed to be ignored. An example is the U.S.-style laws that were instituted in Eastern Europe and Russia in the 1990s and were largely ignored as the population had already developed informal systems to solve many economic problems that worked well enough. Elinor Ostrom and colleagues (Dietz, et al., 2003) have documented even more pernicious outcomes when norms of sharing a public resource, such as water in Africa, are changed by a national government or international organization, leading to a lack of cooperation, high waste, and conflict.

A likely physiologic mechanism at work when institutions undermine shared values, or when social-political environments are so chaotic that forward-looking decisions are difficult to make at all, is the inhibition of OT release during economic transactions and an associated reduction of interpersonal trust. As discussed in the previous section, a sufficient level of trust is necessary for investments that occur over time to be executed. Physiologic stress from living in an unstable environment inhibits OT release; people in this situation are in “survival mode” and are typically present- and self-oriented rather than drawing on shared values that support cooperative behaviors and capture gains to trade. There are spectacular exceptions of selflessness in the face of mortal danger, but the argument here is about economic transactions in a bodega in São Paulo, Brazil or Lagos, Nigeria. Preliminary evidence that environmental factors impact the physiology of trust at the country level are reported in Zak & Fakhar (in press).

This physiologic argument is consistent with individuals responding to environmental conditions in small scale societies in the study of Henrich et al. (2001; also see Camerer, 2004) that was discussed earlier. In these nomadic, agricultural, and pastoral societies, offers in the ultimatum game were all over the map, as were acceptances and rejections of these offers. This occurred even though the subjects interacting in the UG were not identifiable to each other. Variations across societies in the UG were associated with typical society-specific social behaviors. Behavior in these societies had presumably been optimized for the social environment, and behavior only responded imperfectly to the instructions given for the UG during the experiment. Perhaps most interestingly, Henrich and colleagues identify a single factor that explained the incidence of fair offers (those near 50% of the total) across societies: the exposure to

market exchange. Hold on—markets traders are all about getting the best deal, the most for me and the smallest for you. By now I hope that you’ve seen the error in this cartoon of market exchange. In the research of Henrich and colleagues, it appears that those in societies that traded understood that freely-entered into trade necessitates the acquisition of reasonable gains for *both* parties. Repeated trade is most likely to arise when parties work out a fair distribution of gains. This research suggests that markets reinforce the value of equal sharing of gains.

How does this occur? In the elegant experiment of Kimbrough, Smith and Wilson (this volume), traders were allowed to invent their own rules. At first, these were quite individually focused, with an individual maximizing his or her gain in a single transaction. But competition among traders quickly led to stable groups that maximized social surplus and worked out a fair division of it. Markets, and the institutions that underpin exchange, appear to support a value of near-equal sharing. Appropriate laws may do the same thing, moving people toward a change in behavior (and perhaps values) that then make the law itself unnecessary. Or, the behavior may lead the law. An example of the latter is smoking indoors. When I lived in San Diego, California, in the mid-1980s as an undergraduate, it was common for those in bars where young people congregated to go outside to smoke. This was before the state of California formally banned indoor smoking. Yet going outdoors to smoke was not common in many other states prior to laws prohibiting this behavior (and in San Diego this was uncommon in bars frequented by “older” people—those 30 and above). Anti-smoking laws are now being adopted throughout the U.S. and by European countries that have higher proportions of smokers than the U.S. The underlying value issue is the trade-off between the personal freedom to smoke indoors, with the imposition on others to breathe another’s smoke. The former value is individually focused while the latter is other-focused. Donald Eliot at Yale Law School pithily called this “the didactic value of law” at the third conference on values that was convened in this series.

The move from personal exchange to modern mostly impersonal exchange in markets is the key to the division of labor that caused the rapid gains in productivity and wealth since the industrial revolution (Smith, 2003). Because the instantiation of values both varies across individuals and across environments, there must be an enforcement of violations of values. Enforcement in traditional societies as personal—you cheat me, I hurt you or ostracize you. The incentives to cheat, free ride, and steal are rampant during impersonal exchange, necessitating an enforcement body that all agree to, that is, a government. As Lynn Stout discusses in her chapter in this volume, the mistake that legal scholars made, which she traces to Oliver Wendell Holmes, is the belief that humans are value-free, simply weighing costs and benefits of an action when making choices. This has produced laws that, like in the Israeli day-care center example, view punishment as a price one pays for a behavior rather than a violation of mutually shared values, and this may increase rather than reduce violations.

The main implication for economics is that laws that regulate market exchange, if well designed, should take into account innate predilections among most of the populace, to be maximally effective. In particular, it suggests that economies that are moderately regulated will both create the most wealth and have the most personal freedom. The former follows because all regulations have costs (direct costs and transactions costs), and increasing costs reduces exchange. When exchange is curtailed, the opportunities to create wealth are fewer. Having a police officer on every corner

would drastically reduce crime, but at a very high cost that is mostly unnecessary. Competition raises efficiency, and moderate physiologic stress increases cognitive skills. These things are good for the person and the society. But inordinate stress leads to physical decline and values violations that cause poverty. In addition, when every action is dictated from above, the freedom to innovate, which is necessary for technological innovation, is absent. Heavy-handed oversight of social behaviors (in markets and otherwise) also likely crowd-out the working of values. This has been shown experimentally by Fehr & Gächter (2002). Douglass North makes a similar argument when analyzing how institutions evolve (North, 2005).

If the thesis in this paper is correct, it means that there are not Western values and Eastern values, nor are Western and Eastern economic institutions. Rather, values across cultures are just a variation on a theme that is deeply human, strongly represented physiologically, and evolutionarily old. Similarly, the kinds of market institutions that create wealth and enable happiness and freedom of choice are those that resonant with the social nature of human beings who have an innate sense of shared values of right, wrong, and fair. Modern economies cannot operate without these.

References

Bhatt, M., Camerer, C. 2005. Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games and Economic Behavior* 52: 424-259.

Bowles, S. Gintis, H. 2003. Prosocial emotions. *Santa Fe Institute Working Paper #02-07-028*.

Buss, D. 2005. *The Murderer Next Door: Why the Mind is Designed to Kill*. New York: Penguin Books.

Camerer, C. et al. Eds. 2004. *Foundations of Human Sociality : Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.

Canli, T. et al., 2001. An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience* 115(1):33-42.

Canli, T. Amin, Z. 2002. Neuroimaging of emotion and personality: scientific evidence and ethical considerations. *Brain and Cognition*. 50(3):414-31.

Carter, C.S., et al., 2006. *Attachment and Bonding, A New Synthesis*. Cambridge: MIT Press.

CNN.com. 2006. Former Enron exec dies in apparent suicide. <http://archives.cnn.com/2002/US/01/25/enron.suicide/>.

Dabbs, J.M. Dabbs, M.G. 2000. *Heroes, Rogues, and Lovers: Testosterone and Behavior*. New York: McGraw-Hill.

Decety, J. Chaminade, T. 2003. Neural correlates of feeling sympathy. *Neuropsychologia* 41:127-138.

Diamond, J. 1997. *Guns, Germs and Steel: The Fates of Human Societies*. New York: W. W. Norton.

Fehr, E. and Gächter, S. 2002. Altruistic punishment in humans. *Nature* 415:137-140.

Frith, U. Frith, C.D. 2003. Developments and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B* 358(1431): 459-473.

Gneezy, U. Rustichini, A. 2000. Is a fine a price? *Journal of Legal Studies* XXIX: 1-17.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293: 2105-2108.

Grezes, J., Decety, J. 2001. Functional anatomy of execution, mental simulation, observation and verb generation of actions: A meta-analysis. *Human Brain Mapping*, 12: 1-19.

Gruter, M. 1991. *Law and the Mind: Biological Origins of Human Behavior*. Newbury Park, CA: Sage Publications.

Henrich, J. et al., 2001. In search of homo economicus: experiments in fifteen small-scale societies. *American Economic Review* 91(2): 73-79.

Kosfeld, M., Heinrichs, M. Zak, P.J., Fischbacher, U., Fehr, E. 2005. Oxytocin Increases Trust In Humans. *Nature* 435(2):673-676.

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. 2001 A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Nat. Acad. Sci. USA* 98(20): 11832-5.

McLean, B. and Elkin, P. 2003. *Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron*. Portfolio Hardcover.

North, D.C. 2005. *Understanding the Process of Economic Change*. Princeton: Princeton University Press.

O'Doherty, J. et al., 2001. Abstract reward and punishment representations in the human orbital frontal cortex. *Nature Neuroscience* 4:95-102.

Dietz, T., Ostrom, E., Stern, P.C. 2003. The struggle to govern the commons. *Science* 302:1907-1912.

Pedersen, C. 2004. How love evolved from sex and gave birth to intelligence and human nature. *Journal of Bioeconomics* 6:39-63.

Reader, S. N., Laland, K.N. 2002. Social intelligence, innovation, and enhanced brain size in primates. *Proc Natl Acad Sci U S A* 99(7):4436-41.

Rizzolatti, G., Fogassi, L., Gallese, V. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2: 661-670.

Rizzolatti, G. et al. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131-141.

Scheman, N. 1996. Feeling our way toward moral objectivity. Pp. 221-236 in Larry May, et al., Eds. *Mind and morals: Essays on cognitive science and ethics*. Cambridge: MIT Press.

Schwartz, S.H. 1994. Are there universal aspects in the structure and contents of human values? *Journal of Social Issues* 50(4):19-45.

Singer, T. et al. 2004a. Empathy for pain involves the affective but not sensory components of pain. *Science* 303(5661): 1157-1162.

Singer, T. et al. 2004b. Brain responses to the acquired moral status of faces. *Neuron* 41(4):653-662.

Slemrod, J. Ed. 1992. *Why people pay taxes: Tax compliance and enforcement*. Ann Arbor: University of Michigan Press

Smith, V. 2003. Constructivist and ecological rationality. *American Economic Review*, 93(3): 465-508.

Toussaint, L. Webb, J.R. 2005. Gender differences in the relationship between empathy and forgiveness. *Journal of Social Psychology* 145(6):673-85.

Waller, J. 2002. *Becoming Evil : How Ordinary People Commit Genocide and Mass Killing*. Oxford University Press.

Wicker, B. et al. 2003. Both of us disgusted in my insula: the common neural basis for seeing and feeling disgust. *Neuron* 40:655-664.

Zak, P.J., 2006. The neuroeconomics of trust. In *Two Minds. Intuition and Analysis in the History of Economic Thought*, Roger Franz, ed. Springer, in press.

Zak, P.J., Fakhar, A. In press. Neuroactive hormones and interpersonal trust: international evidence. *Economics and Human Biology*.

Zak, P.J., Park, J. Ween, J. Graham, S. 2006. An fMRI study of trust and exogenous oxytocin infusion. Society for Neuroscience 2006 abstract.

Zak, P.J., 2005. Trust: A temporary human attachment facilitated by oxytocin. *Behavioral and Brain Sciences* 28(3):368-369.

Zak, P.J., Kurzban, R., Matzner, W.T. 2005. Oxytocin is Associated with Human Trustworthiness. *Hormones and Behavior* 48: 522-527.

Zak, P.J., Kurzban, R., Matzner, W.T. 2004. The Neurobiology of Trust. *Annals of the New York Academy of Sciences*, 1032: 224-27.

Zak, P.J., Borja, K., Matzner, W.T., Kurzban, R. 2005. The Neuroeconomics of Distrust: Physiologic and Behavioral Differences Between Men and Women. *American Economic Review*, 95(2): 360-363.

Zak, P.J., 2004. Neuroeconomics. *Philosophical Transactions of the Royal Society B*, 359: 1737-1748.

Figures

Figure 1: Survey data on trust in 1994 from 42 countries with varying institutional environments.

