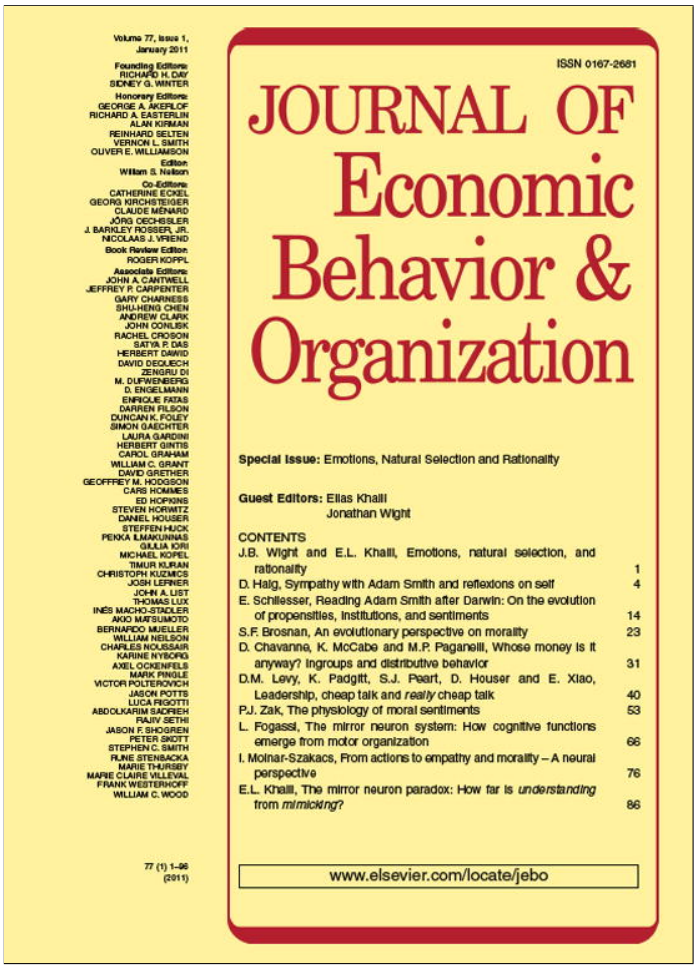

Volume 77, Issue 1,
January 2011

77 (1) 1–96
(2011)

ISSN 0167-2681

JOURNAL OF Economic Behavior & Organization

Special Issue: Emotions, Natural Selection and Rationality

Guest Editors: Elias Khalil
Jonathan Wight

CONTENTS

www.elsevier.com/locate/jebo

Journal of Economic Behavior & Organization 77 (2011) 53–65

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

The physiology of moral sentiments☆

Paul J. Zak[a,b,*]

[a] Center for Neuroeconomics Studies and Department of Economics, Claremont Graduate University, 160 E. 10th St., Claremont, CA 91711-6165, United States
[b] Department of Neurology, Loma Linda University Medical Center, United States

ARTICLE INFO

Article history:
Received 2 July 2009
Received in revised form 2 November 2009
Accepted 4 November 2009
Available online 18 September 2010

JEL classification:
D87
C7
B1
B4

Keywords:
Oxytocin
Empathy
Adam Smith
Neuroeconomics
Game theory
Generosity

ABSTRACT

Adam Smith made a persuasive case that "moral sentiments" are the foundation of ethical behaviors in his 1759 *The Theory of Moral Sentiments*. This view is still controversial as philosophers debate the extent of human morality. One type of moral behavior, assisting a stranger, has been shown by economists to be quite common in the laboratory and outside it. This paper presents the Empathy–Generosity–Punishment model that reveals the criticality of moral sentiments in producing prosocial behaviors. The model's predictions are tested causally in three neuroeconomics experiments that directly intervene in the human brain to "turn up" and "turn down" moral sentiments. This approach provides direct evidence on the brain mechanisms that produce prosociality using a brain circuit called HOME (human oxytocin-mediated empathy). By characterizing the HOME circuit, I identify situations in which moral sentiments will be engaged or disengaged. Using this information, applications to health and welfare policies, organizational and institutional design, economic development, and happiness are presented.

1. Introduction

One of the vexing issues in economics is the conflict between self-regarding and other-regarding behavior. We are indisputably interested in bettering our own conditions. At that same time, human beings show an enormous amount of care and concern for others, often at a cost to themselves. For example, $260 billion was given to U.S. charities in 2005, with $199 billion (77 percent) of this given by individuals (Giving USA, 2006). Over time, the proportion of after-tax income donated to charity has risen from 1.9 percent in 1954 to 2.2 percent in 2005 (Giving USA, 2006). Giving time to help others is another prosocial act. In 2005, 65 million Americans spent an average of 50 h volunteering to help others (White, 2005). Using the 2005 U.S. average wage of $18.48 h^{-1} (Social Security Administration, 2009), this constitutes an additional charitable donation valued at $60 billion. Direct and indirect transfers of $259 billion are an enormous amount; but at the same time it is still in single-digits as a proportion of personal income. A crucial issue in economics is how people navigate this self/other divide.

☆ The John Templeton Foundation provided generous support for this research. I thank Elias Khalil for the invitation to present my research and to contribute to this volume. I also thank George Sarraf for adept research assistance. My colleagues Yannis Venieris and Tom Borcherding provided critical feedback on early drafts that have substantially improved my original ideas. Nevertheless, any errors or omissions are solely my responsibility.

* Correspondence address: Center for Neuroeconomics Studies and Department of Economics, Claremont Graduate University, 160 E. 10th St., Claremont, CA 91711-6165, United States.
E-mail address: paul.zak@cgu.edu.

0167-2681/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2009.11.009

The apparent conflict between self and other can be seen in Adam Smith's two great books, *The Theory of Moral Sentiments* (TMS), first published in 1759, and *An Inquiry into the Nature and Causes of the Wealth of Nations* (WN) first published in 1776. The TMS caused a sensation when it was released, catapulting Smith from a minor professorship at the University of Glasgow to a well-known public intellectual. Indeed, Smith considered TMS so important that he continued to revise and expand it until his death, yet only recently have economists re-discovered its valuable insights (Smith, 1998).

An uncritical comparison of Smith's books leaves the impression that he is inconsistent in his view of human nature. In TMS, Smith argued that "sympathy" or "fellow-feeling" is the basis for moral behaviors. In modern parlance this would be called "empathy" or an emotional response to another's needs. (The word empathy was a 1858 coinage by German philosopher Rudolf Lotze (1817–1881) and therefore unavailable to Smith.) An important insight that Smith had in TMS was that "moral sentiments" are rapid, mostly uncontrolled, and strongly emotional rather than cognitive. Recent neuroscience research confirms Smith's assertion that emotional mirroring readily occurs in humans and has evolutionarily roots in other animals, especially nonhuman primates (De Waal, 2008; Casebeer, 2003; Zak and Barraza, 2009). Primatologist Frans de Waal argues that in social primates like humans, prosocial cooperation is the default behavior rather than a "thin veneer" over a "true" selfish nature (De Waal, 2006).

In WN, Smith promoted the view that individuals, behaving in their own self-interest in competitive markets would produce the greatest possible aggregate welfare Smith (1776/1904). How can Smith's two views of human nature be reconciled? First, self-interest and a desire to help others are not mutually exclusive. One may work hard to help an ailing parent or contribute to a charity. Second, self-interest and empathy may operate in different situations and environments. Someone may clip coupons to save on groceries, and also open up one's home to displaced disaster victims. Third, automatic empathic responses in the brain may or not affect decisions. We may feel empathy for a homeless person, but still choose not to give him money, perhaps out of a fear that we will "enable" him. We are both empathic and self-interested, and constantly adapt to varying social, economic, and institutional environments by changing the amount of self-regarding and other-regarding behavior.

Research in biology and psychology has identified factors in addition to empathy that may affect other-regarding behaviors in humans and nonhumans. For example, animals seeking social status may share resources with conspecifics to sustain coalitions. Similarly, sexual selection can drive males to expend resources to show that they can overcome handicaps and still acquire resources that can be shared with potential mates (Zahavi and Zahavi, 1997; Miller, 2000). A human manifestation of this behavior occurs when men spend lavishly to impress women. Yet, when high status males are established they tend to keep more resources for themselves (Eagly and Wood, 1999).

Another possible inducement for other-regarding behaviors is the threat of punishment. In one-shot anonymous settings, as I describe below, costly punishment of those who violate sharing norms often occurs even though the punisher bears a cost to punish but receives no monetary benefit (Fehr and Gächter, 2000; Camerer, 2003). This can be called "moralistic punishment" (Kurzban et al., 2007). If one anticipates that selfishness will be met with moralistic punishment, then individuals may be more prosocial under the threat of punishment than they would be otherwise. In the TMS, Smith wrote "That whatever appears to be the proper object of gratitude, appears to deserve reward; and that, in the same manner, whatever appears to be the proper object of resentment, appears to deserve punishment" (Smith, 1759/1982). Violations of moral duties, Smith argues, raise our moral hackles, provoking anger and a desire to punish.

This paper puts Smith's ideas in the TMS to the test by developing a neurologically informed model of moral sentiments and then uses three neuroeconomics experiments to test the model. Specifically, I test the roles of empathy, social status, and aggression in motivating other-regarding behaviors. The core innovation of this paper is this: if moral sentiments are real and measurable, they should also be manipulable. Herein I will describe how I have manipulated moral sentiments physiologically to understand if they are real and if so, how they work.

So why involve the brain in this endeavor? A variety of economics experiments have used behavioral manipulations to affect prosocial behavior. For example, Hoffman et al. (1994) framed a two person task to share a fixed sum of money known as the Ultimatum Game (described fully below) as a seller–buyer task and reduced offers from one person to another by nearly 10 percent. Low offers in the Ultimatum Game can also be produced by restricting information about the amount of money to be shared (see Camerer, 2003). Conversely, a 10 percent increase from baseline in offers in the Ultimatum Game was achieved by a perspective-taking exercise (Hoffman et al., 2000). Behavioral variations leave open the question of mechanism: *why* do these modifications affect prosocial tendencies? Economists have speculated why framing and information change prosociality, but this paper goes a step further and seeks to find – and manipulate – physiologic mechanisms in human beings that may affect moral sentiments. Indeed, this paper offers *causal evidence* for moral sentiments by administering drugs to experimental subjects to "turn up" or "turn down" candidate mechanisms in the human brain to see if other-regarding behaviors subsequently change.

2. The Empathy–Generosity–Punishment model

There are several reasons to examine the physiologic foundations for moral sentiments. First, many economic models have been proposed to explain prosocial behaviors (e.g. Rabin, 1993; Sally, 2000, 2001; Camerer, 2003), but all are based on the standard approach in economics of deducing a model and then testing it. The result is a large number of models that roughly equally well explain the data (Camerer, 2003, p. 207). Neuroscientific evidence for moral sentiments goes directly to the brain to reveal the mechanisms producing observed behaviors. Thus, brain-based models seek to get the assumptions

that undergird behavioral models correct by using direct evidence, rather than the typical "guess and verify" method used by economists (Vercoe and Zak, 2010).

Second, economists in the second half of the twentieth century through today have typically used a naive psychology of human choice. It was reasonable to start with rational expectations, but rather than viewing this as a benchmark, many in the profession accepted the assumptions of (hyper)-rationality unskeptically and began to believe that people (at least those in models) are always and everywhere rational, counter to a variety of evidence to the contrary. Elsewhere I have argued using a neuroscientifically informed model that people are neither rational nor irrational; rather, people are "rationally rational" (Zak, 2008b; McKenzie et al., 2009). The rational rationality model predicts that people will invest scarce cognitive resources in solving a decision problem only when the expected payoff is sufficiently large. Otherwise, human beings will expend the minimum resources needed to achieve a "good enough" outcome. "Good enough" means that there is a wide range of acceptable choices. Rational rationality is similar to Herbert Simon's notion of satisficing (Simon, 1991), but clearly identifies when people will satisfice and when they will not. Rational rationality occurs because cognitive resources are constrained and the brain evolved to conserve energy and deploy these resources only as needed. The neuroscience behind rational rationality requires that any economic model identify why individuals would expend scarce brain resources when making a decision rather than rely on previously learned heuristics. The model I propose here meets these objectives by developing a more sophisticated psychology of decision-making by including emotional influences while still optimizing using the preferences-beliefs-constraints framework (Gintis, 2006).

Third, standard economic models typically attribute prosociality to a single factor and as a result these models make non-nuanced predictions that do not accord well with observed decisions. Bowles and Gintis (2002) is an exception where the social emotions guilt and shame are posited to underlie prosocial behaviors. It is unlikely that a single brain mechanism is responsible for moral sentiments. The Empathy–Generosity–Punishment (EGP) model identifies the reasons for, and extent of, prosocial behaviors and shows that there are nonlinearities in decisions. As a result, the EGP model is tested by systematically varying physiologic factors to rank the strength of effects. By manipulating moral sentiments, I catalog both direct effects and cross effects to better understand why and when we are prosocial.

The fourth reason for drawing on neuroscience as the foundation for a model of human decision-making is that participants in many experiments poorly and inconsistently report the reasons for their choices (Smith, 1998). As a result, some other measurement modality besides self-report or revealed preference is needed to assess why people are making observed choices. As I discuss below, the neural substrates for moral sentiments that I have discovered are primarily emotional and largely outside of our conscious awareness. Participants in experiments are therefore expected to provide post hoc rationalizations for their decisions when queried. Such recitations have little value. My approach is direct, causal, and uniform across participants. This aids consistency and removes the reporting burden on participants.

Let me propose the following axiom upon which the analysis here is based.

Axiom. A person's physiologic state affects his or her decisions

We are familiar with this: do not shop for groceries when you are hungry. Set limits when you gamble in Las Vegas. Do not buy the first car you test drive. Knutson et al. recently showed that when the brains' "wanting" system is put into high gear by showing men pornographic pictures, they choose riskier financial portfolios than men shown neutral pictures (Knutson et al., 2008). Similarly, physiologic under-arousal is associated with an avoidance of decisions and variations in choices that suggest suboptimality (Engelmann et al., 2009). For example, cloudy weather in New York and other cities with stock markets has been shown to reduce daily stock market returns (Hirshleifer and Shumway, 2003). The stated axiom asserts that all decisions are physiologically determined.

2.1. Coming HOME

A substantial literature in animal behavior and human psychology has identified another's distress as motivating costly helping behaviors (e.g. Batson, 1991; Davis, 1996; Preston and de Waal, 2002; Sober and Wilson, 1998). Adam Smith also identified distress at motivating moral engagement to relieve our own aversive state, calling this the "healing consolation of sympathy" (Smith, 1759/1982, ii: para 236). Dovidio's arousal: cost-reward model of helping behavior (Dovidio, 1984; Dovidio et al., 1991) posits that in order for people to be motivated to help others, they have to first become aware of another's need for help. Distress elicited through emotional mirroring motivates one to attend to the needs of another. The motivation of distress then produces empathy for the distressed other and a desire to help. Empathy causing the relief of another's distress appears to reward those who help others with a warm glow-type utility flow (Andreoni, 1990).

A similar theory for prosocial behavior that distinguishes between distress and empathy, known as the empathy-altruism hypothesis, was proposed by Batson (1987, 1991), Batson et al. (1987), and Batson and Oleson (1991). Batson's studies have found that low to moderate amounts of distress increase empathy and prosocial engagement, while high degrees of distress are so aversive that they motivate in most a desire to escape and a reduction in helping behaviors. For example, if one observes someone hit his or her head painfully on a doorjamb, most people are motivated to help. Seeing someone with a massive head wound spewing blood can be so distressing that many people will avoid directly engaging to help.

Studies from my lab have found a similar nonlinearity in the physiology linking distress and empathy. In a study where participants watched a 100 second highly emotional video about a father and his 4-year-old son who has terminal brain cancer, subjects reported both distress and increased empathy (Barraza and Zak, 2009). Those watching the video had a 157

percent increase in the neurohormone oxytocin as well as a spike in the stress hormone cortisol. The change is oxytocin was associated with the subjective experience of empathy ($r = .20 > 0$, $p = .01$, $N = 145$) after controlling for the distress one felt. At the same time, distress reduced the change in oxytocin ($p = .05$), even while distress and empathy were highly correlated ($r = .81 > 0$, $p < .001$).

These findings directly link oxytocin to the subjective experience of empathy, and show that distress can inhibit empathy. Related brain imaging studies of the experience of one's own, or another's, pain have produced similar results. For example, responses to another's pain, and to facial expressions indicating pain, locate empathy within the observer's pain matrix: the anterior cingulate cortex, brainstem, cerebellum, and the anterior insula (Hein and Singer, 2008; Singer et al., 2004, 2006; Decety et al., 1997). This suggests that observing another's pain is averse to the observer. While these studies did not examine if observing another's pain was associated with prosocial engagement, they are consistent with a nonlinear effect of distress on empathy.

My lab began running a series of experiments in 2002 that have allowed us to characterize a brain circuit that we call HOME (Human Oxytocin Mediated Empathy) that appears to produce and sustain prosocial behaviors. Identifying this circuit permits sharp hypothesis testing of the physiology of moral sentiments, and therefore I will describe the HOME circuit here.

HOME activates with the release of the neuroactive hormone oxytocin (OT). OT is an evolutionarily ancient molecule that is an essential part of the mammalian attachment system motivating care for offspring. In socially monogamous mammals, OT and a closely related hormone, arginine vasopressin, facilitate attachment to mates to motivate biparental care (Carter, 1998). Maternal (and in some species paternal) care for offspring can be considered a template for more general other-regarding behaviors (De Waal, 2008; Sober and Wilson, 1998). In the human brain, large numbers of OT receptors are found in the amygdala, hypothalamus, and subgenual cortex (Barberis and Tribollet, 1996; Tribollet et al., 1992), brain regions associated with emotions and social behaviors. In addition to its role in the brain, OT is also released peripherally in body (Zak et al., 2005b). Peripheral OT binds to receptors in the heart and vagus nerve (which innervates the heart and gut), reducing heart rate and blood pressure thereby reducing anxiety (Porges, 2001). Thus, OT affects both brain and body, and informs the emotion regulation of fear/safety and approach/withdrawal behaviors.

Along with OT, HOME uses two neurotransmitters, dopamine (DA) and serotonin (SERT). The excitatory neurotransmitter dopamine is associated with goal-directed behaviors, drive, and reinforcement learning. DA motivates people and other animals to seek rewards such as food and sex, and it reinforces these critical behaviors by making them "feel good." Many drugs of abuse such as cocaine and methamphetamine flood the brain with DA, making it difficult to abstain from drug use because they hijack the brain's reinforcement learning system. The DA part of HOME makes prosocial behaviors rewarding, and provides a feedback loop to sustain them.

The activation of HOME also uses the inhibitory neurotransmitter SERT that is well-known for its effects on positive mood and for reducing anxiety. Prosocial behaviors connect us to others and in doing so, reduce our evolutionarily old fear of isolation. Having others around us is essential for human survival, now and during our evolutionary history. HOME makes being around other humans nonaversive and even pleasant.

Fig. 1 shows the three primary components of HOME, OT, DA, and SERT. This is how it works: a positive social stimulus causes OT release that potentiates the discharge of midbrain DA (Liu and Wang, 2003; Petersson et al., 1999). At the same time, OT release causes synaptic SERT to rise (Pfister and Muir, 1989), producing calmness and a positive effect on mood by binding to 5-HT1A (serotonin) receptors in the temporal cortex and prefrontal cortex (Sanabria-Bohórquez et al., 2002). More

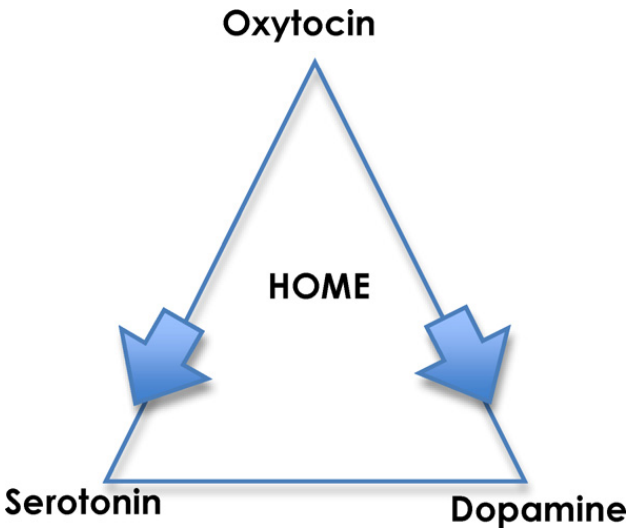

Fig. 1. HOME activates when OT is released following a positive social stimulus, motivating and reinforcing helping behavior by facilitating mid-brain dopamine release, and down-regulating distress by modulating serotonin release.

generally, the DA leg of HOME motivates people to action, and reinforces prosocial behaviors by making them rewarding while the SERT leg reduces the fear of social interaction and improves mood. Importantly, stress markers such as epinepherine and cortisol have a nonlinear effect on HOME: moderate stress increase OT while high levels of stress of fear inhibit OT release and thus the functioning of the HOME circuit.

OT release can be measured in blood, urine, and cerebral spinal fluid, and synthetic OT can be infused into human beings intravenously or intranasally to demonstrate its effect on behavior. Studies using OT infusion in humans have shown that it enhances the ability to infer others' emotions and intentions from facial expressions (Domes et al., 2007). OT also increases the time spent gazing toward the eye region of a face (Guastella et al., 2008), and the recognition of faces (Savaskan et al., 2008). Indeed, mice with the gene for the OT knocked out have social amnesia—they do not appear to remember animals they have previously encountered (Ferguson et al., 2000). Prosocial behaviors require that we understand others' intentions, something HOME allows us to do.

2.2. A model

Using these findings, I recently proposed a model of prosocial behaviors called the Empathy–Generosity model (Zak and Barraza, 2009). Before introducing the Empathy–Generosity–Punishment model that this paper is based upon, I will present the simpler Empathy–Generosity model. I have discovered that the Empathy–Generosity is related a model introduced in a footnote by the prominent Irish statistician and economist Francis Ysidro Edgeworth (1845–1926) in his book *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. In the Edgeworth model, as well as the Empathy–Generosity model, utility is obtained from one's own consumption and a weighted utility of another's consumption (Edgeworth, 1881, p. 53). Modern extensions of Edgeworth are many (Rabin, 1993; Fehr and Gächter, 2000; Andreoni, 1990; Sally, 2000, 2001; Levitt and List, 2007), and use other-regarding preferences as a motivation for prosocial behaviors. These models, though, are not neurologically based and presume motivations for human behavior without fully exploring whether these assumptions are warranted.

The Empathy–Generosity model considers a situation with two people, denoted person 1 and person 2, in which person 1 must decide how much of a fixed amount of resources M to offer to person 2. The resources may be pecuniary or nonpecuniary. Person 1's decision is modeled as

$$\begin{aligned} & \max_{b_1 b_2} E\{U(b_1) + \alpha\tau U(b_2)\} \\ \text{s.t.} \quad & b_1 + b_2 = M \end{aligned}$$

where $U(b_1)$ is the utility person 1 received from consuming benefits b_1, b_2 is the benefit that person 2 receives from person 1, $U(b_2)$ is the utility person 2 obtains from the transfer from person 1, and total resources, M, are finite. By assumption, $U(b)$ is increasing, continuous and strictly concave. Moral sentiments are captured by the parameter α. In his model, Edgeworth called the parameter α on the other's utility "effective sympathy" (1881, p. 53) and considered it a constant across all situations. In the Empathy–Generosity model, I will use Lotze's definition of emotional contagion and will call α "empathy."

The Empathy–Generosity identifies helping another as an individually costly behavior that accounts the motivation for helping by letting empathy depend on the situation the decision-maker faces. Specifically, let $\alpha(\tau)$: $[0,1] \rightarrow [0,1]$ be a continuous hyperbolic function where empathy, α, depends on the observed distress of person 2, τ. The function α has the following properties, $\alpha(0) \geq 0$, $\alpha(1) = 0$, and $\tau^* = \text{argmax}\ \alpha(\tau)$, with $\alpha(\tau^*) > \alpha(0)$. The parameter τ captures the distress that motivates the decision-maker to pay attention to the needs of the other person. The empathy function $\alpha(\tau)$ is hyperbolic because of the relationship between OT and stress hormones, with moderate distress increasing empathy but high degrees of distress causing one to disengage rather than help.

When $\alpha(\tau) = 0$ in the Empathy–Generosity model, person 1 is completely selfish, and when $\alpha(\tau) = 1$ s/he is benevolent, sharing benefits equally with person 2. It is straightforward to prove that when $\alpha < \alpha(\tau^*)$ rises, the benefits to person 2, b_2, increase. Indeed, there is a continuum of (non-unique) equilibria of Empathy–Generosity model parameterized by α. The summary prediction of the model is that empathy, driven by observed distress, motivates costly action to help another.

An open issue is whether person 2 wants help from person 1, or if person 1's help might be insufficient. For example, a small tip to a cabdriver might be returned as being insufficient in order to "call out" the cheapskate. In these cases, person 2 might decline the offer of benefits b_2 from person 1. The Empathy–Generosity model is modified here to include the threat of punishment for stinginess. Let us call this the Empathy–Generosity–Punishment (EGP) model.

The EGP model adds a decision by person 2 that allows him/her to reject or accept person's 1 offer of b_2; if person 1's offer is rejected then person 1 is assumed to lose resources. For concreteness, assume that if person 2 rejects person 1's offer, $b_1 = b_2 = 0$; that is, all resources are lost by both people, though any amount of resources lost upon rejection is sufficient for a well-defined model. Several recent papers in neuroeconomics have shown that individuals (especially men) receive pleasure from punishing those who are stingy (Singer et al., 2004; Delgado et al., 2005; De Quervain et al., 2004; Zak et al., 2005a). Using these findings, let $P(b_2)$ be the utility person 2 receives from punishing person 1 for making a stingy offer. Assume that $P(b_2)$ is decreasing, continuous, and strictly concave. In this setting, person 2 chooses the supremum of the set $\{U(b_2), P(b_2)\}$. The concavity restrictions generate a well-defined, unique solution to person 2's decision problem. Assume that if $U(b_2) = P(b_2)$, then person 2 chooses not to punish. Define the indifference value b_2^* such that $U(b_2) = P(b_2)$ at b_2^* and assume that b_2^* is common knowledge. Then, person 1's decision problem is to offer a split the M benefits between him/herself and

person 2 so that the offer is accepted. Person 1's decision problem now is

$$\begin{aligned} & \max_{b_1 b_2} U(b_1) + \alpha(\tau) U(b_2) \\ \text{s.t.} \quad & b_1 + b_2 = M \\ & b_2 \geq {b_2}^* \end{aligned}$$

As in the Empathy–Generosity model, the EGP model has a continuum of generically nonunique equilibria indexed by α.

The EGP model predicts that prosocial sharing of resources will occur when person 1: (i) witnesses distress by person 2; (ii) the distress raises person 1's empathy, but not too much $[0 < \alpha(\tau) < \alpha(\tau^*)]$; and (iii) the offer to help is large enough to avoid rejection by person 2 ($b_2 \geq b_2^*$). Let us define *generosity* as the amount the offer of b_2 by person 1 exceeds b_2^*.

3. Tests of the EGP model

The fundamental question this paper addresses is how the weight on empathy, α, in the EGP model might be changed through physiologic manipulations. Our experiments will use the ultimatum game (UG) to measure prosocial behaviors (Güth et al., 1982). In the UG, participants are matched in pairs and decision-maker 1 (DM1) offers a split of a fixed endowment M to decision-maker 2 (DM2). If DM2 accepts the offer, both DMs are paid and the game ends. If DM2 rejects the offer, both DMs receive nothing. The structure of the UG therefore permits us to test the EGP model directly.

The most compelling question on prosociality is why would someone ever be generous, that is, why offer more than one's forecast of DM2's minimum acceptable offer in the UG? Secondarily, we would like to know why DM2s would ever engage in costly punishment in a one-shot setting.

The neuroscience literature suggests three ways that the weight individuals put on other-regarding outcomes can be manipulated physiologically. These will provide a way to directly and causally test the EGP model.

3.1. Manipulation 1: oxytocin

The first, and perhaps most obvious factor, that might affect generosity in the UG is empathy as my modern interpretation of Adam Smith suggests. Psychologists such as Aron, Batson, and Cialdini and co-workers have manipulated empathy in a number of ways (Aron and Fraley, 1996; Batson, 1987; Cialdini et al., 1997). For example, participants in experiments are asked to take another's perspective and afterward their empathy and their desire to help are measured. In other studies, social distance is manipulated and again empathy and willingness to help are examined. Perhaps not surprisingly, perspective-taking and reduced social distance increased reported empathy and willingness to help others (though the motivations for this are in question; Maner et al., 2002).

The problem with behavioral manipulations is that they vary across subjects physiologically, and the brain mechanisms they draw upon are unknown. Further, the cited studies only measure a self-reported *desire* to help, rather than actual prosocial behaviors. As discussed above, my lab has recently associated empathy with the brain's release of OT. OT is therefore our target to manipulate empathy directly. Starting in 2004, my lab has shown by assaying OT in blood that when a human being trusts a stranger with his or her money, the recipient's brain releases OT. Receiving larger trust-signaling transfers was associated with higher OT levels, and higher OT was strongly correlated with greater reciprocity of money towards the initial trustee ($p = .01$) (Zak et al., 2004, 2005b). These studies were the first time that OT release was shown to affect non-reproductive behaviors in humans.

To prove that oxytocin causes trusting behaviors, we infused oxytocin into the human brain intranasally, and found we could more than double the number of subjects who trusted a stranger with all their money (Kosfeld et al., 2005). At the same time, OT did not affect objective risk-taking tasks or change cognition or mood. Instead, OT appears to subtly alter the balance between appropriate levels of trust and distrust of strangers, moving people towards great trust. We subsequently made the connection between OT release and the subjective experience of empathy in Barraza and Zak (2009).

In the present study, in order to consistently raise empathy across subjects, we infused 40IU of OT into subjects intranasally using a double-blind protocol and had them play the UG. Peptides closely related to OT have been shown to get into cerebral spinal fluid (Born et al., 2002) and affect brain activity (Kirsch et al., 2005). We therefore expected that participants given OT, relative to placebo, would be more generous in the UG. We did not expect a change in OT to affect the punishment threshold in the EGP model.

3.2. Manipulation 2: arginine vasopressin

Arginine vasopressin (AVP) is a hormone closely related to OT; they differ by only a single amino acid of the nine that make up each of these peptides. Like OT, AVP is also associated with peri-reproductive behaviors in mammals (as well as other functions such as the body's water balance). In mammals that co-parent, AVP causes males to form pair bonds to females and induces males to protect and care for offspring. Behaviorally, AVP rises when offspring are born and is causally associated with mate- and nest-guarding (Young et al., 1999; Young and Wang, 2004; Bester-Meredith et al., 2005). More generally, AVP is associated with "reactive aggression" or an exaggerated response to a provocation in co-parenting rodents and in socially monogamous monkeys (Donaldson and Young, 2008).

In human beings, higher AVP levels in cerebrospinal fluid are related to one's history of aggression, including unprovoked aggression (Coccaro et al., 1998; Kavoussi et al., 1997). Intranasal AVP administration in males causes aggressive facial expressions when seeing an unfamiliar male, but affiliative facial expressions when interacting with a female (Thompson et al., 2007). AVP is part of a brain circuit that facilitates appropriate social behaviors from agonistic to amicable (Donaldson and Young, 2008). We therefore hypothesized that AVP infusion would not affect generosity towards others, but would affect the threshold to punish other for being ungenerous.

3.3. Manipulation 3: testosterone

The EGP model predicts that one can reduce generosity by either increasing distress sufficiently ($\alpha > \alpha(\tau^*)$) or by blocking the action of OT and thereby reducing empathy. We took the latter approach in this third study. Although there are any number of possible ways that OT binding to its receptor can be inhibited, we used testosterone administration (T). T has been shown to inhibit OT binding (Insel et al., 1993).

Correlational studies of salivary T in humans have found that high T males, compared to males with lower T, are more likely to have had physical altercations, are more frequently divorced, spend less time with their children, seek out competition in many areas, have more sexual partners, but also face learning disabilities and more frequently lose their jobs (Raleigh et al., 1984; Mehta and Josephs, 2006). A recent study found that higher T males are more likely to reject low offers in the UG (Burnham, 2007). In general, these studies suggest that high T males are more aggressive and less prosocial (Harris et al., 1996).

These correlations from salivary T should be viewed with some caution as T is highly dependent on a variety of environmental factors (Dabbs and Dabbs, 2000). For example, if one wins a chess match T tends to be higher, and watching one's favorite team lose a soccer match on TV is associated with lower T (Mazur and Booth, 1998; Bernhardt et al., 1998). The high variability in basal T and the uncertainty about what participants did before coming into the lab call into question correlational studies of T and behavior (O'Carroll, 1998). Salivary testosterone assays, while convenient, also face measurement issues, and only a moderate correlation has been found between T measured in saliva and blood serum (Granger et al., 2004).

O'Carroll (1998) has critiqued correlational studies of T writing "Definitive evidence is likely to come from placebo-controlled, double-blind experiments in which circulating T levels are manipulated and appropriately reliable and sensitive assays of behaviour are taken". Directly manipulating T produces causal evidence relating T to behavior, and this is the approach we used. Studies that infused moderate doses of T into healthy males have found little effect on anger or mood (Yates et al., 1999; Pope et al., 2000). Mood effects occur only for supraphysiologic doses of T.

Based on this literature, we administered synthetic T to participants in a double blind crossover paradigm (for details, see Zak et al., 2009). Through its actions on the OT receptor, we expected that T administration would reduce generosity.

In addition, we had previously shown that distrust, measured as small monetary transfers to a stranger in a task known as the trust game, provoked a proportional rise in the bioactive metabolite of T, dihydrotestosterone (DHT) (Zak et al., 2005a). This was more prominent in men than in women and largely explained the lower rates of reciprocity in male DM2s in the trust game compared to female DM2s. Thus, besides reducing generosity, we also expected that raising T pharmacologically would cause men to engage in more individually costly punishment of those making ungenerous offers relative to themselves on placebo.

4. Materials and methods

All the studies reported here used only male participants because of the different effects of the hormones under study between men and women, including varying effects in women over the menstrual cycle. Random email announcements were used to recruit subjects for each study separately in a large public university in Southern California. Institutional Review Board approval from all participating institutions was obtained before running experiments. After obtaining written consent, participants were given a screening by a board-certified medical doctor for health issues that would exclude them from participating; for example, existing medical conditions or medication usage. All three hormones are FDA approved and have a long history of safe use. No adverse events occurred in any of the experiments.

Participants were assigned an identity-masking alpha-numeric code for their data, and were randomized to receive a hormone or placebo. Participants and the experimenters were blind as to which substance was given during the experiment. The OT and AVP infusion experiments administered substances intranasally, a route that has been showed to get approximately 10 percent of administered hormones into the brain after a 1 h loading period (Born et al., 2002). Based on these pharmacokinetics, we had participants wait 60 min after infusion before they made decisions. The OT dose used was 40IU, while a 20IU dose was used for AVP. These dosages were determined using the mid-range in Born et al. (2002), our previous work (Kosfeld et al., 2005), and related research Thompson et al. (2007). Both these experiments used a between-subjects design. The placebo was an identical volume of normal saline (0.91 percent sodium chloride).

The testosterone administration experiment used a 10 g dose of a topical hydrophilic drug called Androgel® (1 percent testosterone gel) prescribed for hypogonadal men (Swerdloff et al., 2000). There are androgen receptors throughout the cortex and subcortical regions in the human brain though which T administration affects behaviors (Chawnshang, 2002). Published pharmacokinetics show that T levels peak 18 h after Androgel® administration (Swerdloff et al., 2000). In order to confirm that T levels were actually raised (i.e. that subjects did not wash the gel off against our instructions), participants

Table 1
Impact of each physiologic manipulation on UG decisions relative to placebo. UG is ultimatum game; generosity is the UG offer minus the punishment threshold; DG is dictator game. All reported effects are statistically significant at $p \leq .05$.

Substance	UG offer	Punishment threshold	Generosity	DG offer
OT	+21%	0	+80%	0
T	−9%	+5%	−27%	0
AVP	0	0	0	0

in this experiment received a blood draw prior to Androgel® administration and again when they returned to the lab 16 h later to make decisions. Approximately 10 percent of Androgel® will enter the blood supply (Swerdloff et al., 2000). This experiment used a with-in subjects design because of the high variability in basal T levels (O'Carroll, 1998). The placebo was an alcohol-based gel of similar color, viscosity and odor.

Sample sizes were kept moderate for all three experiments to limit the number of participants exposed to our physiologic manipulations. The number of participants in the OT experiment who provided useable data was $N = 68$, with each subject making a decision in the UG as DM1 and as DM2. We also asked participants to decide how much to send in the dictator game (DG). The DG is a unilateral transfer by DM1 of an endowment to an unknown DM2; DM2 makes no decision in the DG. Roughly one-half of the participants in each experimental session received the drug and the other half was given the placebo. Additional details on the OT experiment can be found in Zak et al. (2007).

In the AVP experiment, a total of 57 participants provided usable data, with each subject making four decisions in the UG as both DM1 and DM2, and four decisions in the DG. Random rematching to others in the lab for each decision was used to mitigate the possible effects of reputation on decisions. This produced a total of 188 observations for each decision type. More information on the AVP experiment can be found in Stanton (2007).

The testosterone experiment included 25 participants who attended all parts of this crossover study. That is, each participant came to the lab four times; Time 1: 4 pm blood draw and substance administration; Time 2: 8 am next morning blood draw and decision tasks; Time 3: 4 pm blood draw and substance administration; Time 4: 8 am blood draw and decision tasks. The administration of T or placebo was counterbalanced across subjects and balanced within a session. Each participant had between 4 and 8 weeks between time points 1 and 3 as a "wash out" to reduce residual drug and learning effects. Participants made decisions in the UG as DM1 and DM2 and in the DG four times with random rematching to other participants. This study therefore has a total of 200 observations for each decision type. Additional information in this study can be found in Zak et al. (2009).

In all three experiments, participants received instructions and made decisions by computer using proprietary software. No pre-play intersubject communication was permitted and there was no deception of any type. During the drug loading period for all three experiments, participants completed a battery of survey questions that assessed their personality traits, mood, attitudes, and opinions. These included experiences in close relationships-revised (ECR-R) that measures attachment styles (Fraley et al., 2000), the Interpersonal Reactivity Index (IRI) that measures dispositional empathy (Davis, 1980, 1983), the Affective Intensity Measure (AIM) that assesses emotional stability (Larsen et al., 1986), an Anger Inventory (Brain Injury Resource Center, 1998), and the Personal Reaction Inventory (PRI) that measures social behaviors (Snyder, 1987).

5. Results

The effects of all three physiologic treatments on generosity, punishment and altruism are presented in summary form in Table 1. Details on these findings are presented below.

5.1. Generosity

5.1.1. Oxytocin

Participants infused with OT made mean offers in the UG that were 21 percent larger than those given the placebo (OT $4.86 (SD $1.06); placebo $4.03 (SD $1.29); two-tailed Mann–Whitney U test $p = 0.005$; $N = 68$). Indeed, only two subjects, both in the placebo group, offered the subgame perfect Nash equilibrium of $1 in the UG. The punishment threshold was unaffected by OT (OT: $3.03 (SD $1.69); placebo: $2.91 (SD $1.74); two-tailed Mann–Whitney, $p = 0.78$; group mean $2.97). Our primary variable of interest, generosity was 80 percent higher for those on OT compared to participants on placebo (OT: $1.89 (SD $1.06); placebo: $1.06 (SD $1.29); two-tailed Mann–Whitney U-test $p = 0.005$). Physiologically increasing empathy through OT infusion caused participants to be more generous.

5.1.2. Arginine vasopressin

The average UG offer was $5.14 (SD $1.68) for those on AVP group and $5.44 (SD $1.61) for those on placebo (two-tailed Mann–Whitney U test $p = 0.76$; $N = 232$). Counter to our hypothesis, the punishment threshold was unchanged by AVP infusion (AVP: $3.73 (SD $1.61); placebo: $3.54 (SD $1.75); two-tailed Mann–Whitney, $p = 0.54$). Mean generosity was 36 percent lower in the AVP group compared to those on placebo, but this did not reach statistical significance due to large standard errors (AVP: $1.40 (SD $2.80); placebo: $1.90 (SD $2.92); two-tailed Mann–Whitney U-test $p = 0.53$). AVP seems not have to increased reactive aggression leading to punishment in these participants.

5.1.3. Testosterone

We first tested whether T was higher after Androgel® administration. We assayed total, free, and dihydrotestosterone (DHT) to fully characterize the androgenic state of participants as total T is 98 percent protein bound and therefore mostly inactive biologically. Average T levels prior to Androgel® treatment were 4.16 pg/ml (total T), 14.37 ng/ml (free T), and 753.30 pg/ml (DHT). After the 16 h loading period and immediately prior to making decisions, total T was 60 percent higher, free T was 97 percent higher, and DHT was 128 percent higher for those receiving Androgel®; all changes were greater than zero for $p < 1E\text{-}6$ (paired t-tests, $N = 50$). Further, we checked if T levels were increased in every participant given Androgel® compared to himself at baseline. Androgel® did raise T in everyone who received it so no one had to be excluded from the analyses.

Turning to behavior, average proposals in the UG were 9 percent lower for men on T compared to themselves on placebo (T: \$4.63, placebo: \$5.08, one-tailed paired t-test, $N = 200$, $p = .001$). At the same time, the UG punishment threshold was 5 percent higher on T versus placebo (T: \$3.05, placebo: \$2.92, $p = .61$). Testing our primary hypothesis, mean generosity by men on T was 27 percent lower compared to themselves on placebo (T: \$1.57, placebo: \$2.15, $p = .035$).

Since we had blood samples for this study, we examined if generosity and punishment scaled with the level of a man's T. Testing all three measures of T (free, total, and DHT), we found a highly significant negative correlation between generosity and total T ($r = -0.25$, $p = 0.0004$), free T ($r = -0.1908$, $p = 0.0068$), and DHT ($r = -0.3063$, $p = 0.0001$). For example, participants in the lowest decile of DHT had average generosity 560 percent higher than that of men in the highest decile of DHT (\$3.65 vs. \$0.55). Significant correlations were also found for the punishment threshold and total T ($r = 0.1937$, $p = 0.0060$), free T ($r = 0.1529$, $p = 0.0306$), and DHT ($r = 0.2284$, $p = 0.0011$). The punishment threshold for participants in the highest decile of DHT was 86 percent higher than the punishment threshold participants in the lowest decile of DHT (\$4.00 vs. \$2.15). All the results maintain significance when the order of drug vs. placebo administration is controlled (we found that subjects became more generous when they returned the second time to the lab, but this did not drive our results). Thus, when T was increased, participants showed less generosity and an increased desire to punish those for being ungenerous.

5.2. Altruism

As a control, we also had participants make offers in the DG. Our expectation was that none of the hormones would affect DG offers as the DG is relatively self-focused ("how much do you want to send. . .") relative to the UG where one must explicitly think about what DM2 will do when making one's choice as shown in the EGP model. We included the DG to demonstrate that substance infusion did not alter participants' cognitive abilities that might explain their increased generosity (additional tests of this are reported in Section 5.3 below).

5.2.1. Oxytocin

We found that OT did not affect offers in the DG (OT: \$3.77 (SD \$2.21); placebo: \$3.58 (SD \$2.15); two-tailed Mann–Whitney test $p = 0.51$). As expected, transfers in the DG were less than in the UG. A least squares regression of generosity controlling for DG altruism and a OT/placebo indicator showed that OT infusion continued to be associated with generosity (OT coeff. 0.648, two-tailed t-test $p = .014$; DG coeff. = .376, two-tailed t-test $p = .0001$; $N = 68$, $R^2 = .43$).

5.2.2. Arginine vasopressin

This manipulation did not affect DG offers (AVP: \$3.54 (SD \$2.44); placebo: \$3.48 (SD \$2.34); two-tailed Mann–Whitney test $p = 0.90$). Regressing generosity on DG offers and an AVP binary variable showed that AVP continued to be unassociated with generosity (AVP coeff. −0.47, two-tailed t-test $p = .18$; DG coeff. = −0.46, two-tailed t-test $p = .0001$; $N = 232$, $R^2 = .16$).

5.2.3. Testosterone

Average offers in the DG were not different for those on Androgel® compared to placebo (T: \$3.34, placebo: \$3.56, two-tailed paired t-test, $p = 0.86$). No correlation between DG offers and the three measures of T was found. As above, we tested if differences in altruism affected generosity using a least-squares regression. The parametric relationship between T levels and reduced generosity continued to maintain significance when DG offers were included (total T: $\beta = -0.400$, $p = 0.001$; free T: $\beta = -0.057$, $p = 0.013$; DHT: $\beta = -0.001$, $p = 0.000$). Thus indicates that T inhibits generosity even when altruism is taken into account.

5.3. Surveys

For participants in all three experiments, there were no significant differences in anger, empathy, or attachment between those on placebo and those given OT, AVP or Androgel®. For those in the testosterone experiment, these comparisons were made between subjects on placebo compared to themselves on Androgel®. Across conditions subjects were well-matched for age and other demographic factors. The surveys we included measured personality traits and we showed statistically that traits had no effect of on the behaviors in the lab. This indicates that participants' generosity was caused by the change in their physiologic states.

6. Discussion and conclusions

The primary findings presented here are that: (i) 40IU of OT infusion increases generosity by 80 percent compared to placebo but does not affect punishment; (ii) 10 g of T administration reduces generosity 27 percent and raises the punishment threshold as it increases; and (iii) 20IU of AVP infusion has no effect on generosity or punishment.

These findings support the EGP model's predictions that empathy drives prosocial behaviors. A more succinct summary of my findings is: Adam Smith was right. Fellow-feeling or empathy appear to motivate us toward virtue and away from vice. My neuroeconomics studies are just peeling away the physiology that underlies Smith.

What is novel here is that the proposed mechanism linking empathy to virtuous behaviors was physiologic and manipulable. Rather than some region of the brain or psychological state being correlated with empathy, I have sought to narrow the mechanisms of empathy to a single, evolutionarily ancient molecule, oxytocin. The key finding was research from my lab linking the release of OT to the subjective experience of empathy (Barraza and Zak, 2009). My work had already shown that oxytocin was associated with prosocial behaviors such as trust and trustworthiness (Zak et al., 2004, 2005b; Kosfeld et al., 2005), and our recent research has shown that OT affects other virtues such as generosity and sacrifice (Zak et al., 2007; Morhenn et al., 2008). In Barraza and Zak (2009), we also showed that those who were more empathically engaged by the emotional video were also more generous towards a stranger in the UG ($r = 0.24 > 0$, $p = 0.05$). In addition, participants in this experiment were given a choice to donate some of their earnings to charity at the end of this painful experiment (every participant had two blood draws), 32 percent of them did so (mean = \$6.09, SD = \$6.31). Those who donated to charity were those who were also the most empathically engaged and the most generous in the UG ($r = 0.356$, $p = 0.004$).

Another way to understand these results is that human beings have a physiologic moral compass. This itself is not surprising–we are social creatures and our brains evolved to modulate our behaviors to maintain our place in the social milieu. What is surprising is that this tiny and ancient molecule, OT, and the neurotransmitters it modulates in the HOME circuit appear to be responsible for moral sentiments. If the burgeoning research from my lab and others studying OT continue to accumulate findings supporting the physiology of moral sentiments, then a number of important implications can be drawn.

6.1. Life history matters

The development of the HOME system depends critically on the nurturing received in childhood. When parental nurturing is scarce or absent in nonhuman mammals, areas of the brain that typically have high densities of OT receptors are instead sparse and these animals are socially withdrawn throughout their lives (Carter, 2003; Donaldson and Young, 2008). Among the nearly one thousand college students in whom I have measured OT release in blood, about 2 percent have a dysregulation in oxytocin. I call this oxytocin deficit disorder (ODD) (Zak, 2005), as these people have high basal OT levels but do not appear to release OT after receiving a trust stimulus. Accordingly, they do not reciprocate trust. Tellingly, those with ODD have difficulty forming friendships and romantic attachments and are often deceptive. We have recently replicated this finding in patients diagnosed with social anxiety disorder (Hoge et al., 2008). The causes of ODD are likely to include one's developmental history, and we are currently running several studies to characterize the role of life history in the functioning of the HOME system. We expect to find, as in rodents, that one's childhood experiences will affect how HOME operates.

6.2. Institutions matter

Survey responses of perceived trustworthiness by those in one's country vary from 2 percent in Brazil to 65 percent in Norway. Seventy-six percent of this order-of-magnitude variation is explained by the functioning of formal and informal institutions and the level of income (Zak and Knack, 2001). When institutions function well, uncertainty is reduced and planning and predictability simplify decisions. In addition, trust reduces transactions costs, and we have shown that trust is a powerful force to impel economic growth through this route (Zak and Knack, 2001; Knack and Zak, 2002). Low trust, driven by social, political, and economic instability, obstructs growth. Physiologically, predictability and interpersonal trust reduce stress. High levels of stress have been shown to inhibit OT release (Bremner et al., 1995). This causal chain reveals that good institutions facilitate moral sentiments (Zak, 2008a). Indeed, increases in per capita income are associated with greater trust providing a potential reinforcement mechanism between institutions, trust, and moral sentiments.

6.3. Organizations matter

Moral sentiments can be promoted or inhibited by the organizational environment in which one finds oneself. Jeffrey Skilling's Enron was famously a highly competitive organization with a "rank and yank" system that fired the bottom 20 percent of employees each quarter. The stress of making the numbers led many to make up the numbers to survive. Contrast this to Google's motto of "do no evil" and their policy of giving employees paid time every week to work on their own projects (see Zak, 2008a). Google's approach develops employees' sense of mission, accomplishment, and personal fulfillment, something long advocated by my late colleague Drucker (1983). Moral sentiments are essential for any non-coerced collective action since individuals must not only coordinate, but must choose not to free-ride (Zak and Barraza, 2009). We have even found that the food one consumes and the amount of pollution in the environment affects trust as well, likely impacting moral sentiments through the HOME system (Zak and Fakhar, 2006).

6.4. Happiness happens

In the Nicomachean Ethics, Aristotle argued that eudiamonia or happiness comes from leading a virtuous life. This can be scaled up to an entire society: virtuous societies will be happier. In the TMS, Smith wrote "The orderly and flourishing state of society is agreeable to him, and he takes delight in contemplating it. Its disorder and confusion, on the contrary, is the object of his aversion, and he is chagrined at whatever tends to produce it. He is sensible too that his own interest is connected with the prosperity of society, and that the happiness, perhaps the preservation of his existence, depends upon its preservation" (Vol. 1, book iiii, para. 434). I have similarly found that the virtue of trustworthiness is highly correlated with self-reported rates of happiness at the country level ($r = .60 > 0$, $p = .001$, $N = 27$; Zak and Fakhar, 2006). If our goal as behavioral scientists is to maximize human happiness, as thinkers from Bentham (1789/1988) to Haidt (2006) have argued, then designing institutions to foster trust and happiness is an important goal. Raising trustworthiness by improving formal institutions, reducing poverty, raising educational standards, strengthening social ties, cleaning up the environment, and encouraging the consumption of healthy foods (Knack and Zak, 2002; Zak and Fakhar, 2006) can directly elevate happiness. Further, recent findings have called into question the easterlin paradox that claimed to find a ceiling on happiness as income rises (Stevenson and Wolfers, 2008) suggesting that income and happiness may move together after all. Policies that increase the expression of moral sentiments such as trustworthiness are thus of vital importance in developing and developed countries both to increase prosperity as well as happiness.

Moral sentiments are real, measurable and significant. Knowing the physiology of moral sentiments is the key to understanding when we will be selfish and when we will be selfless. Social ties reinforce moral sentiments and are the basis for long-term happiness. In the TMS, Adam Smith understood this, writing "the chief part of human happiness arises from the consciousness of being beloved" (Smith, 1759/1982, p. 41). Indeed, my research identifying the physiology of moral sentiments comes to a similar conclusion: oxytocin is not only the moral molecule, but is also a primary element of the physiology of love (Carter, 1992). It took neuroeconomics studies to bring science to Adam Smith's moral sentiments, putting his intuitions to the test, and finding him on target again.

References

Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. The Economic Journal 100, 464–477.

Aron, A., Fraley, B., 1996. Empathy as including other in the self. In: Paper Presented at the Annual Meeting of the Society of Experimental Social Psychology, Sturbridge, MA.

Barberis, C., Tribollet, E., 1996. Vasopressin and oxytocin receptors in the central nervous system. Critical Reviews in Neurobiology 10, 119–154.

Barraza, J., Zak, P.J., 2009. Empathy toward strangers triggers oxytocin release and subsequent generosity. Annals of the New York Academy of Sciences 1167, 182–189.

Batson, C.D., 1987. Prosocial motivation: is it ever truly altruistic? Advances in Experimental Social Psychology 20, 65–122.

Batson, C.D., 1991. The Altruism Question: Toward a Social-Psychological Answer. Lawrence Erlbaum, Hillsdale.

Batson, C.D., Fultz, J., Schoenrade, P.A., 1987. Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. Journal of Personality 55, 19–39.

Batson, C.D., Oleson, K.C., 1991. Current status of the empathy-altruism hypothesis. In: Clark, M.S. (Ed.), Review of Personality and Social Psychology, vol. 12: Prosocial behavior. Sage, Newbury Park, CA, pp. 62–85.

Bentham, J., 1789/1988. An Introduction to the Principles of Morals and Legislation. Prometheus Books, Amherst, New York.

Bernhardt, P.C., Dabbs Jr., J.M., Fielden, J.A., Lutter, C.D., 1998. Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. Physiology and Behavior 65, 59–62.

Bester-Meredith, J.K., Martin, P.A., Marler, C.A., 2005. Manipulations of vasopressin alter aggression differently across testing conditions in monogamous and non-monogamous peromyscus mice. Aggressive Behavior 31, 189–199.

Born, J., Lange, T., Kern, W., McGregor, G.P., Bickel, U., Fehm, H.L., 2002. Sniffing neuropeptides: a transnasal approach to the human brain. Nature Neuroscience 5, 514–516.

Bowles, S. Gintis, H., 2002. Prosocial emotions. Santa Fe Institute Working Paper #02-07-028.

Brain Injury Resource Center, 1998. Anger inventory. Retrieved from http://www.headyinjury.com/anger.htm.

Bremner, J.D., Krystal, J.H., Southwick, S.M., Charney, D.S., 1995. Functional neuroanatomical correlates of the effects of stress on memory. Journal of Trauma and Stress 8, 527–554.

Burnham, T.C., 2007. High-testosterone men reject low ultimatum game offers. Proceedings of the Royal Society B 274, 2327–2330.

Camerer, C.F., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press, Princeton, NJ.

Carter, C.S., 1998. Neuroendocrine perspectives on social attachment and love. Psychoneuroendocrinology 23, 779–818.

Carter, C.S., 2003. Developmental consequences of oxytocin. Physiology and Behavior 79, 383–397.

Casebeer, W.D., 2003. Moral cognition and its neural constituents. Nature Reviews Neuroscience 4, 840–846.

Chawnshang, C., 2002. Androgens and Androgen Receptor: Mechanisms, Functions, and Clinical Applications. Springer, New York.

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C., Neuberg, S.L., 1997. Reinterpreting the empathy–altruism relationship: when one into one equals oneness. Journal of Personality and Social Psychology 73, 481–494.

Coccaro, E.F., Kavoussi, R.J., Hauger, R.L., Cooper, T.B., Ferris, C.F., 1998. Cerebrospinal fluid vasopressin levels: correlates with aggression and serotonin function in personality-disordered subjects. Archives of General Psychiatry 55, 708–714.

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

Davis, M.H., 1980. A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology 10, 85.

Davis, M.H., 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of Personality and Social Psychology 44, 113–126.

Davis, M.H., 1996. Empathy: A Social Psychological Approach. Westview Press, Boulder, CO.

De Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., Fehr, E., 2004. The neural basis of altruistic punishment. Science 305, 1254–1258.

De Waal, F.B.M., 2006. Primates and Philosophers: How Morality Evolved. Princeton University Press, Princeton.

De Waal, F.B.M., 2008. Putting the altruism back into altruism: the evolution of empathy. Annual Review of Psychology 59, 279–300.

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., Fazio, F., 1997. Brain activity during observation of actions influence of action content and subject's strategy. Brain 120, 1763–1777.

Delgado, M.R., Frank, R.H., Phelps, E.A., 2005. Perceptions of moral character modulate the neural systems of reward during the trust game. Nature Neuroscience 8, 1611–1618.
Domes, G., Heinrichs, M., Michel, A., Berger, C., Herpertz, S.C., 2007. Oxytocin improves 'mind-reading' in humans. Biological Psychiatry 61, 731–733.
Donaldson, Z.R., Young, L.J., 2008. Oxytocin, vasopressin, and the neurogenetics of sociality. Science 322, 900–904.
Dovidio, J.F., 1984. Helping behavior and altruism: an empirical and conceptual overview. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, vol. 17. Academic Press, New York, pp. 361–427.
Dovidio, J.F., Piliavin, J.A., Gaertner, S.L., Schroeder, D.A., Clark III, R.D., 1991. The arousal: cost-reward model and the process of intervention. In: Clark, M.S. (Ed.), Review of Personality Social Psychology, vol. XII: Prosocial behavior. Sage, Newbury Park, CA, pp. 86–118.
Drucker, P.F., 1983. Concept of the Corporation. New American Library, New York.
Eagly, A.H., Wood, W., 1999. The origins of sex differences in human behavior: evolved dispositions versus social roles. American Psychologist 54, 408–423.
Engelmann, J.B., Capra, M.S., Noussair, C., Berns, G.S., 2009. Expert financial advice neurobiologically "offloads" financial decision-making under risk. PLoS ONE 4, e4957, 10.137/journal.pone.0004957.
Edgeworth, F.Y., 1881/1967. Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences. Augustus M. Kelley, New York.
Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90, 980–994.
Ferguson, J.N., Young, L.J., Hearn, E.f., Matzuk, M.M., Insel, T.R., Winslow, J.T., 2000. Social amnesia in mice lacking the oxytocin gene. Nature Genetics 25, 284–298.
Fraley, R.C., Waller, N.G., Brennan, K.A., 2000. An item response theory analysis of self-report measures of adult attachment. Journal of Personality & Social Psychology 78, 350–365.
Gintis, H., 2006. A Framework for the integration of the behavioral sciences. Behavioral and Brain Sciences 30, 1–61.
Giving USA, 2006. The Annual Report on Philanthropy for the Year 2006. Giving Institute, Glenview, IL.
Granger, D.A., Shirtcliff, E.A., Booth, A., Kivlighan, K.T., Schwartz, E.B., 2004. The "trouble" with salivary testosterone. Journal of Psychoneuroendocrinology 29, 1229–1240.
Guastella, A.J., Mitchell, P.B., Dadds, M.R., 2008. Oxytocin increases gaze to the eye region of human faces. Biological Psychiatry 63, 3–5.
Güth, W., Schmittberger, Schwarze, 1982. An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization 3, 367–388.
Haidt, J., 2006. The Happiness Hypothesis. Basic Books, New York.
Harris, J.H., Rushton, H.J., Hampson, E., Jackson, D.N., 1996. Salivary testosterone and self-report aggressive and pro-social personal characteristics in men and women. Aggressive Behavior 22, 321–331.
Hein, G., Singer, T., 2008. I feel how you feel but not always: the empathic brain and its modulation. Current Opinion in Neurobiology 18, 153–158.
Hirshleifer, D., Shumway, T., 2003. Good day sunshine: stock returns and the weather. Journal of Finance 58, 1009–1032.
Hoffman, E., McCabe, K.A., Shachat, K., Smith, V.L., 1994. Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior 7, 346–380.
Hoffman, E., McCabe, K.A., Smith, V.L., 2000. The impact of exchange context on the activation of equity in ultimatum games. Experimental Economics 3, 5–9.
Hoge, E.A., Pollack, M.H., Kaufman, R.E., Zak, P.J., Simon, N.M., 2008. Oxytocin levels in social anxiety disorder. CNS Neuroscience &Therapeutics 14, 165–170.
Insel, T.R., Young, L., Witt, D.M., Crews, D., 1993. Gonadal-steroids have paradoxical effects on brain oxytocin receptors. Journal of Neuroendocrinology 5, 619–628.
Kavoussi, R., Armstead, P., Coccaro, E., 1997. The neurobiology of impulsive aggression. Psychiatric Clinics of North America 20, 395–403.
Kirsch, P., Esslinger, C., Chen, Q., Mier, D., Lis, S., Siddhanti, S., et al., 2005. Oxytocin modulates neural circuitry for social cognition and fear in humans. Journal of Neuroscience 25, 11489–11493.
Knack, S., Zak, P.J., 2002. Building trust: public policy, interpersonal trust, and economic development. Supreme Court Economic Review 10, 91–107.
Knutson, B., Wimmer, G.E., Kuhnen, C.M., Winkielman, P., 2008. Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. NeuroReport 19, 509–513.
Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. Nature 435, 673–676.
Kurzban, R., DeScioli, P., O'Brien, E., 2007. Audience effects on moralistic punishment. Evolution and Human Behavior 28, 75–84.
Larsen, R.J., Diener, E., Emmons, R.A., 1986. Affect intensity and reactions to daily life events. Journal of Personality and Social Psychology 51, 803–814.
Levitt, S.D., List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? Journal of Economic Perspectives 21, 153–174.
Liu, Y., Wang, Z.X., 2003. Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. Neuroscience 121, 537–544.
Maner, J., Luce, C.L., Neuberg, S.L., Cialdini, R.B., Brown, S., Sagarin, B.J., 2002. The effects of perspective taking on motivations for helping: still no evidence for altruism. Personality and Social Psychology Bulletin 28, 1601–1610.
Mazur, A., Booth, A., 1998. Testosterone and dominance in men. Behavioral and Brain Science 21, 386.
McKenzie, R., Turner, J., Zak, P.J., 2009. The neuroeconomics of rational decision making. In: McKenzie, R. (Ed.), Predictably Rational? In Search of Defenses for Rational Behavior in Economics. Springer, pp. 175–202.
Mehta, P.H., Josephs, R.H., 2006. Testosterone change after losing predicts the decision to compete again. Hormones and Behavior 50, 684–689.
Miller, G., 2000. The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature. Doubleday, New York.
Morhenn, V.B., Park, J.W., Piper, E., Zak, P.J., 2008. Monetary sacrifice among strangers is mediated by endogenous oxytocin release after physical contact. Evolution and Human Behavior 29, 375–383.
O'Carroll, R.E., 1998. Placebo-controlled manipulations of testosterone levels and dominance. Behavioral and Brain Science 21, 382–383.
Petersson, M., Hulting, A.L., Andersson, R., Uvnäs-Moberg, K., 1999. Long-term changes in gastrin, cholecystokinin and insulin in response to oxytocin treatment. Neuroendocrinology 69, 202–208.
Pfister, H.P., Muir, J.L., 1989. Influence of exogenously administered oxytocin on central noradrenaline, dopamine and serotonin levels following psychological stress in nulliparous female rats (*Rattus norvegicus*). International Journal of Neuroscience 45, 221–229.
Pope, H.G., Kouri, E.M., Hudson, J.I., 2000. Effects of supraphysiologic doses of testosterone on mood and aggression in normal men: a randomized controlled trial. Archives of General Psychiatry 57, 133–140.
Porges, S.W., 2001. The polyvagal theory: phylogenetic substrates of a social nervous system. International Journal of Psychophysiology 42, 123–146.
Preston, S.D., de Waal, F.B.M., 2002. The communication of emotions and the possibility of empathy in animals. In: Post, S., Underwood, L.G., Schloss, J.P., Hurburt, W.B. (Eds.), Altruistic Love: Science Philosophy Religion in Dialogue. Oxford University Press, New York, pp. 284–308.
Rabin, M., 1993. Incorporating fairness into game theory and economics. The American Economic Review 83, 1281–1302.
Raleigh, M.J., McGuire, M.T., Brammer, G.L., Yuwiler, A., 1984. Social and environmental influences on blood serotonin concentrations in monkeys. Archives of General Psychiatry 41, 405–410.
Sally, D., 2000. A general theory of sympathy, mind-reading, social interaction, with an application to the prisoner's dilemma. Social Science Information 39, 567–634.
Sally, D., 2001. On sympathy and games. Journal of Economic Behavior and Organization 44, 1–30.
Sanabria-Bohórquez, S.M., Biver, F., Damhaut, P., Wikler, D., Veraart, C., Goldman, S., 2002. Quantification of 5-HT(1A) receptors in human brain using p-MPPF kinetic modeling and PET. European Journal Nuclear Medicine and Molecular Imaging 29, 76–81.
Savaskan, E., Ehrhardt, R., Schulz, A., Walter, M., Schächinger, H., 2008. Post-learning intranasal oxytocin modulates human memory for facial identity. Psychoneuroendocrinology 33, 368–374.

Simon, H., 1991. Organizations and markets. Journal of Economic Perspectives 5, 25–44.
Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. Science 303, 1157–1162.
Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J., Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. Nature 439, 466–469.
Smith, A., 1776[1904]. In: Cannan, E. (Ed.), An Inquiry into the Nature and Causes of the Wealth of Nations, 2 vol. Methuen, London, http://oll.libertyfund.org/title/171.
Smith, A., 1759/1982. The theory of moral sentiments. In: Raphael, D.D., Macfie, A.L. (Eds.), vol. I of the Glasgow Edition of the Works and Correspondence of Adam Smith. Liberty Fund, Indianapolis.
Smith, V.L., 1998. The two faces of Adam Smith. Southern Economic Journal 65, 1–19.
Snyder, M., 1987. Public Appearances, Private Realities. Freeman and Company, New York.
Sober, E., Wilson, D.S., 1998. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge.
Social Security Administration, 2009. National Average Wage Index, Retrieved from http://www.ssa.gov/OACT/COLA/AWI.html.
Stanton, A.A., 2007. Neural substrates of decision-making in economic games. Journal of Dissertations 1, 1–59.
Stevenson, B., Wolfers, J., 2008. Economic growth and subjective well-being: reassessing the easterlin paradox. NBER Working Papers 14282, National Bureau of Economic Research.
Swerdloff, R.S., Wang, C., Cunningham, G., Dobs, A., Iranmanesh, A., Matsumoto, A.M., Snyder, P.J., Weber, T., Longstreth, J., Berman, N., the Testosterone Gel Study Group, 2000. Long-term pharmacokinetics of transdermal testosterone gel in hypogonadal men. Journal of Clinical Endocrinology & Metabolism 85, 4500–4510.
Thompson, M.R., Callaghan, P.D., Hunt, G.E., Cornish, J.L., McGregor, I.S., 2007. A role for oxytocin and 5-HT(1A) receptors in the prosocial effects of 3,4 methylenedioxymethamphetamine ('ecstasy'). Neuroscience 146, 509–514.
Tribollet, E., Dubois-Daupin, M., Dreifuss, J.J., Barberis, J.S., 1992. Oxytocin receptors in the central nervous system: distribution, development and species differences. In: Pedersen, C.A., Caldwell, J.D., Jirikowski, G.F., Insel, T.R. (Eds.), Oxytocin in Maternal, Sexual, and Social Behaviors. New York Academy of Sciences, New York, pp. 220–232.
Vercoe, M., Zak, P.J., 2010. Inductive modeling using causal studies in neuroeconomics: brains on drugs. Journal of Economic Methodology 17, 123–137.
White, S.B., 2005. Volunteering in the United States. United States Department of Labor, http://www.bls.gov/opub/mlr/2006/02/ressum.pdf.
Yates, W.R., Perry, P.J., MacIndoe, J., Holman, T., Ellingrod, V., 1999. Psychosexual effects of three doses of testosterone cycling in normal men. Biological Psychiatry 45, 254–260.
Young, L.J., Wang, Z., 2004. The neurobiology of pair bonding. Nature Neuroscience 7, 1048–1054.
Young, L.J., Nilsen, R., Waymire, K.G., MacGregor, G.R., Insel, T.R., 1999. Increased affiliative response to vasopressin in mice expressing the V_{1a} receptor from a monogamous vole. Nature 400, 766–768.
Zahavi, A., Zahavi, A., 1997. The Handicap Principle. Oxford University Press, New York.
Zak, P.J. (Ed.), 2008a. Moral Markets: The Critical Role of Values in the Economy. Princeton University Press, Princeton.
Zak, P.J., 2008b. Rational rationality. In: Psychology Today Blogs: The Moral Molecule, Retrieved from http://blogs.psychologytoday.com/blog/the-moral-molecule/200810/rational-rationality.
Zak, P.J., 2005. Trust: a temporary human attachment facilitated by oxytocin. Behavioral and Brain Sciences 28, 368–369.
Zak, P.J., Barraza, J., 2009. Empathy and collective action. Working paper.
Zak, P.J., Borja, K., Matzner, W.T., Kurzban, R., 2005a. The neuroeconomics of distrust: sex differences in behavior and physiology. American Economic Review Papers and Proceedings 95, 360–364.
Zak, P.J., Fakhar, A., 2006. Neuroactive hormones and interpersonal trust: international evidence. Economics & Human Biology 4, 412–429.
Zak, P.J., Knack, S., 2001. Trust and growth. The Economic Journal 111, 295–321.
Zak, P.J., Kurzban, R., Ahmadi, Swerdloff, R.S., Park, J., et al., 2009. Testosterone administration decreases generosity in the ultimatum game. PLoS ONE 4, e8330, doi:10.1371/journal.pone.0008330.
Zak, P.J., Kurzban, R., Matzner, W.T., 2005b. 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, 224–227.
Zak, P.J., Stanton, A.A., Ahmadi, S., 2007. Oxytocin increases generosity in humans. Public Library of Science ONE 2, e1128.