Ethical immunity: How people violate their own moral standards without feeling they are doing so

Chapter · November 2011

CITATIONS
15

READS
2,365

3 authors, including:

Jason Dana
Yale University
68 PUBLICATIONS **3,978** CITATIONS

SEE PROFILE

George Loewenstein
Carnegie Mellon University
504 PUBLICATIONS **77,918** CITATIONS

SEE PROFILE

9

Ethical Immunity: How People Violate Their Own Moral Standards Without Feeling They Are Doing So

Jason Dana
University of Pennsylvania

George Loewenstein
Carnegie Mellon University

and

Roberto Weber
Carnegie Mellon University

INTRODUCTION

Ethics can play an important role in modern economic and organizational contexts. With frequent opportunities to misappropriate others' wealth, mislead other parties such as investors and consumers, and manufacture harmful products, the extent to which economic actors behave unethically and the conditions that lead them to do so become worthy topics of investigation. Indeed, following large-scale examples of individual and corporate malfeasance (e.g., the Nestle baby milk scandal, the WorldCom and Enron financial collapses, and the frauds committed by Bernie Madoff and others), academics, the press, and the public at large seek to understand why such acts were committed and how they might be prevented in the future.

One important question for understanding these events is whether those who were ultimately responsible started on their path recognizing that they were behaving unethically. By now, a rather broad and thorough experimental literature has shown that the desire to appear ethical to oneself and others can exert a powerful influence on human behavior. Even when given strong incentives for behaving in a selfish fashion, with pains taken to verify subjects' understanding of the setup, experimental subjects regularly engage in other-regarding behavior, such as punishing unethical acts by others with no benefit to oneself (Fehr & Gächter, 2002), behaving honestly when dishonesty is profitable and undetectable (Gneezy, 2005), and implementing distributional outcomes that are egalitarian, again often at the expense of one's own egoistic payoff (Fehr & Schmidt, 1999; Loewenstein, Thompson, & Bazerman, 1989). Yet, if ethics is such a powerful motivation, what can account for the seemingly endless corporate scandals, defrauding of investors and employers, and embezzlement? Are the people who commit these acts simply corrupt, or are there specific situations that tend to give rise to unethical behavior, even on the part of people who would like to view themselves as honorable and honest?

In this chapter, we suggest that a study of *behavioral ethics* that takes its cue from the burgeoning field of behavioral economics can explain when and why ethically minded individuals behave unethically. Much of the focus of behavioral economics has been on ways in which people's behavior falls short of fulfilling their own goals. For example, evidence of procrastination and dynamic inconsistency reveals that people are not always capable of following through on their own intertemporal preferences. Similarly, evidence that different but consequentially equivalent frames of a decision lead to different choices reveals that people sometimes contradict themselves due to minor variations in a choice context. Just as people often fail to achieve their own goals or behave consistently, due to systematic errors and biases in judgment and decision making and to varying subtle features of the choice context, people can also systematically fail to live up to their own ethical standards (see, e.g., Chugh, Bazerman, & Banaji, 2005). That is, they may fail to behave as ethically as even they themselves would wish. The study of behavioral ethics—as apart from normative or legal approaches that specify how individuals should or must behave—aims to understand how even well-intentioned people can sometimes behave unethically. Such an understanding can lead us to a broader understanding of the range of ethical failures in business, taking

us beyond the view that those guilty of malfeasance are just "a few bad apples" (see De Cremer, 2009).

Much of the instruction that occurs in the business ethics courses that have become increasingly popular in business schools teaches students how to use ethical principles to discriminate between behaviors that are and are not ethical. An implicit assumption of such approaches, as Tenbrunsel and Messick (2004) noted, is that "by highlighting and emphasizing the moral components of decisions, executives will be more likely to choose the moral path." The study of behavioral business ethics can provide incremental usefulness beyond the traditional approach by helping to explain when and why people who know the rules, and wish to follow them, end up breaking them. Such an understanding can potentially help people to recognize and avoid these types of "ethics traps"; more important, it can provide insights into the design of institutions that reduce the prevalence of such traps, making it easy and natural for people to behave in conformity with their own ethical principles.

Why do people violate their own ethical standards? One important reason, which is our central focus in this chapter, is that they are able to persuade themselves that they are not doing so: People have a remarkable ability to rationalize unethical behavior or to actively position themselves (e.g., through avoidance of compromising information or through delegation of unethical activities to others) to achieve "ethical immunity" by not facing up to the consequences and obvious interpretations of their actions. As a result, people who would like to be honest, generous, and fair, and to see themselves as such, end up behaving dishonestly, selfishly, and in a way that harms others.

In the following sections, we discuss experimental research examining three different "tactics" people use to avoid holding themselves ethically accountable. Each line of research demonstrates that people behave more ethically when they are morally accountable for their actions—that is, when their actions clearly and directly reflect on their ethical standards. When it is possible to avoid such ethical accountability, however, people often do so and end up behaving less ethically. That is, people often seek to avoid moral accountability, perhaps paradoxically, so that they will not feel compelled to behave in accord with their own standards of ethics.

In the section "Diffusing Responsibility," we show that people faced with ethically difficult decisions often prefer to diffuse responsibility for the choice. This diffusion can occur both vertically (e.g., when people

hire intermediaries to do their "dirty work") and horizontally (e.g., when people fail to behave in an ethical fashion themselves, based on the justification that others can be counted on to do so). In the section "Exploiting Uncertainty," we show that people sometimes prefer to be uncertain about the impact of their actions on others so that they can be more self-interested. Finally, in the section "Seeking Justifications," we show that people, often unwittingly, select and weigh information about what is fair in a way that favors themselves.

The research we discuss contributes to a growing body of literature that views ethical failures as something of which even ethical people are capable (Loewenstein, 1996). Indeed, ethical immunity does not occur because people are not, in fact, really ethical. Rather, it is precisely because people are ethical, or at least truly want to be at some level, that they go through the costly contortions that they do to allow themselves to be selfish or dishonest when in fact no one would punish them for their selfishness or dishonesty but themselves.

A NOTE ON METHODOLOGY: THE VALUE OF ECONOMIC LABORATORY EXPERIMENTS FOR UNDERSTANDING WHEN AND WHY PEOPLE WILL BEHAVE ETHICALLY

By now, a large literature on experimental games incorporating monetary incentives has shown that people put a real value on principles such as fairness and honesty. As examples, people often give up nonnegligible amounts of money so that an anonymous other can benefit (see review in Camerer, 2003), and people often forgo opportunities to lie for money even though it is clear they cannot be caught (Shalvi, Handgraaf, & De Dreu, in press). These types of laboratory studies are helpful for understanding ethical behavior because they enable researchers to create very simple environments in which individuals can choose between acts that are unambiguously "ethical" (telling the truth, sharing wealth equally with another who has an equal right to it) and "unethical" (telling a lie for monetary gain, taking wealth from others without any right to it). In such studies, confounding selfish motives can be carefully controlled and hence eliminated as potential causes of behavior.

Take the example of ethical behavior in distributions of wealth. In the simple laboratory dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994), a player in the role of "dictator" can divide an endowment of money with an anonymous other in any way he or she sees fit, with both individuals having equal right to the wealth. The game can be played only once, the identities of the players are never revealed (even to the experimenter), and the recipient is given no chance to retaliate or otherwise respond. In this way, we can distinguish a player's inherent desire to act ethically, by dividing wealth equally, from desires for impression management, quid pro quo, or avoiding punishment. Despite removing all of these selfish motives for sharing, giving across a variety of dictator game experiments remains stubbornly positive (Camerer, 2003).

While laboratory experiments often show that individuals exhibit a high propensity toward some ethical behaviors, this does not mean that they necessarily provide evidence about the degree to which people are or are not ethical or generous in other closely related contexts. Indeed, *no* research can provide such evidence because all research, including field research, necessarily focuses on a limited set of specific situations, with a particular set of characteristics (Falk & Heckman, 2009). Therefore, to say that "people are not selfish" or "people do not lie" from the evidence presented in a single laboratory or nonlaboratory paradigm is of questionable validity (as much as it would be to say that "people always lie" because one finds one domain, poker playing, in which they do so). The merit of the experiments that we review is that they facilitate understanding of the precise factors that cause people to be more or less ethical, holding other factors constant.

THREE TACTICS FOR ACHIEVING ETHICAL IMMUNITY

Diffusing Responsibility

One situation in which people can avoid accountability for difficult ethical decisions is when responsibility for these decisions and for the resulting consequences is diffused (Darley & Latane, 1968). The actions of other people often diffuse responsibility in a fashion that enables individuals to take self-interested actions they would eschew if acting unilaterally.

Here, we define two different types of diffusion of responsibility. *Vertical* diffusion of responsibility occurs when an intermediary is placed between the decision maker and the stakeholders in the decision, making the decision maker feel less responsible for the outcome. For example, firms may use "firing consultants," who provide little other service than carrying out the act of firing their employees, or firms may outsource functions to a contractor that subsequently pays its employees wages that are below the usual standards of the firm. *Horizontal* diffusion of responsibility occurs when decision makers rely on others to act in a stakeholder's best interests rather than doing so themselves. The "bystander effect" (Darley & Latane, 1968) is a classic example, whereby people can fail to intervene and help a person in distress when others are present but almost never do so when alone.

Recent research suggests that vertical diffusion by way of intermediaries changes the perception of responsibility for unethical outcomes, both by external observers and by those directly involved. For example, players who are unfair through intermediaries are less likely to be punished than players who make unfair choices themselves (Bartling & Fischbacher, 2010; Coffman, 2010). Thus, the same acts, when carried out through intermediaries, are often perceived as less unethical than when the same acts are carried out directly.

As Hamman, Loewenstein, and Weber (2010) demonstrate, even the extent to which individuals hold *themselves* accountable for their unethical actions is mitigated when acting through intermediaries. In one set of studies, "principals" could decide which of several "agents" to hire to decide how much of a fixed sum of money given to the principal would be shared with a third party (the "recipient"). Principals tended to fire those who gave a lot and hire those who gave little, to the point at which those agents who were "employed" were those who gave almost none of the dictators' money to recipients. Motivated to be hired, the agents' choices became increasingly favorable to the dictator through time, leaving the other party almost nothing. The result was that, relative to conditions in which dictators made the sharing decision individually (i.e., not through agents), much less was shared. Moreover, despite the fact that they had systematically hired agents who acted selfishly on their behalf, dictators who made the sharing decision via agents ended up feeling less responsible for the low payoffs realized by recipients. Delegation, it seems, makes it easier to secure selfish outcomes because one does not feel as morally

accountable when someone else—even a person one has hired to do the job—is doing one's dirty work.

A similar conclusion regarding the perils of vertical diffusion of responsibility comes from an experiment by Ellman and Pezanis-Christou (2010). They considered production by two subjects, who acted as members of a firm, which yielded costs for a passive third party. Both subjects were involved in choosing a level of production for the firm, which determined their own payoff and the payoff of the third party. Ellman and Pezanis-Christou considered alternative organizational structures in which the decision is made vertically (one subject set the level of production and the other carried it out) or horizontally (both subjects jointly determined the level of production, either by reaching consensus or by having their suggested choices averaged), and they also varied the presence of communication between the two subjects.

Without communication, ethical production levels that limit the harm to the third party occurred significantly more often under a horizontal structure in which both subjects' suggestions were averaged to determine what the firm did than under vertical structures in which the responsibility was diffused (Ellman & Pezanis-Christou, 2010). However, under an alternative horizontal structure, in which firm members had to reach consensus, ethical behavior was just as low as under vertical structures—mainly because when teams compromise, they tend to choose the less ethical of the two initial suggestions. When communication between the subjects was allowed, however, the vertical structure yielded the most ethical outcomes of any organizational structure. Communication apparently provided subordinates with voice and greater responsibility in the process of selecting production levels, erasing the effects of vertical diffusion.

Although vertical diffusion led to greater selfishness in the experiments just described (when no communication was present), other research provides evidence that horizontal diffusion of responsibility can have a similar effect. Considerable research in psychology demonstrates bystander effects in which an increase in the number of potential interveners in a situation reduced the likelihood that ethical action will be taken (Darley & Latane, 1968). However, one problem with interpreting bystander effects is that they potentially confound two different phenomena. On the one hand, and consistent with the argument we have been advancing, people may not actually care much about whether the ethical action is taken if they do not feel responsible for its occurrence. On the other hand, and consistent

with a free-rider problem, people may truly care about whether the ethical action is taken but would prefer not to incur the costs of taking the action themselves if there is a good chance someone else will do so.

Dana, Weber, and Kuang (2007) isolated the first effect with an economic game. In the game, two decision makers chose between two actions that yielded a monetary payoff for them and for a passive bystander. Each decision maker could ensure a fair outcome that yielded the same payoff for everyone, while an unequitable outcome that benefitted the two decision makers and harmed the third party resulted only if both decision makers selected it. In this setup, any decision maker who was interested in the equitable outcome could guarantee its implementation, but each decision maker could evade accountability for choosing the selfish option because his or her decision only led to this outcome if the other player made the same choice. Further, because all players would receive the same amount if one chose the fair option, no free-rider problem was present; that is, there was no material advantage to letting someone else ensure equity. The fact that subjects chose the selfish option almost twice as often as players in a baseline condition who did not have the option to diffuse responsibility supports the idea that horizontal diffusion of responsibility can encourage unethical behavior by making people feel less accountable for their own behavior.

Exploiting Uncertainty

Another important factor that can influence the extent to which individuals feel accountable for unethical behaviors is the presence of uncertainty. The research we review presents several examples in which individuals rely on uncertainty, usually regarding the consequences of their actions, as a justification for behaving unethically while maintaining the perception that one is ethical. Some of the studies show that people use the veil of uncertainty to hide unethical behavior (including possibly from themselves); other studies go a step further and show that people sometimes choose to remain willfully ignorant to provide themselves with such a veil.

As mentioned, many economic experiments measuring fair or ethical behavior in distributional choices suggest that many people are motivated by a pure sense of doing what is "right" in that they give when there can be no possibility of personal gain from doing so. A commonly drawn conclusion from this research is that people (e.g., dictators in the dictator

game) genuinely value being kind to others. However, several experiments suggest that fairness motives are not so simple. For example, in experiments using the dictator game, people who would share money often kept all of it when the other party did not know a game was being played, that is, that money could have been shared (Dana, Cain, & Dawes, 2006; see also Broberg, Ellingsen, & Johannesson, 2007; Lazear, Malmendier, & Weber, 2010). Thus, when others are uncertain about our actions, we are much less likely to behave fairly, even when there is no chance at retribution (see also Andreoni & Bernheim, 2009).

A related set of experiments explored how behavior changes when decision makers are allowed to remain uncertain about the consequences of their own actions. The findings of these studies indicate that, when individuals can remain uncertain about the degree of harm imposed on others by their actions, unethical conduct becomes much easier.

In a binary version of the dictator game, Dana et al. (2007) showed that a majority of dictators preferred an equal and welfare-maximizing option to a selfish option that gave them more but also lowered the total payout to both parties. Although one might interpret this finding as showing that people care about the other party, an additional treatment added a superficial, and seemingly irrelevant, level of payoff uncertainty by obscuring the relationship between payoffs and actions. In this treatment, the decision maker's own payoffs for the two options were known and were the same as in the baseline case, but the payoffs to the other party were left uncertain. These payoffs could either be the same as in the baseline game, or reversed, so that the action that benefitted the decision maker also benefitted the recipient and maximized total payoff. The actual payoffs were determined by a coin flip, and the outcome was not initially known to the decision maker. However, the decision maker could easily reveal the true payoffs by clicking on a "reveal game" button, in which case the true payoffs would be known.

If the ethical choice in the baseline game is to select the equality option, a comparable action is easily implemented in this "hidden information" treatment: A decision maker simply has to reveal the true payoffs and select whichever is the more generous option. However, dictators exploited uncertainty about the consequences so that they could be more selfish (Dana et al., 2007). Roughly half of dictators chose not to reveal the true payoffs, ultimately leading to a majority of choices being unfair. Thus, not knowing how another is impacted by one's action, even when one can

easily resolve such uncertainty, provides the decision maker with ethical immunity to pursue self-interest while not feeling directly responsible for creating an inequitable outcome.

The finding by Dana et al. (2007) has been replicated in other experiments, which manipulated the exact payoffs and whether the default option was ignorance or information about the payoffs (Feiler, 2005; Grossman, 2010; Larson & Capra, 2009). In combination, all of these experiments suggest that people will sacrifice personal gain for the benefit of others when they know the exact consequence of their choices, but when there is even a little bit of uncertainty, which is irrelevant from a decision–theoretic point of view because it can be resolved at no cost, the same individuals will capitalize on this uncertainty about the consequences of their action—a form of ethical immunity—to pursue material gain.

Haisley and Weber (2010) provide an additional example of exploiting uncertainty to be more self-interested by showing that individuals manipulate their perceptions of ambiguity (i.e., lotteries in which there is uncertainty about the probabilities) to justify self-interested behavior. In their experiment, decision makers made a series of binary choices, with one choice option more equitable but the other (selfish) option yielded more money to the dictator and less to the recipient, on average. In case of the "self-interested" choice, the payoff to the recipient was also contingent on a lottery.

The experiment (Haisley & Weber, 2010) varied whether the lottery payment to the recipient was presented as "known" uncertainty (i.e., an explicit 0.5 chance to win) or "ambiguity" (the probability of winning could be anywhere between 0 and 1). When the probability of an adverse payout to the recipient was known, the dictator had very little room to manipulate subjective perceptions of likelihood. But, under ambiguity, the subject could plausibly convince him- or herself that the lottery faced by the recipient might be more favorable. That is, the authors hypothesized that dictators motivated to act selfishly, while wanting to believe they were not doing so, would rely on the ethical immunity provided by the ambiguous probabilities to convince themselves that what they were doing was "not that bad" for the recipient.

Consistent with the appeal of uncertainty as a means for obtaining ethical immunity, self-interested choices were more frequent under ambiguity than under known risk (Haisley & Weber, 2010). Moreover, subjects biased their subsequent (incentivized) estimates of the payoff to the

recipient upward under ambiguity, but not under known risk, indicating that they had convinced themselves under ambiguity that these lotteries were indeed more attractive for the other party.

This research suggests that environments with greater uncertainty about the outcomes that result from one's choices are likely to yield less-ethical behavior. This is important, as many real-world situations in which individuals must make ethical choices are also frequently accompanied by uncertainty (e.g., about how much another will be harmed by actions that would benefit oneself).

Seeking Justifications

People avoid responsibility for behaving generously or ethically not only by avoiding information and situations that would hold them to account. Research on self-serving biases in judgments of fairness suggests that individuals also play "mind games" with themselves that serve a similar function. While supporting the general conclusion that people have a desire to be fair, this research shows that they gravitate toward the principles of fairness that most favor themselves. A classic example comes from a study by van Avermaet (reported in Messick, 1985) in which participants were instructed to fill out questionnaires until told to stop. When they finished, they were left with money to pay themselves and another participant who had already left. Participants were told one of four things: (a) The other subject put in half as much time and completed half as many surveys, (b) the other subject put in half as much time but completed twice as many surveys, (c) the other subject had put in twice as much time but completed half as many surveys, or (d) the other subject put in twice as much time and completed twice as many surveys.

Clearly, participants' sense of ethics served as a powerful constraint on behavior: Almost no one kept all of the money, which would be unjustifiably selfish and unfair because the other participant did similar work. How they shared the money, however, provides an interesting insight into human nature. Participants who worked twice as long and completed twice as many surveys kept twice as much money on average, a simple application of a merit principle to pay. Participants kept more than half of the money, however, both in the condition where they worked longer and completed less and the condition where they completed more and did not work as long. Again, behavior was consistent with a merit principle, but the

principle chosen, on average, systematically favored the subject making the allocation. Finally, when participants completed only half as much and worked only half as long, they kept, on average, half of the money, consistent with a rule of equal division rather than merit (van Avermaet reported in Messick, 1985).

Similar results—suggesting that what constitutes a "fair" allocation is malleable toward judgments that favor the self—have obtained in numerous other experiments (Frohlich, Oppenheimer, & Kurki, 2004; Konow, 2000). What we can take away from this research is that most people are not unabashedly selfish; they have a desire to abide by some notion of what is fair or ethical. Yet, judgments of fairness can be systematically biased to favor the self. Much as people judge the actions they delegate to others or that they take under uncertainty as less diagnostic about their own ethical qualities, people judge fairness in a way that enables them to avoid the conclusion that they are seeking outcomes that are self-interested and unfair.

A series of experiments by behavioral economists (Loewenstein, Issacharoff, Camerer, & Babcock, 1992; Babcock, Loewenstein, Issacharoff, & Camerer, 1995) shows that this self-serving bias was caused by selective interpretation of information and also that it led, both in laboratory experiments and in a field study, to failures and impasse in negotiation. Simulating pretrial bargaining, Loewenstein et al. conducted bargaining experiments in which subjects were presented with case materials (depositions, police reports, etc.) from an actual lawsuit, randomly assigned to the role of either plaintiff or defendant, and asked to negotiate a settlement in the form of a payment from defendant to plaintiff. At the outset, the experimenters gave the defendants a monetary endowment to finance the settlement, and the division of the endowment the subjects agreed on through bargaining was what they took home as pay. The longer it took the parties to agree to a settlement, the more both were penalized by having the endowment of money they were dividing shrink. If they failed to settle, the defendant's payment to plaintiff, based on the smaller endowment size, was determined by a neutral judge who had reviewed all of the case materials. Before negotiating, both plaintiffs and defendants were asked to predict how the neutral judge would rule in the case and were also paid for the accuracy of this prediction.

Participants' estimates of a fair settlement were biased in a self-serving manner, leading to low settlement rates. Direct evidence that the self-serving bias played a role in this failure to settle came in the form of

the predictions of the judge's ruling. Plaintiffs' predictions of the judge's award to them were, on average, substantially higher than those of defendants, despite the facts that the estimates were secret and had no bearing on the settlement and that both parties were paid to be accurate in their estimates. Further, the larger the discrepancy between a particular plaintiff's and defendant's estimates, the lower was their likelihood of settlement; hence, they both left the experiment worse off in terms of payment. This evidence suggests that self-serving biases are unintentional; people are often unable to avoid being biased, even when it is in their best interest to do so.

In subsequent experiments employing the same paradigm (Babcock et al., 1995), settlement rates were markedly improved by assigning participants their roles only after reading the transcripts. In this way, any motivation to interpret evidence as favorable to one side over another while reading and evaluating the materials was removed. Without a self-interested conclusion to reach, interpretations of fairness, as measured by predictions of the judge's ruling, looked more like those of a neutral third party than an interested party. More interestingly, diminishing the self-serving bias in this fashion dramatically increased settlement rates. This finding suggests that self-serving biases work by way of distorting the way that people seek out and weigh information when they perceive that they have a stake in the conclusion.

Using a similar approach, Haisley and Weber (2010), in the study of ambiguity and fairness discussed previously, used individuals' own initial, and naturally negative, perceptions of ambiguity to constrain the extent to which they could perceive ambiguity as favorable. Recall that the experiment found that dictators who were motivated to view ambiguity as favorable in that it justified self-interest did so and, as a result, behaved more selfishly. One treatment simply elicited preferences for ambiguity versus simple risk from dictators, prior to their finding out that they would have an incentive to view ambiguity favorably, This mild intervention, in which dictators simply had to express a preference (against ambiguity) prior to being motivated to view it favorably, was sufficient to constrain their subsequent actions and judgments—they no longer viewed ambiguity as favorable when it could help them, and as a result, they behaved significantly less selfishly. Thus, obtaining initial "unbiased" judgments of what is fair or ethical, or even about the likely consequences of uncertain

events or processes, may be a powerful instrument for limiting the ability of individuals to deflect ethical accountability for their actions.

Sah and Loewenstein (2010) examined the rationalizations that physicians employ that enable them to accept gifts from pharmaceutical companies and medical device manufacturers, which most outsiders to the profession view as barely disguised bribes. Physicians in the early stages of their careers were asked in a survey about the acceptability of receiving various gifts from industry. Some of the physicians were first asked questions about the sacrifices they made in medical training. Others were also asked the sacrifice question and in addition were asked a question that implicitly suggested the idea that these hardships might justify acceptance of gifts. A control group was simply asked about the acceptability of receiving gifts with no prior priming questions. Reminding physicians of the sacrifices they had made in obtaining their education resulted in gifts being evaluated as more acceptable, and even though most residents disagreed with the hinted-at rationalization, exposure to it further increased the perceived acceptability of gifts. Reminders of hardships such as being overworked and having large amounts of debt apparently led physicians to conclude that they deserved gifts.

A research study using a die-under-cup paradigm further demonstrates the value decision makers place on having self-justifications for ethically questionable behavior (Shalvi, Dana, Handgraaf, & De Dreu, 2010; task adapted from Fischbacher & Heusi, 2008). Participants were asked to roll a die that was placed under a cup by shaking it. Then, through a small hole in the top of the cup, participants privately observed the result and were to be paid in dollars whatever number they said they rolled. They were further instructed to roll the die a few additional times to convince them that it was fair. This procedure made it transparent that even after the participants left, no one could observe what they rolled. The results revealed two interesting patterns. While the distribution of reports differed significantly from that of a fair die roll, the amount of apparent lying was modest, with about 37% of participants reporting that they rolled a six. Perhaps more interestingly, the distribution of reports very well fit the distribution that would be expected if one were reporting the best of three rolls, raising the possibility that people were willing to report the best roll that they had even though they were instructed that they were to report only the first roll.

To test this possibility directly, the procedure was changed in two ways (Shalvi et al., 2010). First, participants were allowed to inspect the die

prior to rolling to ensure that it was legitimate. Then, they rolled the die under the cup once and observed the result. The experimenter then passed around a box, and each participant swept the cup (and thus the die) into the box, effectively hiding their result from ever being seen. Reports from this procedure were significantly lower on average, indicating less lying for material gain. Further, the distribution of reports differed significantly from the best-of-three distribution. It seems that participants needed justifications to be able to lie, even though they were the only ones privy to both the lie and the justifications. Having the opportunity to roll the die multiple times thus benefitted the participants, who could roll until they saw a better number and then feel that their report was less of a lie.

CONCLUSIONS

In this chapter, we show that, although most people are motivated to be ethical, self-interest often manages to trump ethics. But, the unethical conduct is not simply the result of people rationally deciding to forgo ethical considerations in pursuit of self-interest. Instead, the way in which people arrive at unethical conduct is more psychologically complex. Specifically, when the context surrounding an ethical choice provided individuals with the opportunity to not have their ethical selves held to account, people often exploited this lack of accountability to behave unethically without feeling they had done so. Put differently, while we may find it desirable to behave ethically, we may find it even more desirable to behave unethically without feeling like we have done so.

Ethical Immunity in Organizations

Although ethics traps exist in most, if not all, spheres of life, they are likely to be especially common and treacherous in business settings, where people often have limited responsibility for, as well as ability to monitor the consequences of, their individual actions. Indeed, common features of business environments are likely to facilitate the use of each of the three methods of achieving ethical immunity that we discussed in this chapter.

First, decision making in most businesses tends to be diffused, both vertically and horizontally, making it easy for people to avoid accountability

for their own actions by "pointing the finger" at others. For example, the responsibility for ensuring accurate bookkeeping typically rests with unit accountants, internal auditors, the company chief financial officer (CFO), and external auditors. If all of these individuals perceive only limited involvement, or sometimes perhaps even passive involvement, then we might expect to see the kinds of unethical conduct demonstrated in experiments.

Second, decision making in business contexts is also typically characterized by high degrees of uncertainty. For example, in product safety decisions uncertainty about the degree of harm caused or the conditions under which harm may occur may provide decision makers with ethical cover. Moreover, the fact that businesses can often manipulate the kind of research conducted and disseminated regarding product safety—as the tobacco industry infamously did for decades—creates a striking parallel with those experiments in which people's ethical conduct is diminished when they have the ability to strategically fail to acquire relevant information.

Finally, business contexts may offer many of the kinds of justifications we discussed that facilitate unethical conduct. For instance, the obligation created by representing an organization or client may allow individuals to believe that the ethical act is equivalent to the one that best satisfies the firm's need for profit. Indeed, arguments by scholars—such as Milton Friedman's well-known claim that "the social responsibility of business is to increase its profits" (1970)—may present ethically conflicted managers with the necessary justification for rationalizing unethical conduct. Another important element of business contexts is competition, which may contribute to the often-used justification that people feel like they have no choice but to commit an unethical act.

Attacking Ethical Immunity Through Greater Self-Accountability

Beyond helping to predict and explain people's likelihood to behave unethically, the study of behavioral ethics can also suggest means of discouraging unethical behavior and encouraging ethical behavior. As noted in the introduction, we are skeptical of attempts to engender ethical behavior that rely on people to recognize when they face an ethical dilemma and to successfully navigate themselves around ethical traps. While normative and legal approaches may be useful in helping professionals to define ethical boundaries, they are unlikely to provide much traction when it comes to preventing those same people from crossing those boundaries.

Instead, we believe that, just as "choice architecture" (Thaler & Sunstein, 2008) can be used to design institutions that help boundedly rational people to make better decisions, institutions and information mechanisms can be designed to hold individuals accountable for their own behavior, even if only to themselves, encouraging ethical behavior by disarming mechanisms that enable ethical immunity. To enhance ethical behavior, all the research just reviewed suggests, we need to create decision-making settings that establish a kind of *self-accountability* that makes it more difficult for people to avoid being confronted with the fact that they are violating their own ethical standards.

In one of the experimental conditions of a study discussed, Haisley and Weber (2010) provided a demonstration of how greater self-accountability can enhance ethical behavior. In one condition, prior to making their choices, dictators were asked to decide whether they liked or did not like ambiguity for themselves. In this condition, the impact of ambiguity on selfish behavior disappeared. That is, by having subjects state, up front, their attitudes toward known risk versus ambiguity, they subsequently became "constrained" by these initial attitudes and were not able to self-servingly manipulate their perceptions of uncertainty.

Efforts to increase self-accountability can already be seen in a variety of professional and corporate settings, although they have generally not been interpreted in such terms. For example, new standardized conflict-of-interest disclosure forms for medical journals likewise require authors not only to list any conflicts they face but also to write affirmatively that they have no conflicts if none are listed. Increasing numbers of states are requiring unemployed persons to file frequently to renew unemployment benefits, probably because affirmatively stating that one is unemployed is more difficult for a person who has recently obtained employment than is failing to report such a change in employment status.

Closing Comments

The research reviewed in this chapter can be viewed as a glass half empty (as we have done), emphasizing that people are not as inherently ethical as prior research might suggest. However, it could also be viewed as a glass half full, in the sense of identifying a broad class of situations in which people do, in fact, behave ethically, honestly and generously. The key feature of these situations is the existence of moral accountability or, equivalently,

the lack of what Dana et al. (2007) have called "moral wiggle-room." The good news is that when people are, in effect, held to account for their behavior, all of the research suggests that they are very likely to behave in an ethical fashion.

REFERENCES

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica, 77*(5), 1607–1636.

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *American Economic Review, 85*, 1337–1342.

Bartling, B., & Fischbacher, U. (2010). *Shifting the blame: On delegation and responsibility*. Working paper.

Broberg, T., Ellingsen, T., & Johannesson, M. (2007). Is generosity involuntary? *Economics Letters, 94*(1), 32–37.

Camerer, C. (2003). *Behavioral game theory: Experiments on strategic interaction*. Princeton, NJ: Princeton University Press.

Chugh, D., Bazerman, M., & Banaji, M. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. Moore, D. Cain, G. Loewenstein, & M. Bazerman (Eds.), *Conflicts of interest: Problems and solutions from law, medicine and organizational settings*. London: Cambridge University Press.

Coffman, L. (2010). *Intermediation reduces punishment (and reward)*. Working paper.

Dana, J., Cain, D. M., & Dawes, R. (2006). What you don't know won't hurt me: Costly (but quiet) exit in a dictator game. *Organizational Behavior and Human Decision Processes, 100*, 193–201.

Dana, J., Weber, R., & Kuang, J. (2007). Exploiting moral wriggle room: Behavior inconsistent with a preference for fair outcomes. *Economic Theory, 1*, 67–80.

Darley, J., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377–383.

De Cremer, D. (2009). Being unethical or becoming unethical: An introduction. In D. De Cremer (Ed.), *Psychological perspectives on ethical behavior and decision making* (pp. 3–13). Greenwich, CT: Information Age.

Ellman, M., & Pezanis-Christou, P. (2010). Organisational structure, communication and group ethics. *American Economic Review*, *100*, 2478–2491.

Falk, A., & Heckman, J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science, 326*, 535–538.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137–140.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics, 114*, 817–868.

Feiler, L. (2005). *Patterns of information avoidance in binary choice dictator games*. Working paper.

Fischbacher, U., & Heusi, F. (2008). *Lies in disguise, an experimental study on cheating.* TWI Working Paper 40, Thurgau Institute of Economics, University of Konstanz, Germany.

Forsythe, R., Horowitz, J., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior, 6*, 347–69.

Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. *The New York Times Magazine*.

Frohlich, N., Oppenheimer, J., & Kurki, A. (2004). Modeling other-regarding preferences and an experiment. *Public Choice, 119*, 91–117.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review, 95,* 384–394.

Grossman, Z. (2010). *Understanding strategic ignorance*. Working paper.

Haisley, E., & Weber, R. (2010). Self-serving interpretations of ambiguity and other-regarding behavior. *Games and Economic Behavior, 68*, 614–625.

Hamman, J. R., Loewenstein, G., & Weber, R. A. 2010. Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, *100*(4): 1826–46.

Konow, J. (2000). Fair shares: accountability and cognitive dissonance in allocation decisions. *American Economic Review, 90*, 1072–1091.

Larson, T., & Capra, C. M. (2009). Exploiting moral wiggle room: Illusory preference for fairness? A comment. *Judgment and Decision Making, 4*(6), 467–474.

Lazear, E. P., Malmendier, U., & Weber, R. A. (2010). *Sorting, prices, and social preferences.* Working paper.

Loewenstein, G. (1996). Behavioral decision theory and business ethics: Skewed tradeoffs between self and other. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of conduct: Behavioral research into business ethics* (pp. 214–227). New York: Russell Sage Foundation.

Loewenstein, G., Issacharoff, S., Camerer, C., & Babcock, L. (1992). Self-serving assessments of fairness and pretrial bargaining. *Journal of Legal Studies, 12*, 135–159.

Loewenstein, G., Thompson, L., & Bazerman, M. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology, 57*, 426–441.

Messick, D. (1985). Social interdependence and decisionmaking. In G. Wright (Ed.), *Behavioral decision making* (pp. 87–109). New York: Plenum.

Sah, S., & Loewenstein, G. (2010). Effect of reminders of personal sacrifice and suggested rationalizations on residents' self-reported willingness to accept gifts: A randomized trial. *Journal of the American Medical Association, 304*(11), 1204–1211.

Shalvi, S., Dana, J., Handgraaf, M. J. J., & De Dreu, C. K. W. (2010). *Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior*. Working paper.

Shalvi, S., Handgraaf, M. J. J., & De Dreu, C. K. W. (in press). Ethical maneuvering: Why people avoid both major and minor lies. *British Journal of Management*. doi: 10.1111/j.1467–8551.2010.00709

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research, 17*, 223–236.

Thaler, R., & Sunstein, C. (2008). *Nudge: improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.