Dishonesty in Everyday Life and Its Policy Implications

Nina Mazar and Dan Ariely

Dishonest acts are all too prevalent in day-to-day life. This article examines some possible psychological causes for dishonesty that go beyond the standard economic considerations of probability and value of external payoffs. The authors propose a general model of dishonest behavior that includes internal psychological reward mechanisms for honesty and dishonesty, and they discuss the implications of this model in terms of curbing dishonesty.

Companies such as Enron, Tyco, WorldCom, and Adelphia are associated with some of the biggest financial scandals in U.S. corporate history since the Great Depression. These corporations exemplify how the boom years of the 1990s have been accompanied by a serious erosion in business principles and, especially, ethics in the accounting and auditing profession. The Brookings Institution estimated that the Enron and WorldCom scandals alone would cost the U.S. economy approximately $37–$42 billion of gross domestic product in the first year alone (Graham, Litan, and Sukhtankar 2002). For comparison, this is approximately the amount that the U.S. federal government spends each year on homeland security (De Rugy 2005).

In addition to corporate scandals, almost all companies present their employees with the conflict between selfishly pursuing their own financial goals and being honest. Brokerage companies represent perhaps the most clear examples of such companies. Brokers are rewarded on the basis of the volume or profitability of the business they transact; such a system provides ample opportunities for conflicts of interest. Although brokers are supposed to act in their clients' best interests, the commissions system can tempt them to choose personal gains over their clients' interests; for example, brokers may experience pressure to buy and sell when they stand to gain larger commissions, they may recommend stocks or funds that are suitable for their interests but not for the client's interest, they may delay the trades ordered by their clients to invest their own money first, or they may misuse knowledge of a large impending order (Davis 2004; McDonald 2002).

Companies and their employees are not alone in the playground of dishonesty (Murphy and Laczniak 1981). In addition to corporate scandals and individuals within companies who behave dishonestly, consumers also behave in ways that are ethically questionable (Bagozzi 1995; Vitell 2003). For example, in a recent survey conducted by Accenture (2003) on insurance fraud, approximately 25% of U.S. adults approved of overstating the value of claims to insurance companies, and more than 10% indicated that submitting insurance claims for items that were not lost or damaged or for treatments that were not provided is acceptable. According to Accenture's press release, the Insurance Services Office estimates that the cost of fraud in the U.S. property and casualty industry is approximately 10% of total claims payments, or $24 billion annually. Similar incidents of consumer fraud can be found in the retail industry. According to the National Retail Federation (Speights and Hilinski 2005), "wardrobing," or the return of used clothing, was estimated to cost approximately $16 billion in the United States in 2002.

Another domain that is central to consumers' unethical behavior involves intellectual property theft, such as music, film, and software piracy. Although the standards and morals linked to such behaviors are not yet well established, it is clear that the economic implications of these endeavors are large. As the Office of the United States Trade Representative estimated, intellectual property theft worldwide costs U.S. companies at least $250 billion a year, a staggering statistic considering that the copyright industries make up approximately 6% of the U.S. gross domestic product ($626 billion) and employ 4% of the U.S. workforce (U.S. Department of Justice 2004). Perhaps the largest contribution to consumer dishonesty comes from tax deception, including issues such as omitting income and inflating deductions. A recent Internal Revenue Service (IRS) study based on special audits of randomly selected individual income tax returns for the 2001 tax year estimates that the "tax gap"—the difference between what the IRS estimates taxpayers should pay and what they actually pay—is somewhere between $312 billion and $353 billion annually. These numbers translate into an overall noncompliance rate of 15% to 16.6% (Herman 2005).

These examples represent only a subset of everyday deception by companies, individuals within companies, and individual consumers. Together, they contribute to the U.S. economy losing hundreds of millions of dollars in tax revenues, wages, and investment dollars, as well as hundreds of thousands of jobs each year.[1] As the damages to society's

Nina Mazar is a postdoctoral associate (e-mail: ninam@mit.edu), and *Dan Ariely* is Luis Alvarez Renta Professor of Management Science (e-mail: ariely@mit.edu), Sloan School of Management, Massachusetts Institute of Technology.

[1]Not everybody loses in every dishonest act; those who committed the act can gain.

ISSN: 0743-9156 (print), 1547-7207 (electronic)

Vol. 25 (1) Spring 2006, 117–126

welfare become more apparent, substantial money is being invested in special governmental task forces (e.g., the Department of Justice's task force for intellectual property issues), programs, and laws to fight dishonest behavior. However, little has proved successful so far. For example, although the IRS is ramping up its audits on high-income taxpayers and corporations, focusing more attention on abusive shelters, and launching more criminal investigations, the overall tax gap has not changed much as a reaction to these measures. As IRS officials publicly state, too many audits result in no change in the amount of taxes paid. That means that these large and expensive efforts for increased compliance seem to be a huge waste of time and money for the taxpayer and the IRS (Herman 2005).

Similar disappointing results can be observed, for example, in industry antipiracy measures, such as investing in technologies for better copy protection of CDs and DVDs, blocking unauthorized downloads, improving consumer identifiability, and increasing the number of litigations. Evidence for the number of litigations was reported in a recent article at law.com, stating that music industry lawsuits against individuals have been ineffective at cutting peer-to-peer music swapping. Although the industry measures seem to have contributed to an increased awareness of copyright laws and lawsuit campaigns, people seem relatively unintimidated by them. An April 2004 survey revealed that 88% of children between the ages of 8 and 18 years understood that peer-to-peer music downloading is illegal, but despite this, 56% of these children admitted to continued music downloading (Von Lohmann 2004).

Given this limited success in curbing dishonest behavior, there are two possible approaches for understanding and limiting future dishonesty. The first approach assumes that the current strategy for curbing dishonesty is the correct one but that it is not practiced sufficiently or with sufficient force. With this approach, the major two variables—the probability of being caught and the magnitude of punishment—should increase, thus reducing or eliminating dishonesty. At the extreme, this would mean "cutting off the right arm" for minor crimes and cutting more important organs for more severe crimes. The second approach, which is the one we take in this article, questions whether the typical path taken to reduce dishonesty (i.e., increasing the probability of being caught and the magnitude of punishment) is the correct path. In the next section, we elaborate on these two approaches. The first approach falls under the standard rational model, and the second one falls under the psychological theory of honesty.

Economic and Psychological Theories of (Dis)Honesty

Economic Theories of the Effect of External Incentives

The standard economic perspective of the *Homo economicus* is one in which the individual is a rational, selfish human being who is interested only in maximizing his or her own payoffs. This rational individual knows what he or she wants and does not want and is able to perform corresponding trade-offs to select the option that is expected to deliver the greatest positive surplus (Hobbes and Macpherson 1968; Smith and Skinner 1997, 1999). From this perspective, there is nothing special about the decision to be honest, which depends solely on the expected external benefits (e.g., more money, a better position) and expected external costs (e.g., paying a fine, loosing a job) to the individual (Hechter 1990; Lewicki 1984); the higher the external rewards from being dishonest, the higher is the extent to which an individual engages in dishonest behavior (see Figure 1). Such a cost–benefit trade-off means that decisions about honesty are like every other decision that individuals face.

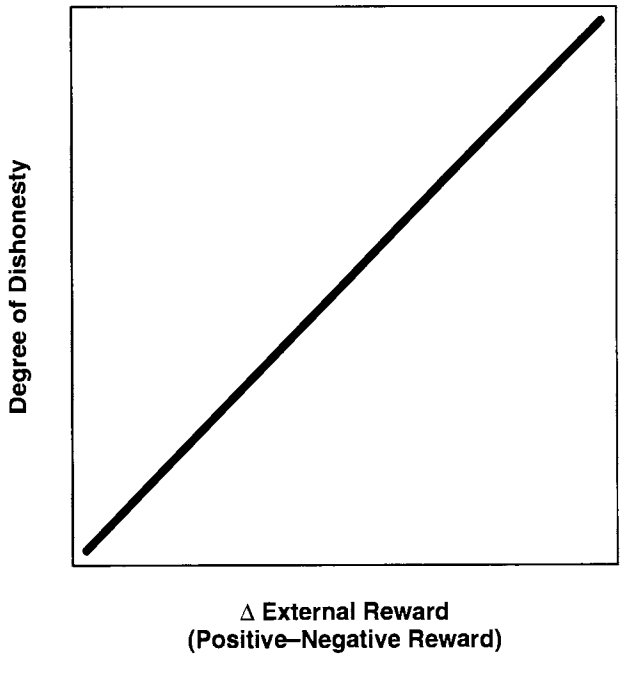
Because this view of dishonesty has been adopted in general, and has been adopted in legal theorizing in particular, efforts to limit dishonesty have assumed that the only ways to make progress are restricted to the external costs and benefits of a dishonest act. The ensuing emphasis on the pervasiveness of police force and magnitude of punishment are the two simplest ways to manipulate these external costs (see Figure 1).

Psychological Theories of the Effect of Internal Incentives

Internal Rewards for Virtuous Behavior

In contrast to the classic economic perspective, there is ample evidence from different academic fields, such as psychology, sociology, anthropology, behavioral and experimental economics, and neuroeconomics and neuroscience, that in addition to the external reward mechanisms, there are also internal reward mechanisms and that these exert influence on people's decisions. Economists such as Ernst Fehr (see, e.g., Fehr and Fischbacher 2003, 2004; Fehr and

Figure 1. The Relationship Between Expected Net Benefits of Dishonesty and the Propensity for Dishonesty According to a Theory That Includes Only External Rewards



Gachter 2002) and James Andreoni (see, e.g., Andreoni, Harbaugh, and Vesterlund 2003; Andreoni and Miller 2006) have repeatedly demonstrated altruism and reciprocity in social dilemma games. For example, in an ultimatum game, two people have the opportunity to split a certain amount of money through a one-time-only, anonymous interaction. One of the two players, the proposer, offers a division of the money. The second player, the receiver, must decide whether he or she wants to accept or reject this proposition. If the receiver rejects the proposed division, both players go home without any money. If the receiver accepts the division, each player receives the amount that the proposer offered. From a game theoretical, rational point of view, the proposer should split the money unequally, in favor of him- or herself. After all, it is a one-time game, so reciprocation is not an issue. In addition, offering even $.01 to the second player should make the receiver accept the offer because that player will be better off with this amount than with nothing. Instead, however, a majority of the offers are typically split equally, and many offers that are not split equally are rejected. These types of results demonstrate that people care about more than maximizing their own monetary payoffs and that these other considerations include social utility and the care for others' outcomes. Uri Gneezy (2005) conducted a related study that emphasized this point in the context of (dis)honesty. Gneezy used a simple deception game with asymmetric information to demonstrate how people act selfishly, insofar as they maximize their own payoffs, but they are also sensitive to the costs that their lies impose on others. These results are important when we consider daily deception because there are many differences in wealth (e.g., between employees and employers, between consumers and corporations) and in perceptions of the cost that deception creates for the other party. Essentially, these results suggest that people will be more dishonest when they face wealthier counterparts and when the cost of the deception for these counterparts seems lower.

Thus, an important question for economists is, Why do people consider more than "just" their material payoffs? To address this question, Joseph Henrich and colleagues (2001) undertook a large cross-cultural study of behavior in ultimatum, public good, and dictator games. They analyzed the behavior of people from 15 small-scale societies in 12 countries exhibiting a wide variety of economic and cultural conditions. The sample consisted of foraging societies; societies that practice slash-and-burn horticulture; nomadic herding groups; and sedentary, small-scale agriculturalist societies. Two of their main findings in support of the hypothesis of internalized reward mechanisms are that (1) the observed behavior varied substantially across societies and (2) people's preferences were not exogenous as the classical model would predict but rather were shaped by their society's characteristic economic and social interactions of everyday life. In other words, socialization is a key to the development of internalized reward mechanisms.

Adding to the behavioral evidence for internal rewards, recent findings from neuroscience and neuroeconomics provide further credence for the existence of internalized reward mechanisms and point to the brain structure that might be implicated in their activation. Thanks to the recent, rapid diffusion of brain-imaging studies, such as positron emission tomography or event-related functional magnetic resonance imaging, groundbreaking developments have added more pieces to the puzzle of reward processing. Several studies in these fields have identified certain regions in the dorsal and ventral striatum, especially the caudate nucleus and the nucleus accumbens, to represent a brain's pleasure center that can be activated through different forms of rewards (see also related studies in rats and primates; e.g., Olds and Milner 1954; Schultz et al. 1992). For example, human beings show significantly increased striatal activity during the anticipation of monetary gains (Knutson et al. 2001), pleasant tastes (O'Doherty et al. 2002), or beautiful faces (Aharon et al. 2001). More important for our point, however, is that the same brain regions are also activated in anticipation of satisfying social outcomes (i.e., social rewards). Rilling and colleagues (2002) show how the striatum "lit up" when people rewarded cooperators. Seemingly contrary to these findings, De Quervain and colleagues (2004) report similar findings when people punished defectors; this result holds even if the punishment was possible only at a personal material cost. Together, these two studies suggest that people feel good about complying with internalized social norms and values; that is, someone who cooperates *should* be rewarded, and someone who defects *should* be punished to reestablish socially desirable behavior.

Psychology has long argued on behalf of internal reward mechanisms. Most notably, Sigmund Freud and colleagues (Freud 1933; Freud and Strachey 1962; Freud, Strachey, and Gay 1989) lectured extensively about the superego, that is, the part of the self that represents society's moral norms and values that the individual internalizes during the course of his or her early life (see also Campbell [1964] on the internalization of moral norms). The superego acts as an internal judge, rewarding or punishing the individual depending on compliance with these norms and values.

Activation of Internal Rewards Lessens Sensitivity to External Incentives

On the basis of the evidence of internal reward mechanisms described, Mazar, Amir, and Ariely (2005) studied how external and internal reward mechanisms work in concourse to influence people's decisions to be (dis)honest. The general procedure was to contrast financial reward incentives that favor dishonest behavior with people's inherent motivation for honest behavior. Taking the motivation for financial rewards as self-evident, Mazar, Amir, and Ariely tested the extent to which the decision to be (dis)honest is driven by the possible consequences of being caught cheating (external costs) relative to the internal rewards of honesty. In one of the experiments, participants were presented with a set of 50 multiple-choice, general-knowledge questions and were promised $.10 per answer they solved correctly within 15 minutes.

There were four different between-subjects conditions that differed in procedure after participants finished answering the questions on the original test sheets. The four conditions were (1) experimenter graded (control), (2) self-graded (self), (3) self-graded plus shredding (self+), and (4) self-graded plus shredding plus self-payment (self++). In the control condition, participants were told that when they finished answering the questions, they should ask the experimenter for a bubble answer sheet (a sheet with circles to indicate the answers, much like those used for exams that

are scanned and graded electronically) on which to transfer their answers and then take both forms to the experimenter, who would check their performance and pay them.

In the self condition, participants were told that when they finished answering the questions on the original test sheet, they should ask the experimenter for a premarked bubble answer sheet, which included the same circles as in the control condition but with one of the answers premarked as being the correct answer. The respondents' task was to transfer their own answers to the bubble answer sheet and then to count and write down the number of correct answers. When this was completed, they were asked to take both forms to the experimenter, at which point they would be paid accordingly. In this condition, participants had the possibility to cheat at the risk that the experimenter might discover it.

The self+ condition was similar to the self condition, except that participants were instructed that after they transferred their answers to the premarked answer sheet, they should walk to a shredder, shred their original test sheet, and take only the answer sheet to the experimenter, at which point they would be paid accordingly. This condition offered a lower probability of being caught cheating than the self condition. Indeed, there was no legal way of proving that anyone cheated. The experimenter could be suspicious only about an atypical high number of correct answers.

Because this social interaction could make cheating participants feel uncomfortable, Mazar, Amir, and Ariely (2005) decreased the chance of being caught even further in the final condition (self++). In this condition, participants were instructed to shred both their test sheet and the answer sheet and then to walk over to a large jar with money that was sitting at the end of the table and take the amount that they earned.

This study had four important findings: First, participants in the external condition solved significantly fewer questions than the participants in the three self conditions. Because there was no reason to believe that the skill level of the participants was different in any of these conditions, this result implies that external reward mechanisms exist, and people are dishonest when it is good for them. This result is also along the lines of Schweitzer and Hsee's (2002) findings, according to which participants are more likely to be dishonest as the private information they have becomes more ambiguous.

Second, there was no significant difference between the three self conditions, though the probability of being caught further decreased dramatically from the self to the self+ to the self++ condition (thus, the fear of being caught should also have decreased). This result provides direct evidence that the external disincentives for dishonesty are only a part of a more complex picture.

Third, the magnitude of dishonesty in the three self conditions was relatively small. Participants cheated only 20% of the possible average magnitude and thus were far from maximal dishonesty, as the standard rational model would predict. This result suggests that people seem to possess internal reward mechanisms for honesty because when given the opportunity and incentive to be dishonest, they exhibited dishonest behavior, but this dishonest behavior was limited in its extent (probably by the strength of their internal reward mechanisms).

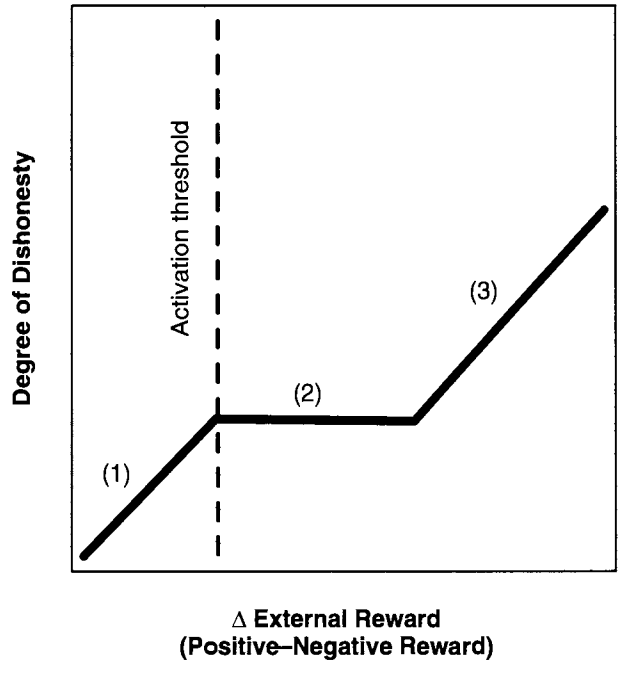
Fourth, these results suggest that, at some level, the act of cheating itself can activate the internal reward mechanism. This means that though low levels of dishonesty might go "unnoticed" from the point of view of the internal reward mechanism, at some point, the magnitude of dishonesty itself can activate this mechanism and limit dishonest behavior.

Together, these findings suggest that the relationship between the external and the internal reward mechanisms is complex. In particular, we hypothesize that the internal reward mechanism is either active or inactive (in the same way that people categorize actions and people as honest or dishonest rather than in a more continuous way) and influences the tendency for acting (dis)honestly as a step function: First, below a certain level of dishonesty, the internal reward mechanism may not be activated at all and thus does not influence behavior; that is, the propensity for dishonesty in such cases is a function of external cost–benefit considerations (see Figure 1). Second, beyond the activation threshold, when the dishonest act is noticeable, the internal reward mechanism is activated and, at least within a certain range, exerts its maximal force independently of the level of external rewards. As a consequence, when a person's internal standards become activated, they override the sensitivity to external rewards such that the propensity for dishonesty becomes independent of increased external rewards for being dishonest (within a certain range). Third, arguably, it is likely that when the external rewards become very large, they become tempting and may ultimately prevail. That is, at some point, a person's internal standards could be overridden or deactivated such that the causes for dishonesty preclude such internal considerations and are based solely on planned dishonesty (a cost–benefit analysis as theorized by the standard rational model). Figure 2 sketches the hypothesized relationship of these components, as Mazar, Amir, and Ariely's (2005) findings illustrate. These findings and the contrast between Figures 1 and 2 suggest that a model for honesty that assumes that people behave like a *Homo economicus* would lead to different predictions about a person's decision to be (dis)honest and, to the extent that these assumptions are wrong, to incorrect prescriptions for the prevention of deception.

Changing the Activation Threshold for Internal Rewards

Note that all the research we discussed supports the theory that internal reward mechanisms indeed exist. However, there is also research that suggests that it is possible to move the activation threshold, that is, to cause internal reward mechanisms for honesty to be more active, or to "kick in" earlier. Representative of this kind of evidence is the extensive body of research in psychology on objective self-awareness by Duval and Wicklund (1972). According to Duval and Wicklund, objective self-awareness represents attention directed inward that induces self-evaluation in relation to standards that are salient or accessible in the immediate situation, which in turn increases motivation to meet the standard (see also Carver and Scheier 1998). Particular situations such as being in front of a real or implied audience (Duval and Wicklund 1972), being individualized, standing in front of a mirror (Carver and Scheier 1978), or writing short stories about oneself (Fenigstein and Levine 1984) can increase an individual's awareness of him- or her-

Figure 2. The Relationship Between Expected Net Benefits of Dishonesty and the Propensity for Dishonesty Based on the Results of Mazar, Amir, and Ariely's (2005) Study



Notes: Mazar, Amir, and Ariely's (2005) results show that dishonesty itself can activate the internal reward mechanisms, which in turn interrupt the standard relationship between external rewards and the propensity for being dishonest (as postulated by classic economics).

self as an object in the world. When awareness of the self is increased, people are also more likely to be aware of discrepancies between how they want to view themselves (the ideal self) and how they actually behave. Given this tension and the discomfort it can create, people who are aware of it might work actively to reduce this discrepancy by either shifting attention away from the self or changing their behavior to act more in accordance with their ideal self. In the domain of deception, this means that higher self-awareness might lead to more honest behavior.

Beaman and colleagues (1979) provide a well-known example of the effect of self-focused attention on increasing the alignment between behavior and internal standards. In this famous experiment, which was conducted during Halloween, trick-or-treating children entered a house and were told by the experimenter to take only one candy; then, the experimenter left the children alone in front of the candies. They found that children who were individualized by being asked their names and addresses were more likely to take only one candy. In addition, these children were even more likely to take only one candy when there was a mirror directly behind the candy bowl. Mazar, Amir, and Ariely (2005) use a different manipulation to influence self-awareness; specifically, they use a manipulation that is more directly related to honesty to test whether an increased awareness leads to more honest behavior on a math test. In their experiment, participants were told to write down either as many of the Ten Commandments as they could remember (increased self-awareness of honesty) or the names of ten books that they read in high school (control). They had two minutes for this task before they moved on to an ostensibly separate task: the math test. The task in the math test was to search for number combinations that added up to exactly ten. There were 20 questions, and the duration of the experiment was restricted to five minutes. After the time was up, students were asked to recycle the test form they worked on and indicate on a separate collection slip how many questions they solved correctly. For each correctly solved question, they were paid $.50. The results showed that students who were made to think about the Ten Commandments claimed to have solved fewer questions than those in the control. Moreover, the reduction of dishonesty in this condition was such that the declared performance was indistinguishable from another group whose responses were checked by an external examiner. This suggests that the higher self-awareness in this case was powerful enough to diminish dishonesty completely. Yet another method that proved successful in increasing self-awareness with the consequence that students cheated less on that math test was to make students sign an honor code before beginning the test. This finding is particularly interesting because it provides some evidence for the effectiveness of a simple commitment device that schools, companies, and the government can easily adopt.

A different kind of research that can be interpreted in line with the self-awareness theory is based on Daniel Schachter and colleagues' work on the role of emotional arousal for moral behavior. For example, Schachter and Latane (1964) tested the effects of a tranquilizer in a situation in which participants could cheat. They found that participants cheated significantly more after having taken the tranquilizer, a sympathetic inhibitor, than did the placebo controls (see also Schachter and Singer 1962). These results suggest that the activation of the threshold is based on arousal and that when arousal is decreased (even by artificial means, such as a tranquilizer), the threshold is less likely to be activated, and thus dishonesty is likely to be more prevalent. This idea is further developed by Dienstbier and Munter (1971), who show that it is not emotional arousal per se that influences the tendency to cheat but rather a person's understanding or interpretation of the meaning and significance of that arousal (see also Dienstbier 1972, 1975).

Self-Deception

Thus far, the discussion has focused on deceptive acts and the activation of a threshold that ignites the internal rewards. In our conceptualization of the threshold, it is important that the amount of deception in and of itself can activate the threshold. An added complexity comes from the idea of self-deception, in which people can reframe an act in a way that makes them not perceive it as dishonest. Under such reframing (self-deceptions), dishonest acts would not contribute to the activation of the threshold and thus would not influence the tendency for honesty. As a consequence, even if an individual has internalized standards for honesty, acts of dishonesty that do not influence the activation of the threshold (self-deceptions) would not influence behavior, and the relationship between external rewards and the tendency for dis-

honesty would remain in Zone 1 of Figure 2 (never reaching the threshold).

In general, self-deception represents a biased, self-serving information flow within an individual—that is, an active but unconscious misrepresentation of reality to the conscious mind (Trivers 2000). Although it seems to be a paradox that a person could deceive him- or herself, casual observation suggests that people are effective in maintaining unrealistically positive views of themselves. People maintain beliefs in their intelligence, competence, and moral worth in the face of their sometimes foolish, incompetent, and immoral behavior. Similarly, people frequently estimate that they are, for example, better, more intelligent, or more beautiful than average. Because typically far more than 50% of people estimate this "better-than-average" effect, aggregating these data clearly violates logic (e.g., Alicke et al. 1995). Other researchers have shown how people can be led to believe false or biased "facts" about their own past (see research on memory distortion and suppression of unwanted memories; e.g., Anderson, Cohen, and Taylor 2000; Loftus 1994; Schachter and Dodson 2002), and they can convince themselves of certain motivations for their behavior, thus hiding their true intentions. Therefore, self-deception can be successful even in the most extreme cases: For example, doctors who participated in genocide in Nazi Germany managed to convince themselves of the rectitude of their actions (Lifton 1986).

Although self-deception can be beneficial in the short run because it enables people to maintain self-regard, the inflated self-perception can be problematic in the long run when that bogus self-regard has costs. Norton and Ariely (2005) show this in the context of an IQ test. In one of Norton and Ariely's experiments, people participated in two sequential IQ tests. Half of the participants in the first test were provided with an answer key on the bottom of their test sheet, whereas the other half did not have an answer key. When the first test was over, each participant learned how many questions he or she solved correctly. After that, they were asked to predict how many questions they would solve on a second, similar test without an answer key. The prediction task was designed such that people had a monetary incentive to be accurate. Payment depended more on the accuracy of their predictions and less on their performance such that the dominant strategy was to predict accurately.

The results of this experiment show that in the first round, participants who had the answer key performed better on the test than participants without the answer key and that there was no performance difference on the second IQ test when none of the participants were provided with an answer key. This result suggests that the presence of the answers allowed participants to solve more problems correctly. More important, it seems that participants who had the answer key in the first test believed that this better performance was actually due to their greater intelligence, as reflected in their higher estimates of their performance on the second test. As a consequence, participants who deceived themselves by having an inflated perception of themselves made less money. Notably, when Norton and Ariely (2005) asked a different group of people to predict how they would perform in such an experiment, participants predicted that if they had the answer key on the first test but not on the second, they would perform better on the first test but not on the second test. This result suggests an added danger of self-deception, namely, that people believe that they are immune to it.

Researchers have explained this paradox of not knowing about self-deception under the assumption that a person can simultaneously store both true and false knowledge, with a bias toward the true knowledge being stored in the unconscious and the false (i.e., misrepresented) knowledge being stored in the conscious (Greenwald 1997). Trivers (2000) points out that from an evolutionary standpoint, this way of organizing knowledge has the advantage that self-deception not only must act primarily in the service of fooling oneself but also can act in the service of deceiving others. The latter is true because an outside observer initially interacts with the conscious mind of the deceiver, and if the deceiver is also deceiving him- or herself, the conscious mind comprises only the false information. The true information would be hidden in the unconscious. If this is the case, cues, such as a higher-pitched voice, increased pupil size, or lip pressing, which in general accompany attempted, conscious deception, should not be available to the observer, thus making the deceit more difficult to detect (DePaulo and Morris 2004). As a consequence, even if people are fully self-aware (i.e., internal reward mechanism is active) and the net utility of deception is negative (i.e., costs loom larger than benefits), deception that eludes the person who is committing the dishonest act (i.e., self-deception) might not completely vanish.[2]

Policy Guidelines for Reducing Dishonesty

The standard rational model of decisions about honesty and dishonesty assumes that people trade off only external costs and benefits of an outcome. In contrast, the psychological model we have sketched assumes that decisions about honesty also include considerations of internal reward mechanisms. In addition, we argue that decisions about honesty can sometimes be not cognizant, for example, at levels of activation below the threshold or when self-deception is involved. As a consequence, making the right policy recommendation to decrease dishonesty depends particularly on the analysis of what is driving the deceit in a particular situation. Our literature review suggests that there are four general drivers of dishonesty: (1) lower external costs and relatively higher benefits of deception; (2) lack of social norms, which results in a weak internal reward mechanism; (3) lack of self-awareness, which primes the activation of the internal reward mechanism; and (4) self-deception. In the following subsections, we elaborate on how these different perspectives on dishonesty can translate into different approaches for reducing or curbing dishonesty.

When Dishonest Behavior Is Caused by External Rewards

If the cause for deception lies solely in greater external benefits than costs of the dishonest act, the solution is sim-

[2]We assume that the external reward mechanism as postulated by classic economics is always active in at least normally healthy people.

ple: The costs for dishonest actions must be greater than their expected benefits. This can be achieved by increasing either the probability of being caught or the severity of the punishment. Thus, if the cause of dishonesty is based solely on an imbalance of external costs and benefits, the standard legal approach of controlling the external costs is appropriate. This theory implies that it is appropriate to introduce governmental task forces, such as the Department of Justice's task force on intellectual property, which, among other combat strategies, invests in increasing the number of specially trained prosecutors. The same is true for the IRS's increase in audits and the music industry's aggressive increase in filing lawsuits against individual deceivers.

Even if the cause for deception is related to the cost–benefit analysis, there might be ways to increase the effectiveness and efficiency of measures to combat dishonest behaviors. For example, if the probability of being caught and the magnitude of punishment are evaluated differently, and research suggests that the probability of punishment is more important than increasing the severity of the punishment (Nagin and Pogarsky 2003), it might be best to allocate efforts accordingly. However, even if legislators decide to invest more effort in the probability of detection and decrease the magnitude of punishment (e.g., moving from a $500 fine for not stopping at a stop sign with 10% probability of being caught to a $100 fine with a 50% probability of being caught), there is still the question of what is the optimal probability for deterrence.

Informing the question of what is the optimal probability for deterrence, research by Barkan, Zohar, and Erev (1998) suggests that the best approach is eliminating altogether the probability of being caught, that is, moving to nonprobabilistic punishments. Their main argument is that events that have low probability are unlikely to occur (by definition) and thus can have a perverse effect on learning, such that people who violate the rule and are not caught receive a *positive* reward for the violation, which causes them to underestimate the probability of being caught and, over time, increases their tendency to behave in this undesired way (see also Erev et al. 2003). According to this perspective, a person who expects that driving through a red light would involve a $500 fine in 5% of the cases is more likely to drive through it than a person who has the same expected value but with certainty of being caught (i.e., a definite $25 fine). More important, over time, the person in the probabilistic punishment setting is going to discount the probability of the punishment further (as long as he or she is not caught), which in turn will lead to an even greater tendency for violation. Eliminating the probabilistic component from all undesirable behaviors is impossible, but it is clear that there are some cases (e.g., driving through an intersection at a red light) in which this is possible and desirable.

When Dishonest Behavior Is Caused by the Internal Reward Mechanism

If the reason for dishonest actions lies in a lack of internalized social norms, our primary recommendation would be to invest in educational efforts and socialization to increase the strength of the internal reward mechanism. The key questions in this context are, How can this best be done, and is there a critical age period for the internalization of such mechanisms (as in language and visual development)? For example, educational efforts can be integrated in schools, social clubs, or religious institutions. Another possibility that is increasingly exercised by the government and the music, film, and software industries is to feature public messages in an attempt to build a social norm so that a particular type of behavior (e.g., illegally downloading music or movies) becomes socially undesirable and frowned on. Other efforts could illustrate how such acts can hurt the rank-and-file workers, not just the big corporations, by reducing their job security or pay (see, e.g., television advertisements and movie trailers launched by the Motion Picture Association of America in 2003 and 2004).

When the effects of such efforts on the development of socially based internal reward mechanisms are understood, it is important to go a step further and ask what the limits of such efforts should be and whether society should allow all ideologies to participate in the creation of such internal rewards (e.g., what about racial prejudices or particular cults?). The question about the types of internal reward mechanisms that society could develop or not develop is particularly important if we consider the likely possibility that this process might have a critical period in which younger people are much more sensitive to such influences and that when the critical age is reached, these mechanisms remain relatively stable for the rest of their lives. Given the higher sensitivity of younger adults to social influence and advertising, society might want to place careful boundaries for the development of such socially based internal representations by different interested parties, such as religious, financial, and social institutions.

If dishonest behavior occurs not because of a lack of social norms but simply because of a lack of self-awareness and, thus, the degree to which these internalized social norms are activated, it is important to make use of contextual cues that increase awareness when deception is about to happen, namely, at the point of temptation. For example, the IRS could slightly change its forms by making them more personal or by asking people to sign an honor code of sorts before they begin filling out the forms. Another possibility worth trying might be to include a survey that asks tax payors questions such as how much they care about their country, how important honesty was to their parents, how many people they think lie on their taxes, or what the typical profile is of tax deceivers.

The consideration of internal rewards also suggests that the theory of optimal punishment (optimal punishment trades off the benefits of deterrence and the cost of punishing innocent people) should be reconsidered with these inputs in mind. If the punishment magnitude is determined in a way that makes the costs slightly higher than the benefits and if these costs also include internal costs, the optimal punishment will be lower by that amount. For example, if the expected benefit for a particular crime is Y and the internal reward for honesty is X, the standard rational model would prescribe a punishment with an expected magnitude of $-(Y + \varepsilon)$, whereas the model that includes internal rewards would prescribe $-(Y + \varepsilon) + X$. The complexity is that not everyone has the same level of internal reward mechanisms, and to the degree that these are unobservable, it is difficult to assess the true level of optimal punishment (though it is possible that someday there will be a test for

this). Conversely, signs of repeated criminal behavior, for example, can be taken as an indication for a lower level of internal reward mechanisms, causing the magnitude of X to be updated as lower. This type of framework, in which X is an individual variable, has the potential to help build a theory of repeated punishment with the same desired principles of optimal punishment but with more effectiveness. (Currently, repeated crime is punished more severely, such as in California's "three-strikes-and-you're-out" approach, but there is no theory or logical guideline for the magnitude of these policies.)

When Dishonest Behavior Is Caused by Self-Deception

Finally, deception of the self-deception variety is difficult to fight. Self-deception due to a self-serving bias is very robust. Several academics have shown that paying people to be more realistic in their view, making people write essays that argue the other side's point of view, or educating people about the self-serving bias are not successful in debiasing people (see, e.g., Babcock and Loewenstein 1997; Babcock, Loewenstein, and Issacharoff 1997). Therefore, the most successful way to fight self-deception might be to eliminate the incentives that spawn the bias and simply eliminate the situations that can give rise to this type of behavior. For example, Bazerman, Loewenstein, and Moore (2002; see also Bazerman and Loewenstein 2001) argue that if deceptive audits by accounting professionals are mainly due to self-deception, it might be more effective to pass laws or enforce standards among the accounting profession that bar auditors from offering both consulting and tax services to clients, prohibit hiring accountants through clients, and allow only limited-time contracts. As these examples show, fighting deception caused by self-deception requires serious interventions that limit substantially the freedom and self-determination of people in certain situations.

Conclusion

In summary, there is no question that dishonesty is prevalent in daily life. The standard economics perspective considers one cause for dishonesty—external reward mechanisms—and thus emphasizes the probability of being caught and the magnitude of punishment as the only ways to overcome dishonesty. In contrast, the psychological perspective we present herein suggests that dishonesty is also influenced by internal reward mechanisms and that such mechanisms should be taken into account when considering effective measures for limiting dishonesty in daily life. Moreover, the psychological perspective suggests that internal and external rewards are not simply additive but also take a particular functional form (see Figure 2). With this functional form in mind, the psychological approach for reducing dishonesty could be based on increasing the long-term effectiveness of internal rewards (education), increasing the short-term effectiveness of internal rewards (contextual cues), or eliminating the possibility of dishonest acts when the cause could be attributed to self-deception. When the role of internal rewards is better understood, both preventions and punishments of dishonesty can be made more effective and efficient.

References

Accenture Inc. (2003), "One-Fourth of Americans Say It's Acceptable to Defraud Insurance Companies, Accenture Survey Finds," press release, (February 12), (accessed January 31, 2006), [available at http://www.accenture.com/xd/xd.asp?it=enweb&xd=_dyn%5Cdynamicpressrelease_577.xml].

Aharon, Itzhak, Nancy Etcoff, Dan Ariely, Christopher F. Chabris, Ethan O'Connor, and Hans C. Breiter (2001), "Beautiful Faces Have Variable Reward Value: fMRI and Behavioral Evidence," *Neuron*, 32 (3), 537–51.

Alicke, Mark D., M.L. Klotz, David L. Breitenbecher, Tricia J. Yurak, and Debbie S. Vredenburg (1995), "Personal Contact, Individuation, and the Better-Than-Average Effect," *Journal of Personality and Social Psychology*, 68 (5), 804–825.

Anderson, Stephen J., Gillian Cohen, and Stephanie Taylor (2000), "Rewriting the Past: Some Factors Affecting the Variability of Personal Memories," *Applied Cognitive Psychology*, 14 (5), 435–54.

Andreoni, James, William T. Harbaugh, and Lise Vesterlund (2003), "The Carrot or the Stick: Rewards, Punishments, and Cooperation," *The American Economic Review*, 93 (3), 893–902.

——— and John H. Miller (2006), "Analyzing Choice with Revealed Preference: Is Altruism Rational?" in *Handbook of Experimental Economics Results*, Vol. 1, Charles Plott and Vernon L. Smith, eds. Amsterdam: Elsevier Science, forthcoming.

Babcock, Linda and George Loewenstein (1997), "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives*, 11 (1), 109–126.

———, ———, and Samuel Issacharoff (1997), "Creating Convergence: Debiasing Biased Litigants," *Law & Social Inquiry*, 22 (4), 913–25.

Bagozzi, Richard P. (1995), "Reflections on Relationship Marketing in Consumer Markets," *Journal of the Academy of Marketing Science*, 23 (4), 272–77.

Barkan Rachel, Dov Zohar, and Ido Erev, (1998), "Accidents and Decision Making Under Risk: A Comparison of Four Models," *Organizational Behavior and Human Decision Processes*, 74 (2), 118–44.

Bazerman, Max H. and George Loewenstein (2001), "Taking the Bias Out of Bean Counting," *Harvard Business Review*, 79 (1), 28.

———, ———, and Don A. Moore (2002), "Why Good Accountants Do Bad Audits," *Harvard Business Review*, 80 (11), 96–102.

Beaman, Arthur L., Bonnel Klentz, Edward Diener, and Soren Svanum (1979), "Self-Awareness and Transgression in Children: Two Field Studies," *Journal of Personality and Social Psychology*, 37 (10), 1835–46.

Campbell, Ernest Q. (1964), "The Internalization of Moral Norms," *Sociometry*, 27 (4), 391–412.

Carver, Charles S. and Michael F. Scheier (1978), "Self-Focusing Effects of Dispositional Self-Consciousness, Mirror Presence, and Audience Presence," *Journal of Personality and Social Psychology*, 36 (3), 324–32.

——— and ——— (1998), *On the Self-Regulation of Behavior*. New York: Cambridge University Press.

Davis, Ann (2004), "Open Secrets; Head of the Line: Client Comes First? On Wall Street, It Isn't Always So; Investing Own

Money, Firms Can Misuse Knowledge of a Big Impending Order; Mischief in the 'Back Books,'" *The Wall Street Journal*, (December 16), A1.

DePaulo, Bella M. and Wendy L. Morris (2004), "Discerning Lies from Truth: Behavioral Cues to Deception and the Indirect Pathway of Intuition," in *The Detection of Deception in Forensic Contexts*, Pär Anders Granhag and Leif A. Strömwall, eds. New York: Cambridge University Press, 15–40.

De Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schelthammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr (2004), "The Neural Basis of Altruistic Punishment," *Science*, 305 (5688), 1254–58.

De Rugy, Veronique (2005), "What Does Homeland Security Spending Buy?" working paper, American Enterprise Institute for Public Policy Research, (April 1), (accessed January 31, 2006), [available at http://www.aei.org/publications/pubID.21483/pub_detail.asp].

Dienstbier, Richard A. (1972), "The Role of Anxiety and Arousal Attribution in Cheating," *Journal of Experimental Social Psychology*, 8 (2), 168–79.

——— (1975), "An Emotion-Attribution Approach to Moral Behavior: Interfacing Cognitive and Avoidance Theories of Moral Development," *Psychological Review*, 82 (4), 299–315.

——— and Pamela O. Munter (1971), "Cheating as a Function of the Labeling of Natural Arousal," *Journal of Personality and Social Psychology*, 17 (2), 208–213.

Duval, Shelley and Robert A. Wicklund (1972), *A Theory of Objective Self Awareness*. New York: Academic Press.

Erev, Ido, Paul Ingram, Ornit Raz, and Dror Shany (2003), "On the Possibility of Gentle Rule Enforcement," working paper, Department of Industrial Engineering and Management, Technion.

Fehr, Ernst and Urs Fischbacher (2003), "The Nature of Human Altruism," *Nature*, (October 23), 785–91.

——— and ——— (2004), "Social Norms and Human Cooperation," *Trends in Cognitive Sciences*, 8 (4), 185–90.

——— and Simon Gachter (2002), "Altruistic Punishment in Humans," *Nature*, (January 10), 137–40.

Fenigstein, Allan and Michael P. Levine (1984), "Self-Attention, Concept Activation, and the Causal Self," *Journal of Experimental Social Psychology*, 20 (3), 231–45.

Freud, Sigmund (1933), *New Introductory Lectures on Psycho-Analysis*. New York: W.W. Norton.

——— and James Strachey (1962), *The Ego and the Id*. New York: W.W. Norton.

———, ———, and Peter Gay (1989), *An Outline of Psycho-Analysis*, stand. ed. New York: W.W. Norton.

Gneezy, Uri (2005), "Deception: The Role of Consequences," *American Economic Review*, 95 (1), 384–94.

Graham, Carol, Robert E. Litan, and Sandip Sukhtankar (2002), "The Bigger They Are, the Harder They Fall: An Estimate of the Costs of the Crisis in Corporate Governance," working paper, The Brookings Institution, (accessed February 1, 2006), [available at http://www.brookings.edu/views/papers/graham/20020722Graham.pdf].

Greenwald, Anthony G. (1997), "Self-Knowledge and Self-Deception: Further Consideration," in *The Mythomanias: The Nature of Deception and Self Deception*, Michael S. Myslobodsky, ed. Mahwah, NJ: Lawrence Erlbaum Associates, 51–71.

Hechter, Michael (1990), "The Attainment of Solidarity in Intentional Communities," *Rationality and Society*, 2 (2), 142–55.

Henrich, Joseph, Robert Boyd, Sam Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath (2001), "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 91 (2), 73–78.

Herman, Tom (2005), "Study Suggests Tax Cheating Is on the Rise; Most Detailed Survey in 15 Years Finds $250 Billion-Plus Gap; Ramping Up Audits on Wealthy," *The Wall Street Journal*, (March 30), D1.

Hobbes, Thomas and C.B. Macpherson (1968), *Leviathan*. Baltimore: Penguin Books.

Knutson, Brian, Charles M. Adams, Grace W. Fong, and Daniel Hommer (2001), "Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens," *Journal of Neuroscience*, 21 (16), RC159 (1–5).

Lewicki, R.J. (1984), "Lying and Deception: A Behavioral Model," in *Negotiation in Organizations*, Max H. Bazerman and Roy J. Lewicki, eds. Beverly Hills, CA: Sage Publications, 68–90.

Lifton, Robert J. (1986), "Reflections on Genocide," *Psychohistory Review*, 14 (3), 39–54.

Loftus, Elizabeth F. (1994), "The Repressed Memory Controversy," *American Psychologist*, 49 (5), 443–45.

Mazar, Nina, On Amir, and Dan Ariely (2005), "(Dis)Honesty: A Combination of Internal and External Rewards," working paper, Sloan School of Management, Massachusetts Institute of Technology.

McDonald, Ian (2002), "Brokers Get Extra Incentive to Push Funds," *The Wall Street Journal*, (April 8), C17.

Murphy, Patrick E. and Gene R. Laczniak (1981), "Marketing Ethics: A Review with Implications for Managers, Educators, and Researchers," in *Review of Marketing*, Ben M. Enis and Kenneth J. Roering, eds. Chicago: American Marketing Association, 251–66.

Nagin, Daniel S. and Greg Pogarsky (2003), "An Experimental Investigation of Deterrence: Cheating, Self-Serving Bias, and Impulsivity," *Criminology*, 41 (1), 501–527.

Norton, Mike and Dan Ariely (2005), "Self-Deception: How We Come to Believe We Are Better Than We Truly Are," working paper, Sloan School of Management, Massachusetts Institute of Technology.

O'Doherty, John P., Ralf Deichmann, Hugo D. Critchley, and Raymond J. Dolan (2002), "Neural Responses During Anticipation of a Primary Taste Reward," *Neuron*, 33 (5), 815–26.

Olds, James and Peter Milner (1954), "Positive Reinforcement Produced by Electrical Stimulation of Septal Area and Other Regions of Rat Brain," *Journal of Comparative and Physiological Psychology*, 47 (6), 419–27.

Rilling, James K., David A. Gutman, Thorsten R. Zeh, Giuseppe Pagnoni, Gregory S. Berns, and Clinton D. Kilts (2002), "A Neural Basis for Social Cooperation," *Neuron*, 35 (2), 395–405.

Schachter, Stanley and Chad S. Dodson (2002), "Misattribution, False Recognition and the Sins of Memory," in *Episodic Memory: New Directions in Research*, Alan Baddeley, Martin Conway, and John Aggleton, eds. London: Oxford University Press, 71–85.

——— and Bibb Latane (1964), "Crime, Cognition, and the Autonomic Nervous System," *Nebraska Symposium on Motivation*,

Vol. 12, David Levine, ed. Lincoln: University of Nebraska, 221–75.

——— and Jerome Singer (1962), "Cognitive, Social, and Physiological Determinants of Emotional State," *Psychological Review*, 69 (5), 379–99.

Schultz, Wolfram, Paul Apicella, Eugenio Scarnati, and Tomas Ljungberg (1992), "Neuronal Activity in Monkey Ventral Striatum Related to the Expectation of Reward," *Journal of Neuroscience*, 12 (12), 4595–4610.

Schweitzer, Maurice E. and Christopher K. Hsee (2002), "Stretching the Truth: Elastic Justification and Motivated Communication of Uncertain Information," *Journal of Risk and Uncertainty*, 25 (2), 185–201.

Smith, Adam and Andrew S. Skinner (1997), *The Wealth of Nations: Books I–III*. New York: Penguin Books.

——— and ——— (1999), *The Wealth of Nations: Books IV and V*. London: Penguin Books.

Speights, David and Mark Hilinski (2005), "Return Fraud and Abuse: How to Protect Profits," *Retailing Issues Letter*, 17 (1), 1–6.

Trivers, Robert (2000), "The Elements of a Scientific Theory of Self-Deception," in *Evolutionary Perspectives on Human Reproductive Behavior*, Peter Moller and Dori LeCroy, eds. New York: New York Academy of Sciences, 114–31.

U.S. Department of Justice, Office of the Attorney General (2004), "Report of the Department of Justice's Task Force on Intellectual Property," (October 24), (accessed February 1, 2006), [available at http://www.usdoj.gov/criminal/cybercrime/IPTaskForceReport.pdf].

Vitell, Scott J. (2003), "Consumer Ethics Research: Review, Synthesis and Suggestions for the Future," *Journal of Business Ethics*, 43 (1–2), 33–47.

Von Lohmann, Fred (2004), "Is Suing Your Customers a Good Idea?" (September 29), (accessed February 1, 2006), [available at http://www.law.com/jsp/article.jsp?id=1095434496352].