Journal of Experimental Social Psychology 45 (2009) 708–719

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior

Francesca Gino [a,*], Max H. Bazerman [b]

[a] Kenan-Flager Business School, University of North Carolina, Campus Box 3490, McColl Bld 4725, Chapel Hill, NC 27516, United States
[b] Harvard Business School, Harvard University, Baker Library 453, Soldiers Field Road, Boston, MA 02163, United States

ARTICLE INFO

Article history:
Received 20 June 2008
Revised 27 March 2009
Available online 12 April 2009

Keywords:
Unethical behavior
Implicit biases
Gradual changes
Ethical decision-making
Whistle blowing

ABSTRACT

Four laboratory studies show that people are more likely to accept others' unethical behavior when ethical degradation occurs slowly rather than in one abrupt shift. Participants served in the role of watchdogs charged with catching instances of cheating. The watchdogs in our studies were less likely to criticize the actions of others when their behavior eroded gradually, over time, rather than in one abrupt shift. We refer to this phenomenon as the *slippery-slope effect*. Our studies also demonstrate that at least part of this effect can be attributed to implicit biases that result in a failure to notice ethical erosion when it occurs slowly. Broadly, our studies provide evidence as to when and why people accept cheating by others and examine the conditions under which the slippery-slope effect occurs.

Introduction

Companies such as Enron, Tyco, Parmalat, and WorldCom are often cited as examples of disasters that resulted from unethical management behavior. Both the popular press and academic studies have noted that managers and leaders in modern organizations either lack strong ethical standards (Andrews, 1989; Longnecker, 1985; Molander, 1987; Pitt & Abratt, 1986) or are willing to abandon them in the face of economic incentives or competitive pressures (Gellerman, 1986; Hosmer, 1987). For instance, while accounting firms in this country are charged with providing independent financial statements and reporting client mismanagement, they do so at the risk of displeasing their clients and losing lucrative service contracts (Moore, Tetlock, Tanlu, & Bazerman, 2006). As a result, "independent" auditors often have been found to be complicit in their clients' unethical practices (Levitt & Dwyer, 2002).

Of course, some individuals who observe the unethical behaviors of others do speak up and report ethical misconduct. For instance, Jeffrey Wigand publicly revealed that executives of US tobacco companies knew that cigarettes were addictive when they approved the addition of known carcinogenic ingredients to cigarettes (Mollenkamp, Levy, Menn, & Rothfeder, 1998). Similarly, Enron employee Sherron Watkins exposed former Enron chairman and CEO Kenneth Lay's questionable accounting practices in 2002.

Under what conditions are people likely to ignore the unethical behavior of others? This paper addresses this question by examining the predictable conditions that affect whether or not observers of unethical actions speak up and the underlying mechanisms that explain the decision to speak up. Specifically, we focus on gradual erosion in others' unethical behavior and investigate its influence on observers' tendency to accept such behavior.

We present the results of four experiments in which we vary the process of deterioration of others' behavior – gradually, through small changes, or in a single, abrupt shift. We predict and show that people are more likely to accept the unethical behavior of others if the behavior develops gradually (along a slippery slope) rather than occurring abruptly. The studies provide the first empirical investigation of what we call the "slippery-slope effect" within the context of ethical judgment. Our studies also examine why the phenomenon occurs. We show that implicit biases explain at least part of the tendency to ignore others' unethical actions that occur on a slippery slope.

Over the last three decades, scholars have used different terms to describe psychological phenomena that can occur without an individual's awareness or intention, including "automatic", "unconscious", "implicit", or "spontaneous" (Blair, 2001). While phenomena that differ along these dimensions can be clearly distinguished on a conceptual level (Bargh, 1994), in this paper we use the term "implicit" to describe phenomena that occur without awareness or intention, and we use the term "explicit" to describe phenomena that occur with awareness and intention. We demonstrate that observers are not fully aware of their ethical lapses when ethical erosion is gradual, and we compare their behavior

* Corresponding author.
E-mail address: fgino@unc.edu (F. Gino).

0022-1031/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2009.03.013

to the case in which ethical erosion occurs as an abrupt change. Thus, the slippery-slope effect is, in part, a psychological phenomenon that can occur without an individual's awareness or intention – i.e., it is subject to implicit biases.

Defining unethical behavior

Numerous definitions of unethical behavior exist in the literature (e.g., Brass, Butterfield, & Skaggs, 1998; Treviño, Weaver, & Reynolds, 2006). For our current purposes, we rely on Jones' (1991) broad conceptualization of unethical behavior as reflecting any action that is "either illegal or morally unacceptable to the larger community" (p. 367). Examples include violations of ethical norms or standards (whether they are legal standards or not), stealing, cheating, lying, or other forms of dishonesty. In our studies, unethical behavior occurs when participants accept exaggerated estimates made by others and increase their payoff by doing so. Consistent with this view of unethical behavior, Reynolds (2006) has found that both the presence of harm and the violation of a behavioral norm are positively (and distinctively) associated with moral awareness. For now, we avoid stating whether intentionality is a requirement for a behavior to be unethical, but return to this issue when discussing our hypotheses.

The slippery-slope effect in ethical judgment

The question of when people report others' ethical misconduct is addressed by research on organizational whistle-blowing (e.g., King, 1997; Miceli, Dozier, & Near, 1991; Near & Miceli, 1986). Whistle-blowing refers to the disclosure by organizational members of wrongdoing within an organization to the public or to those in positions of authority inside the organization (Near & Miceli, 1985). The literature on whistle-blowing aims to explain individual reactions to acts of perceived unethical behavior (Gundlach, Douglas, & Martinko, 2003) and has identified individual, organizational, and situational factors that affect such reactions (Near & Miceli, 1995, 1996).

At the individual level, religious and ideal values (Miceli & Near, 1992; Sims & Keenan, 1998), moral standards and judgment (Miceli et al., 1991), power and justice (Near, Dworkin, & Miceli, 1993), and individual locus of control (Miceli & Near, 1992) have all been shown to affect the tendency to blow the whistle. Scholars have also investigated individual differences between employees who do or would report colleagues' unethical behavior and inactive observers (e.g., Sims & Keenan, 1998). Whistleblowers tend to have higher levels of pay, seniority, and education than inactive observers (Miceli & Near, 1984). Furthermore, whistleblowers typically have more years of service and higher levels of professional status than inactive observers (Miceli & Near, 1988).

In addition, previous research has identified various organizational (e.g., whistle-blowing culture) and situational (e.g., nature of the wrongdoing) factors that affect whistle-blowing. For example, people are more likely to report wrongdoing when the organizational climate supports whistle-blowing (Near & Miceli, 1995) or when the organization is perceived to be responsive to complaints (Miceli & Near, 1988). In addition, Greenberger, Miceli, and Cohen (1987) have suggested that whistle-blowing is less likely when wrongdoing is ambiguous or when group members must assist each other with the task in question.

We argue that one unexamined, critical factor that could affect when people report others' ethical misconduct is the nature of the wrongdoers' behavior. Here, we focus on a particular aspect of others' unethical behavior: how such behavior evolves over time. We consider two types of change in others' ethicality. First, ethical behavior might change abruptly, thus providing a clear signal to observers that an actor crossed ethical boundaries. Second, ethical behavior might erode gradually over time, thus weakening the signal to observers. We suggest that these two types of changes in the erosion of others' ethicality differentially influence observers' likelihood to report unethical behavior.

For an illustration of this argument, imagine that an accountant employed by a well-known accounting firm is in charge of the audit of a large corporation with a strong reputation. The accounting firm and the client have an excellent relationship, and the accounting firm receives tens of millions of dollars in fees from the client each year. For 3 years, the accountant in charge of the audit viewed and approved the client's high-quality, ethical financial statements. Suddenly, in the fourth year, the corporation stretches the limits of the law, and even breaks it in certain areas. Would the accountant sign a statement certifying that the financial statements were acceptable according to government regulations? Now suppose that the accountant saw and approved of the company's high-quality, highly ethical financial statements for 1 year. The following year, the corporation begins stretching the law in a few areas, without appearing to break the law. In the third year, the firm stretches the ethicality of its returns a bit more. Some of the company's accounting decisions now possibly violate federal accounting standards. By the fourth year, the corporation stretches the limits of the law in more areas and occasionally breaks the law.

We believe that the accountant is much less likely to refuse to sign the statements in the second version than in the first version of this example, even if the unethical behavior is identical in year four of both stories. There are several reasons this might be true. First, the "induction mechanism" might be at play (Tenbrunsel & Messick, 2004). The induction mechanism argues that people use past actions as benchmarks when they evaluate new actions. Thus, if the behavior of others we observed in the past was ethical and acceptable, then similar or slightly different behaviors are also ethical and acceptable. The induction mechanism is comparable to the process of routinization, which suggests that once a practice becomes a routine, people consider it to be acceptable and ordinary (Kelman & Hamilton, 1989; Tenbrunsel & Messick, 2004). Similarly, unethical acts might become routine if they develop gradually over time. According to Ashforth and Anand (2003), unethical acts can "become an integral part of day-to-day activities to such an extent that individuals may be unable to see the inappropriateness of their behaviors" (p. 4). In addition, Brief, Buttram, and Dukerich (2001) describe institutionalization of corruption as occurring in part due to "repetitive cycles of routinized activity" (p. 480). While this research remains silent on whether or not these processes occur intentionally and does not provide empirical evidence for their existence, it does suggest that gradual changes in the unethical behavior of others might influence the likelihood that observers will accept such behavior compared to the case of abrupt changes. Building on this work, we predict a slippery-slope effect in ethical judgment by hypothesizing that:

Hypothesis 1. People are more likely to accept the unethical behavior of others if the unethical behavior develops gradually over time than if the unethical behavior occurs abruptly.

The role of implicit biases in explaining the slippery-slope effect

One premise of the literature on whistle-blowing is that both whistleblowers and inactive observers do in fact notice others' unethical behavior and perceive it as unacceptable. In addition, most of the ethics literature focuses on intentional and conscious choices by actors. By this logic, if observers are intentionally tolerating behavior that they view to be unethical, these observers are themselves acting unethically (Miceli, Near, & Schwenk, 1991). An example of this behavior is the case of auditors who knowingly fail to report malfeasance by their clients. Thus, this body of research suggests that when seeing evil, individuals make an explicit decision regarding whether to speak up or not by weighing the potential costs and benefits of their choice. This perspective is consistent with rational crime theory (Allingham & Sandmo, 1972; Becker, 1968), which is often used as a possible explanation for why individuals engage in unethical behavior. According to this framework, when faced with an opportunity to behave dishonestly, an individual engages in a cost–benefit calculation, which determines his/her ultimate decision about whether or not to behave dishonestly (Hill & Kochendorfer, 1969; Leming, 1980; Michaels & Miethe, 1989; Steininger, Johnson, & Kirts, 1964; Tittle & Rowe, 1973; Vitro & Schoer, 1972).

By contrast, other scholars have suggested that our ethical decisions are compromised by the same systematic errors that characterize our thinking and decision-making processes overall (Messick & Bazerman, 1996), some of which are beyond our conscious awareness (Chugh, Bazerman, & Banaji, 2005). According to this view, unethical actions result from unintentional behavior that is fueled by implicit biases and automatic cognition (Kern & Chugh, in press). Banaji, Bazerman, and Chugh (2003) use the term "bounded ethicality" to refer to individuals' limits in recognizing the ethical challenge inherent in a situation or decision. Their research suggests that much unethical behavior is unintentional and inconsistent with the preferred action that someone would engage in if he/she had greater awareness of his/her own behavior (Banaji et al., 2003; Bazerman & Banaji, 2004). According to this perspective, most people may believe that their behavior is ethical and unbiased; at the same time, most also fall prey to unconscious biases in ethically relevant domains—for instance, succumbing to implicit forms of prejudice, in-group favoritism, conflicts of interest, and the tendency to over-claim credit (Banaji et al., 2003).

We build on this research and argue that implicit biases play an important role in explaining the failure to report or react to the unethical behavior of others. People tend to view themselves as moral and honest (e.g., Messick & Bazerman, 1996; Tenbrunsel, 1998), competent (e.g., Messick, Bloom, Boldizar, & Samuelson, 1985), and deserving (e.g., Ross & Sicoly, 1979; Taylor, 1989). This positive self-view reduces individuals' ability to see and recognize ethical challenges, especially when unethicality is ambiguous. Here, we focus on the ambiguous case of individuals who observe the ethicality of others eroding gradually. Prior work on visual perception has shown that people frequently fail to notice gradual changes that occur right in front of their eyes (Simons, 2000). In Simons' research, the information people miss is visual, and the mental processes that might explain the failure to notice changes are perceptual. We suggest that similar mechanisms operate when ethical erosion occurs on a slippery slope. Thus, the information might not be visual and the processes might not be perceptual. Specifically, due to implicit biases, individuals may fail to see wrongdoing that occurs in front of them. This reasoning leads us to hypothesize that:

Hypothesis 2. People's propensity to accept others' unethical behavior when ethical erosion occurs gradually is partially the result of implicit biases that prevent observers from noticing the erosion.

Study 1

Our first study tests Hypothesis 1, which predicts that people are more likely to accept others' ethical misconduct when it increases gradually rather than abruptly.

Method

Participants

Four experimental sessions were conducted in the computer laboratory of a university in the Northeastern United States. Seventy-six individuals (39% male, M_{age} = 28, SD = 10.63) participated in the experiment. Most participants (70%) were students from local universities. They were recruited using ads in which they were offered money to participate in an experiment on decision-making.

Design and procedure

The study included one between-subjects factor, which manipulated the type of change in the inflated estimates of others that participants evaluated. The procedure was identical across each experimental session, and participants were randomly assigned to one of two conditions: the *abrupt-change* condition (N = 38) or the slippery-slope condition (N = 38). Each experimental session was conducted on computers. Participants received a $10 show-up fee and had the opportunity to win up to an additional $25 during the experiment. During recruitment and again in the experiment's instructions, participants were told that their earnings would be a function of their choices during the experimental session. Participants were not given any feedback on their decisions until after the experiment.

Upon arrival at the computer laboratory, participants were registered and randomly assigned to terminals. The room was arranged so that participants could not see each other during the experiment. Participants received copies of the instructions, which told them how to play the experiment using the computer. The researcher also read the instructions aloud, and the participants were given an opportunity to ask questions. Thus, before the session began, participants knew what each phase entailed and how payoffs would be computed.

The experiment consisted of a few practice rounds followed by three phases of 16 rounds. In the practice rounds, participants were shown a series of three pictures of jars containing different amounts of American pennies ($9.75, $10.50, and $11.25) on their computer screens and were told the true amount of money in the jars. This allowed them to get a sense of the pictures they would see during the experiment. In addition, given that prior work has shown that people tend to underestimate the amount of money contained in jars (see, for instance, Bazerman & Samuelson, 1983), the three practice rounds were designed to reduce this tendency toward underestimation.

In each round of the actual experiment, participants saw a picture on the computer screen of a jar containing pennies. Pictures differed in each round. In many rounds, pictures of the jar containing different amounts of money were presented. In addition, pictures of jars containing the same amount of money were presented, but the pictures were different. Thus, to avoid an experience confound, if in Round 3 of both Phases 2 and 3 the jar contained $10.04, a different picture was used in each phase. (In such cases, the jar was shaken between photographs to change the position of the coins.) To avoid biases in estimates due to the shape of the jar, the same jar was used in all rounds.

In Phase 1, participants played the role of "estimator"; in both Phases 2 and 3, they played the role of "approver". When playing estimators, for each picture viewed in Phase 1, they were asked to guess the amount of money contained in the jar within the range of $0.00–18.00. When playing approvers, they instead were asked to either accept or reject an estimate of the amount of money contained in the pictured jar; the estimate was shown below the picture. If they wanted to accept the estimate, they first had to sign a document stating that they honestly believed the estimate was accurate within a range of 10%. They could sign the document by pressing the "SIGN" button pictured on the computer screen; after pressing it, their signature would appear on the document. If they wanted to reject the estimate, they had to press the "NO" button pictured on the computer screen. (We used this signing process to make the ethicality of participants' approval more vivid to them.)

In Phase 2, the estimates came from the experimenter. In Phase 3, the estimates were those of another participant who had been shown the same picture; specifically, the estimate was randomly selected from among the estimates provided by participants in Phase 1 (during the same experimental session). When playing Phases 2 and 3, participants did not know whether the estimates came from the experimenter or from another participant.

As Table 1 shows, in Phase 2, in the abrupt-change condition, the estimate equaled the true value of the money in the jar, which was $10 plus random error for the first 10 rounds. At Round 11, while the value of the money in the jars remained at $10 plus random error, the estimates jumped to $14 plus random error and stayed at that level for the remaining five rounds. By contrast, in the slippery-slope condition, the estimates from the abrupt-change condition were increased by 40 cents in Round 2, 80 cents in Round 3, $1.20 in Round 4, and so on. Then, from Rounds 11–16, the estimates were identical to the abrupt-change condition. The true amount of money contained in the jar only varied as a function of the small random error. We added the random error because round estimates might have made participants suspicious or cognizant that the estimates were provided by the experimenter.

At the end of the experiment, participants were asked to complete a questionnaire that asked their gender, status, and age. They then were paid their earnings, debriefed, thanked, and dismissed.

Participants' payoff

In both conditions, participants' payoffs were computed based on the roles they played in each of the 16 rounds as follows. When playing the role of estimator, each participant received 8% of her estimate every time the approver accepted it. Thus, the estimator had an incentive to provide a high estimate in Phase 1. Yet if the approver rejected the estimate, the estimator would receive no payment. When a participant played the role of approver, the payoff depended on whether the approver accepted or rejected the estimator's estimate. If she rejected it, she would receive no payment. If she approved it, she would get 4% of the estimate. However, 10% of the time (randomly determined), a third party double-checked the estimate. This check was automatically made by the computer. In that case, the approver would incur a penalty of $5 if she approved an estimate that was not within 10% of the true value. All these payoffs were common knowledge.

Note that the approver's payoff schedule includes a conflict of interest. Approvers were paid more when they allowed inflated estimates, but at the risk of a potential penalty for approving egregiously exaggerated estimates. This payoff scheme was designed to parallel the payoffs faced by regulators or watchdogs, such as auditors charged with verifying that organizations report truthful information to the public. These watchdogs typically have a financial incentive to please those they are supposed to be monitoring, which might mean turning a blind eye to small infractions. However, watchdogs face a substantial penalty if they are caught allowing a great deal of cheating.

Participants were told that their payoffs at the end of the experiment were computed as the sum of their performance in Phase 1 and their performance in either Phase 2 or Phase 3. Whether they were paid based on either Phase 2 or Phase 3 was predetermined by the experimenter but unknown to participants.

Dependent measure

Our hypotheses were tested on the data collected in Phase 2, in which our manipulation occurred. Phases 1 and 3 were used to avoid deceiving participants and to allow us to manipulate the estimates provided in Phase 2. Indeed, we wanted participants to believe that they might be judging estimates made by other participants. Thus, the experiment might have consisted only of Phase 2 had we told participants that the estimates they were asked to approve came from other participants when shown the same pictured jar. However, this would have involved deception, which we wanted to avoid.

Our dependent measure is the number of approvals (i.e., "yeses) of the estimator's estimate over Rounds 11–16. As mentioned above, in the abrupt-change treatment, the abrupt change happened in Round 11. In subsequent rounds, participants in both conditions were asked to review the same estimate (see Table 1, estimate 11 through estimate 16). To test our first hypothesis, we compared the cumulative number of approvals in Rounds 11–16 in the two conditions. Based on Hypothesis 1, we expected to find a statistically significant difference for this measure, with a significantly higher number of approvals in the slippery-slope condition than in the abrupt-change condition.

Results

Rate of approval

We computed the percentage of participants approving the estimator's estimates in Phase 2 within each condition. As shown in Table 2, the percentage of "yes"es was higher in the slippery-slope condition than in the abrupt-change condition in each round following the abrupt change, i.e., Rounds 11–16 (all *p*-values are lower than 0.05). These results are consistent with our first hypothesis, suggesting that people are more likely to accept the unethical behavior of others if the unethical behavior develops gradually over time than if it occurs abruptly.

At the aggregate level (over Rounds 11–16), the mean rate of approval was 52% in the slippery-slope condition and 24% in the abrupt-change condition. On average, in the slippery-slope condition, participants approved the estimate 3.11 times over the last six rounds of Phase 2 ($SD = 2.18$). In the abrupt-change condition, the number of approvals for each participant over Rounds 11–16 was significantly lower ($M = 1.45$, $SD = 1.64$; $t\ [74] = 3.75$, $p < 0.001$), thus supporting our first hypothesis.

Table 1
Estimates provided by the experimenter in each condition and true values.

Estimate	Slippery-slope condition ($)	Abrupt-change condition ($)	Randomly generated error ($)	True values ($)
1	10.01	10.01	+0.01	10.01
2	10.37	9.97	−0.03	9.97
3	10.84	10.04	+0.04	10.04
4	11.21	10.01	+0.01	10.01
5	11.58	9.98	−0.02	9.98
6	12.00	10.00	+0.00	10.00
7	12.44	10.04	+0.04	10.04
8	12.79	9.99	−0.01	9.99
9	13.22	10.02	+0.02	10.02
10	13.59	9.99	−0.01	9.99
11	14.02	14.02	+0.02	10.02
12	14.01	14.01	+0.01	10.01
13	13.98	13.98	−0.02	9.98
14	13.99	13.99	−0.01	9.99
15	14.02	14.02	+0.02	10.02
16	14.03	14.03	+0.03	10.03

Table 2
Rate of approval in Rounds 11–16 for each condition, Study 1.

Round	Slippery-slope condition (%)	Abrupt-change condition (%)	*p*-Values
11	26	3	0.003
12	61	24	0.001
13	53	29	0.036
14	61	32	0.011
15	50	21	0.008
16	61	37	0.039

The impact of Phase 1 estimates on Phase 2 approvals

To examine whether approval rate in Phase 2 was affected by the estimates participants provided in Phase 1, we used the following hierarchical model: $PH2\_APPROVAL_{ij} = \alpha_0 + \beta_1 \; TREAT_i + \beta_2 \; PH1\_EST_{ij} + subjects' \; RANDOM \; EFFECTS + \varepsilon$, where index i refers to participants and index j refers to rounds. We only considered Rounds 11 through 16. The dependent variable was approval (yes = 1, no = 0) for each participant and for each considered round in Phase 2. Explanatory variables were: (i) a dummy variable indicating the experimental condition ($TREAT_i$), and (ii) the estimates participants provided in Phase 1 ($PH1\_EST_{ij}$). The parameter β_1 measured the effect of our manipulation on the propensity of participants to approve estimates in Phase 2, while the parameter β_2 measured the effect of the estimates participants provided in Phase 1 on approval in Phase 2.

This analysis reveals that participants were more likely to approve an estimate in Phase 2 when they overestimated the amount of money contained in the jar in Phase 1 ($\beta_2 = 0.04$, $t = 4.66$, $p < 0.001$). Yet, even after controlling for this effect, the effect of our manipulation was still in the predicted direction and statistically significant ($\beta_1 = 0.27$, $t = 6.43$, $p < 0.001$): participants were more likely to approve an estimate in Phase 2 when playing in the slippery-slope condition than when playing in the abrupt-change condition.

Post-test. We conducted a study on a non-overlapping sample of participants to document that individuals do see misrepresentation of others in our experimental context as an ethical violation, as opposed to acceptable behavior in a laboratory setting. During Rounds 11–16 in Phase 2 of the main experiment, participants were asked to review inaccurate estimates. Did they perceive inaccurate or exaggerated estimates to be unethical? To verify that they did not have difficulty determining the accuracy and ethicality of these estimates, a manipulation check was needed. We did not include a manipulation check in the main study for fear that it would be too intrusive and arouse suspicion about the experiment's purpose. Instead, we ran a manipulation check on a separate, non-overlapping sample of participants. We recruited 36 individuals (45% male, $M = 27$, $SD = 8.50$) to complete a 10-min study for which they were paid \$5. These participants were recruited from the same pool of participants used for Study 1. The pilot study consisted of a short paper survey that included the same pictures we used in Round 1 and Rounds 11–16 of the main study. Specifically, participants were asked to look at seven different photographs of jars with pennies in them. Each jar was presented on a different page. The first page after the instructions included an example of a pictured jar with an estimate of the true amount of money contained in the jar. This page was included so that participants had a sense of how much money corresponded to a certain level of coins in the jar.

On each page of the survey after the example, participants saw a picture and, under the picture, an estimate of the money contained in the jar. Participants were told that the estimates were provided by individuals who participated in another study in which they were asked to guess the amount of money contained in the jar and were paid based on the accuracy of their estimates. On each page, we also asked participants to indicate how appropriate, acceptable, and ethical the estimate was along a 7-point scale (ratings ranged from 1 = not acceptable [appropriate, ethical] at all, to 7 = very acceptable [appropriate, ethical]). Finally, we asked them to indicate how much they thought the estimate differed from accuracy; possible answers were: within 75% of accuracy, 50%, 25%, 10%, 5%, 3%, and 1%. (The answer "within 75%" refers to an estimate that is not very accurate, while the answer "within 1%" refers to a very accurate estimate.) We found that these four scales were highly and significantly correlated. For our purposes, the most important correlations are between accuracy and ethicality, between appropriateness and acceptability, and between acceptability and ethicality. The higher the perceived inaccuracy of the estimate, the more unethical participants considered it to be ($r = 0.43$, $p < 0.001$). The more appropriate participants considered the estimate to be, the more acceptable they rated it ($r = 0.86$, $p < 0.001$). Finally, the higher the perceived acceptability of the estimate, the more ethical they considered it to be ($r = 0.73$, $p < 0.001$). Given these correlations, we created just one measure for the perceived ethicality of the estimates. The 4-item scale was reliable (Cronbach's alpha was 0.87). Overall, we found that the use of exaggerated estimates was a successful manipulation for unethical behavior, as participants perceived inaccurate estimates to be unethical.

Discussion

The results of Study 1 are consistent with Hypothesis 1: people are significantly more likely to accept the unethical behavior of others if the unethical behavior develops gradually than if it occurs abruptly. We should note that by the end of Phase 2, participants in the abrupt-change condition approved 37% of the estimates. While this percentage is significantly lower than the one in the slippery-slope condition (consistent with our prediction), it is also significantly greater than zero. This result suggests that individuals in the abrupt-change condition may adapt to unethical behavior and go along with it even if they rejected it initially. Alternatively, participants might have second-guessed their judgments as they continued to receive estimates that they had initially judged as inflated.

Study 2

Our first study demonstrated how the slippery slope phenomenon affects the acceptance of the unethical behavior of others. One limitation of Study 1 is that we do not know if participants believed their approval decisions or not. In other words, we do not know whether they actually changed their opinion of the amount of money in the jars or whether their estimates were (in the aggregate) unbiased, and they simply approved estimates that they perceived to be biased. Our second study examines this issue and investigates whether the type of change in the estimates received from others (abrupt vs. gradual) results in a change in participants' estimates of the amount of money in pictured jars.

The second experiment differs from the first in the number of rounds participants played in the role of approvers. While participants in Study 1 played approvers in all of the rounds of Phase 2, Study 2 participants were first exposed to the process of degradation of the other party's estimates and then were asked to estimate the amount of money contained in the jar (with payment for accuracy) for Rounds 13–16. This estimation task allowed us to test whether or not participants' judgment was distorted by the process they went through in the first 12 rounds (either abrupt change or slippery slope). We expected the estimates of participants in the slippery-slope condition to be significantly higher than those of participants in the abrupt-change condition.

Method

Participants

As in Study 1, we recruited participants through ads offering money to participate in an experiment on decision-making. Seventy-four individuals (42% male; $M_{age} = 26$, $SD = 9.86$) agreed to participate. Most participants (78%) were students from local universities in a city in the Northeastern United States.

Design and procedure

Study 2 followed the same procedure used in Study 1 with only one difference: while in Study 1, participants played the role of approver in all 16 rounds of Phase 2, in Study 2 they played the role of approver in the first 12 rounds of Phase 2 and the role of estimator in Rounds 13–16. With this new design, we can compare the estimates of participants in Rounds 13–16 of Phase 2 between conditions to test whether or not the slippery slope creates a greater change in the participants' best estimate of the money in the pictured jars.

Payment

As in Study 1, participants received $10 as a show-up fee and also had the opportunity to win up to $25 during the experiment. The payoff structure was the same as in Study 1 except for Rounds 13–16 of Phase 2. Because participants played the role of estimators during those rounds, they received 8% of their estimate each time it fell within 10% of the true amount of money in the pictured jar.

Dependent measure

As in Study 1, we tested our hypothesis on the data collected in Phase 2, in which our manipulation occurred. In Study 2, we were interested in two dependent measures. The first is the number of approvals of the estimator's estimates in Rounds 11 and 12. Indeed, as in Study 1, the abrupt change occurred in Round 11; the estimates that approvers received in Rounds 11 and 12 were the same across conditions. Thus, to find further support for the hypothesis that people are less likely to act unethically by accepting the incorrect assessments of others when the behavior of others erodes slowly rather than in one abrupt shift, we compared the cumulative number of approvals in Rounds 11 and 12 in the two conditions. Similar to our prediction in Study 1, we expected to find a statistically significant difference for such measure, with a higher number of approvals in the slippery-slope condition than in the abrupt-change condition.

Our second dependent measure is given by the estimates participants provided over Rounds 13–16 while playing the role of estimators. To test the hypothesis that the slippery slope influences participants' honest estimate of the amount of money in the pictured jars, we compared the average estimate in Rounds 13–16 in the two conditions. We expected to find higher estimates in the slippery-slope condition than in the abrupt-change condition.

Results

Verification process

We first replicated the analysis conducted on the data from Study 1 by computing the rate of approvals in Phase 2 within each condition for Rounds 11 and 12. As shown in Table 3, the percentage of "yes"es was higher in the slippery-slope condition than in the abrupt-change condition in each round following the abrupt change (both p-values <0.05). At the aggregate level (over Rounds 11 and 12), the mean rate of approval was higher in the slippery-slope condition than in the abrupt-change condition (43% vs. 12%, χ^2 [1, $N = 200$] = 24.10, $p < 0.001$). These results replicate the findings of Study 1 and provide further support for Hypothesis 1.

Table 3
Rate of approval in Rounds 11 and 12 for each condition, Study 2.

Round	Slippery-slope condition (%)	Abrupt-change condition (%)	p-Values
11	27	8	0.032
12	59	16	<0.001

Estimation process

At the aggregate level, over Rounds 13–16, the mean estimate per participant was significantly higher in the slippery-slope condition ($M = 11.14$, $SD = 1.32$) than in the abrupt-change condition, ($M = 10.23$, $SD = 0.74$), t (72) = 3.65, $p = 0.001$. As shown in Table 4, there were significant differences across conditions in the value for the estimates provided by participants in each round over Rounds 13–16 (all p-values lower than 0.05). In addition, it is useful to note that slippery-slope participants deviated from the $10 benchmark by roughly four times as much as participants in the abrupt-change condition. On the other hand, their estimates are far from the $14 benchmark of the estimates that they were asked to approve in Rounds 11 and 12.

The impact of Phase 1 estimates on Phase 2 approvals

To examine whether approval rate in Phase 2 was affected by the estimates that participants provided in Phase 1, we conducted the same mixed-model analysis we ran for Study 1. We estimated the following hierarchical model: $PH2\_APPROVAL_{ij} = \alpha_0 + \beta_1\ TREAT_i + \beta_2\ PH1\_EST_{ij} + subjects'\ RANDOM\ EFFECTS + \varepsilon$, where index i refer to participants and index j refers to rounds. We only considered Rounds 11 and 12, as Study 2 participants played the role of estimators over Rounds 13–16 of Phase 2. The effect of the estimates that participants provided in Phase 1 is positive: participants were more likely to approve an estimate in Phase 2 when they overestimated the amount of money contained in the jar in Phase 1 ($\beta_2 = 0.04$, $t = 2.60$, $p = 0.01$). Yet, even after controlling for this effect, the effect of our manipulation is still in the expected direction and statistically significant ($\beta_1 = 0.28$, $t = 4.19$, $p < 0.001$); participants were more likely to approve an estimate in Phase 2 when playing in the slippery-slope condition than when playing in the abrupt-change condition.

Discussion

Consistent with our first hypothesis, the results of Study 2 show that people are more likely to accept the incorrect assessments of others when the behavior of others erodes slowly rather than when it erodes in one abrupt shift. Study 2 also demonstrates that the process that participants went through in the first 12 rounds (reviewing estimates that changed in a slippery-slope manner vs. a single abrupt change) significantly affected their estimates of the amount of money contained in pictured jars in subsequent rounds. Our results are consistent with the hypothesis that the slippery slope leads to some implicit change in participants' estimate of the true value of the commodity. But, the results also suggest that they do not come to fully believe the highly biased estimates that they are asked to approve. Indeed, in the slippery-slope condition, participants' estimates in Rounds 13–16 were above the real value of the money in the pictured jars and they were greater than the estimates of participants in the abrupt-change condition, but they were also lower than the estimates participants saw in Rounds 11 and 12. Thus, an alternative explanation for the findings of Study 2 is that participants in the slippery-slope condition were more likely to think that they could get away with an inflated estimate compared to participants in the

Table 4
Mean of estimates (in $) across participants in Rounds 13–16, Study 2. Standard deviation is reported in parentheses.

Round	Slippery-slope condition	Abrupt-change condition	p-Values
13	11. 38 (1.70)	10.19 (0.73)	<0.001
14	11.26 (1.77)	10.28 (0.58)	0.002
15	10.95 (1.31)	10.23 (1.00)	0.026
16	11.06 (1.48)	10.21 (1.20)	0.008

abrupt-change condition, especially in the case of an estimate that was lower than the highly inflated estimates they saw in Rounds 11 and 12. We will rule out this alternative explanation in Studies 3 and 4.

These results are also interesting in light of research showing that individuals are more likely to commit sins of omission (e.g., ignoring inflated estimates) than sins of commission (making inflated estimates) (e.g., Sugarman, 1986). Spranca, Minsk, and Baron (1991) coined the term *omission bias* to refer to people's tendency to prefer (i.e., consider more moral and ethical) harm caused by omission over equal or lesser harm caused by acts. In the case of our study, the omission bias might have led participants to approve inflated estimates but be less open to making inflated estimates themselves.

While Study 2 provides some evidence that our approvers' implicit valuations were affected by the slippery slope, it does not provide a stringent test of the partial role of implicit processes in explaining the slippery-slope effect. Our third study addresses this issue and begins to uncover the psychological mechanisms behind the slippery-slope effect.

Study 3

Our third study attempts to replicate the results from Studies 1 and 2 and to test the possibility that some degree of the approval of unethical behavior occurs outside the awareness of the judge. Testing for bounded ethicality in our context, we examine whether unethical behaviors partially occur as the result of people unconsciously "lowering the bar" over time through small changes in their acceptance of others' ethicality. We continue to investigate the relationship between abrupt vs. gradual changes in unethical behavior and indications of implicit processes that affect the propensity of third parties to approve of the actions of others.

One possible concern in the prior studies is that participants perceived that it was in their best interest to approve estimates that they actually considered inflated. To examine the possible role of these perceptions and incentives, our third study adds a second manipulation to the one used in our first two studies related to changes in others' estimates. Specifically, we manipulate whether or not participants have incentives to accept/reject others' unethicality. Several researchers have suggested that incentives may affect the decision to engage in unethical behavior (e.g., Carr, 1968; Tenbrunsel, 1998). This conjecture was also tested empirically by Hegarty and Sims (1978), who found that people were more likely to pay bribes when they were rewarded for such behavior. Yet, other studies have found no empirical link between the size of incentives and behavior (Hegarty & Sims, 1978; Treviño & Youngblood, 1990).

Our first two studies included a financial incentive for approvers. Specifically, approvers were paid more if they accepted inflated estimates, but at the risk of a potential (but not very likely) penalty for approving egregiously exaggerated estimates. The presence of this financial incentive might have motivated people to accept the unethical behavior of others. We think that it is quite likely that a main effect for incentives will exist, such that approvers will approve inflated estimates more frequently in the incentives vs. no-incentive condition. But, we are more interested in testing our prediction that the effect between gradual and abrupt changes in estimates will be maintained even in the no-incentive condition. This latter prediction is consistent with our second hypothesis and contrasts with the possibility that the entire slippery-slope effect is driven by the intentional unethical behavior of self-interested actors trying to maximize their own outcomes.

In addition to the incentives manipulation, Study 3 participants were asked to complete an additional task at the end of Phase 2, i.e., a 12-item word-completion task. Word-fragment completion tests have been shown to assess implicit cognitive processes (Bassili & Smith, 1986; Tulving, Schacter, & Stark, 1982). Thus, the word completion task allowed us to test whether or not participants' choice of words (and their implicit thought processes) was influenced by the process they went through in reviewing estimates in Phase 2 of Study 3 (in either the abrupt-change or slippery-slope condition). We expected that participants in the slippery-slope condition would use significantly fewer words related to unethical behavior than participants in the abrupt-change condition. Support for this prediction would suggest that the greater acceptability of others' unethical behavior when such behavior develops gradually rather than abruptly occurs without the observer's conscious awareness. Such results would be consistent with recent research that argues that many unethical acts occur without the conscious awareness of the actor who engages in the unethical conduct (Banaji et al., 2003; Bazerman & Banaji, 2004) and would provide evidence consistent with *Hypothesis* 2.

Finally, in Study 3 we measured participants' reaction times (in seconds) during their approval decision in each round of Phase 2. Reactions times have been used in prior research as a valid measure of implicit biases (e.g., Greenwald, McGhee, & Schwartz, 1998). Prior research has suggested that implicit biases reflect automaticity in judgment. These automatic judgments require a shorter time to be made than conscious judgments. Thus, if unconscious processes explain (at least in part) the slippery-slope effect, then we would expect shorter reaction times in the slippery-slope condition than in the abrupt-change condition. Furthermore, we would expect reaction times to be positively correlated to the number of ethically-related words participants used in the word completion task. Such correlations would indeed suggest that the more accessible (un)ethical constructs are, the longer people take to approve or disapprove inflated estimates.

Method

Participants

One hundred forty-eight individuals (49% male, $M_{age} = 29$, $SD = 12$) participated in Study 3. Participants were recruited using ads offering money to participate in an experiment on decision-making. Most participants (74%) were students from local universities in a city in the Eastern United States. At the beginning of the experiment, participants were randomly assigned to one of the four conditions. Thirty-eight individuals participated in the no-incentives/abrupt-change condition, 36 in the incentives/slippery-slope condition, and 37 individuals participated in the other two conditions.

Design and procedure

Study 3 employs a 2 (change in estimates: slippery-slope vs. abrupt-change) × 2 (incentives: incentives vs. no incentives) design in which both change in estimates and incentives are between-subject factors. In each condition, the procedure used in Study 3 followed the one used in Study 1, with two differences. First, while participants in Study 1 played all three phases sequentially with no interruption, in Study 3 we added an additional task at the end of Phase 2. Specifically, participants were asked to complete a 12-item word-completion task in which each word stem (e.g., F R _ _ _) could be completed with an unethical behavior-related word (e.g., fraud) or with other words unrelated to ethicality (e.g., frame, fresh, frogs, fries, fruit, front, etc.). This task was designed to test whether being in the slippery-slope condition is more likely to activate thoughts relevant to relaxed ethical standards than being in the abrupt-change condition. For instance, when people are asked to complete the word "C H _ _ _", if they are thinking about cheating, they will be more likely to fill in "cheat" than alternatives such as "chime" or "chimp".

Study 3 also differed from Study 1 by including a no-incentive condition. The no-incentive condition differed from the incentive condition only in its payoff structure. As specified below, while the payoff in the incentive condition varied depending on participants' choices during the experiment, participants in the no-incentive condition were paid a fixed fee for their participation.

Participants in the incentive conditions received $10 as a show-up fee and had the opportunity to win up to an additional $25 during the experiment. The payoff structure was the same as that used in Study 1. Participants in the no-incentive conditions received $10 for their participation in addition to the $10 show-up fee.

Results

Rate of approval

At the aggregate level, over Rounds 11–16, the mean rate of approval was higher in the slippery-slope/no-incentives condition than in the abrupt-change/no-incentives condition (31% vs. 16%, χ^2 [1, $N = 450$] = 16.66, $p < 0.001$). As for the two conditions with incentives, replicating the findings of Studies 1 and 2, the rate of approval was 56% in the slippery-slope/incentives condition and only 32% in the abrupt-change/incentives condition, χ^2 (1, $N = 438$) = 30.75, $p < 0.001$.

The average number of approvals over Rounds 11–16 was used as the dependent variable in a univariate ANOVA with change in estimates (slippery slope vs. abrupt change) and incentives (incentives vs. no incentives) as between-subjects factors. This analysis reveals a significant main effect for both factors. On average, participants approved a higher number of estimates in the incentives condition ($M = 2.67$, $SD = 2.36$) than in the no-incentives condition ($M = 1.37$, $SD = 1.72$), F (1, 144) = 16.23, $p < 0.001$, $\eta^2 = 0.101$. Furthermore, participants approved a higher number of estimates in the slippery-slope condition ($M = 2.66$, $SD = 2.24$) than in the abrupt-change condition ($M = 1.39$, $SD = 1.87$), F (1, 144) = 15.57, $p < 0.001$, $\eta^2 = 0.098$. The interaction term was not significant (F [1, 144] < 1, $p = 0.35$, $\eta^2 = 0.006$).

As Table 5 shows, the percentage of "yes"es was higher in the slippery-slope condition than in the abrupt-change condition in each round following the abrupt change, i.e., Rounds 11–16. In each round, the difference in the rate of approvals between the two conditions is statistically significant (all p-values are lower than 0.05). These results are consistent with our first hypothesis, that people are more likely to accept the unethical behavior of others if the unethical behavior develops gradually, over time, than if the unethical behavior occurs abruptly.

Table 5 also reports the round-by-round results for the incentives vs. no-incentives comparisons. In each round, the difference in the number of approvals between the two conditions is statistically significant (all p-values are lower than 0.05), suggesting that the presence of incentives exacerbates the demonstrated effect of failing to act against others' gradual ethical erosion. The rate of approval over Rounds 11–16 by condition is reported in Table 6.

Overall, these results are consistent with *Hypothesis* 2 and suggest that the effect between gradual and abrupt change in estimates found in Studies 1 and 2 is maintained even when incentives are not present.

Perceived acceptability

We also examine whether approval rate in Phase 2 was affected by the estimates that participants provided in Phase 1. We estimated the following hierarchical model: $PH2\_APPROVAL_{ij} = \alpha_0 + \beta_1 TREAT_i + \beta_2 INCENTIVES_i + \beta_3 PH1\_EST_{ij}$ + *subjects' RANDOM EFFECTS* + ε, where index i refer to participants and index j refers to rounds. Explanatory variables were: (i) a dummy variable indicating the experimental condition for change in estimates (slippery slope vs. abrupt change) ($TREAT_i$), (ii) a dummy variable indicating the experimental condition for incentives ($INCENTIVES_i$), and (iii) the estimates participants provided in Phase 1 ($PH1\_EST_{ij}$). The effect of the estimates provided in Phase 1 is positive: participants were more likely to approve an estimate in Phase 2 when they overestimated the amount of money contained in the jar in Phase 1 ($\beta_3 = 0.06$, $t = 10.01$, $p < 0.001$). Yet, even after controlling for this effect, the effect of our manipulations was still in the expected direction and statistically significant; participants were more likely to approve an estimate in Phase 2 when playing in the slippery-slope condition than when playing in the abrupt-change condition ($\beta_1 = 0.22$, $t = 7.83$, $p < 0.001$), and they were also more likely to approve an estimate in Phase 2 when playing in the incentives condition than when playing in the no-incentives condition ($\beta_2 = 0.20$, $t = 7.09$, $p < 0.001$).

We conducted a similar analysis by including interaction terms for our variables of interest. We found no significant effects for the interactions.

Reaction time

We used reaction time as the dependent variable in a repeated-measure ANOVA in which incentives and change in estimate served as between-subjects factors (repeated-measure on round). This analysis revealed a significant main effect for change in estimate; on average, participants in the slippery-slope condition spent less time making their decisions (to approve or not approve the estimates of others) than participants in the abrupt-change condition (5.63 s vs. 7.38, F [1, 144] = 52.85, $p < 0.001$, $\eta^2 = 0.27$). The main effect of incentives was also significant; participants in the incentives condition spent more time making their decisions than participants in the no-incentive condition (6.80 s vs. 6.25, F [1, 144] = 5.08, $p < 0.05$, $\eta^2 = 0.03$). The change in estimate × incentives interaction only reached marginal significance (F [1, 144] = 3.53, $p = 0.06$, $\eta^2 = 0.02$). This interaction is depicted in Fig. 1. The presence of incentives lead to a significantly longer time spent on the approval decisions in the abrupt-change conditions (F [1, 73] = 7.35, $p < 0.01$, $\eta^2 = 0.09$), but not in the slippery-slope conditions ($p = 0.77$). These results are consistent with our second

Table 5
Rate of approval in Rounds 11–16 for the slippery-slope and abrupt-change conditions, and for the incentives condition (Study 3).

Round	Change in Phase 2 estimates			Incentives		
	Slippery-slope condition (%)	Abrupt-change condition (%)	p-Values	No-incentives condition (%)	Incentives condition (%)	p-Values
11	23	8	0.010	7	25	0.002
12	48	24	0.002	27	45	0.019
13	47	28	0.019	23	52	<0.001
14	49	32	0.032	27	55	<0.001
15	44	16	<0.001	20	40	0.008
16	55	31	0.003	35	51	0.049

Table 6
Rate of approval in Rounds 11–16 by condition, Study 3.

Round	No incentives		Incentives	
	Slippery-slope condition (%)	Abrupt-change condition (%)	Slippery-slope condition (%)	Abrupt-change condition (%)
11	11	3	36	14
12	32	21	64	27
13	30	16	64	41
14	32	21	67	43
15	35	5	53	27
16	46	24	64	38

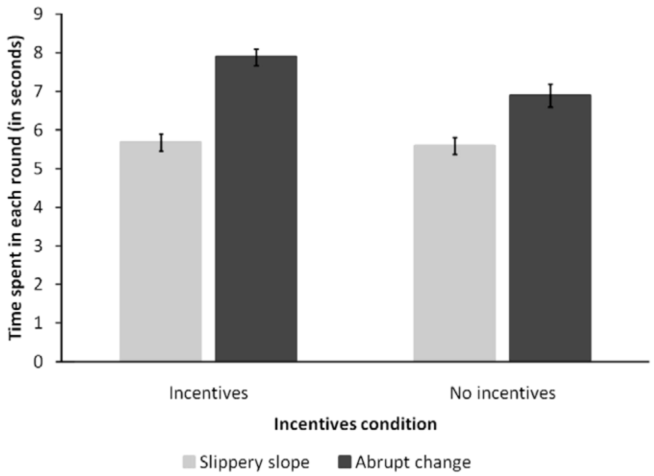

Fig. 1. Mean number of seconds spent in each round (over Rounds 11–16 of Phase 2) by condition, Study 3. Error bars represent standard errors.

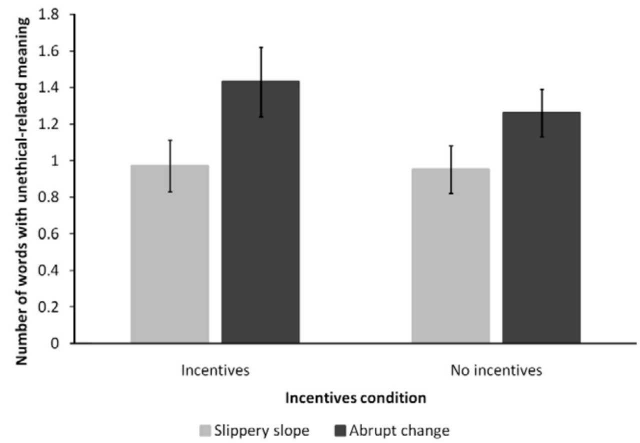

Fig. 2. Mean number of words with unethical-related meaning by condition, Study 3. Error bars represent standard errors.

hypothesis, which predicted that implicit biases partly explain the slippery-slope effect.

Word-completion task

Two independent judges, who were blind to the experimental condition and to the hypotheses of the study, counted the number of words related to unethicality and classified them as suggestive or not suggestive of relaxed ethical standards. Agreement between raters was determined by counting the number of "hits", defined as classifications on which the two raters agreed. The inter-rater reliability coefficient was 0.99, a high agreement level. We averaged the evaluations to create a score for the number of words with unethical-related meaning. We conducted an ANOVA with this score as the dependent variable and incentives (incentives vs. no incentives) and change in estimates (slippery slope vs. abrupt change) as the between-subjects factors. The mean values by condition are reported in Fig. 2. This analysis revealed a significant main effect for change in estimates; as we expected, participants in the abrupt-change condition used a significantly higher number of words related to unethical behavior ($M = 1.34$, $SD = 0.98$) than did participants in the slippery-slope condition ($M = 0.96$, $SD = 0.81$), $F(1, 142) = 6.75$, $p = 0.01$, $\eta^2 = 0.05$. No other significant effect was found ($p = 0.52$ for the main effect of incentives and $p = 0.64$ for the interaction effect). These results suggest that people were more likely to perceive the behavior of others as ethically related when it eroded abruptly rather than gradually. Consistent with our second hypothesis, these results also suggest that the greater acceptability of others' unethical behavior when such behavior develops gradually occurs without the observer's conscious awareness.

Finally, we examined whether the number of ethically-related words participants generated was correlated with reaction times. We found a significant and positive correlation ($r = 0.18$, $p < 0.05$), suggesting that the more accessible (un)ethical constructs were, the longer participants took to approve or disapprove inflated estimates. This finding suggests that the relatively faster reaction times of participants in the slippery-slope condition may be due to the lack of or reduced weighing of the ethical ramifications of approving inflated estimates.[1]

Discussion

The results of Study 3 replicate those of Studies 1 and 2: people are more likely to accept the unethical behavior of others when changes in others' unethical behavior occur gradually over time than when they occur abruptly, thus supporting *Hypothesis* 1. Furthermore, the results of Study 3 are consistent with *Hypothesis* 2 and show that the slippery-slope effect can be explained at least in part by implicit biases, as suggested by the word-completion results.

In Study 3, we also added a no-incentive manipulation to rule out the possibility that the entire effect is due to differences in the intentional assessment of what strategy will maximize one's own payoff. The effect between gradual and abrupt change in estimates was maintained even when incentives were not present, providing evidence consistent with the argument that people perceive gradual changes in others' unethical behavior as acceptable.

The word-completion results also suggest that abrupt changes in others' unethical behavior lead the judge's mind to focus more on ethical issues than the gradual changes. In addition, the word completion results provide support for the argument that at least part of this process occurs outside of intentional awareness.

Finally, additional support for the role of implicit biases in explaining the slippery-slope effect comes from the results about reaction times, which show that participants in the slippery-slope conditions spent less time deciding whether or not to approve the estimates of other participants than those in the abrupt-change condition.

Study 4

Study 3 found that even when incentives were not present, the rate of approval of inflated estimates was still higher in the slippery-slope condition than in the abrupt-change condition. While the no-incentive condition used in Study 3 eliminates personal monetary gains for approving exaggerated estimates, participants might have had other reasons to accept the estimates of others.[2] To address this issue, we conducted a fourth study in which we changed the incentives for approvers. Differently from our previous studies, in Study 4 approvers were incentivized to detect fraud.

[1] We thank an anonymous reviewer for suggesting this analysis and interpretation of such positive correlation.

[2] We thank an anonymous reviewer for suggesting this possibility.

Method

Participants

As in our other studies, we recruited participants through ads offering money to participate in an experiment on decision-making. Sixty-five individuals (54% male; $M_{age} = 24$, $SD = 4.92$) participated in the study for pay. Most participants (78%) were students from local universities in a city in the Southeastern United States.

Design and procedure

Study 4 followed the same procedure used in Study 1 with only one difference: while participants in Study 1 had incentives to approve inflated estimates, Study 4 participants had the incentive to simply detect fraud. Specifically, participants in Study 4 earned additional money in each round in which they identified estimates that were more than 110% of the real amount in the pictured jars.

Participants received $5 as a show-up fee and had the opportunity to win additional money during the experiment. The payoff structure was the same as in Study 1 with regard to Phase 1 (when participants played the role of estimators) but different with regard to Phases 2 and 3 (when participants played the role of approvers). When playing the role of approvers, participants received $1 every time they correctly identified whether or not the estimates were within a 10% range of accuracy regarding the real amount of money in the pictured jar (i.e., they approved accurate estimates or rejected inflated ones).

Dependent measure

As in the other studies, we tested our hypothesis on the data collected in Phase 2, in which our manipulation occurred. Given the nature of the new instructions to approvers, our dependent variable has a different meaning than the approval rate used in Studies 1 through 4 even if it is computed the same way. Approving an estimate in Study 4 meant stating that the estimate was within the 10% range of accuracy.

Results

Rate of approval of inflated estimates

The estimates participants received in Phase 2 over Rounds 11–16 are outside the 10% range of accuracy specified in the instructions to the fraud-detection task. Thus, if participants were to correctly estimate the amount of money in the pictured jar in each round, we would expect them to always reject Phase 2 estimates. We computed the rate of approval over Rounds 11–16 in Phase 2 within each condition. As shown in Table 7, the percentage of "yeses was higher in the slippery-slope condition than in the abrupt-change condition in each round following the abrupt change (all p-values are lower than 0.05, except in Round 13, in which $p < 0.10$). At the aggregate level (over Rounds 11–16), the mean rate of approval was higher in the slippery-slope condition than in the abrupt-change condition (19% vs. 3%, χ^2 $[1, N = 390] = 27.28$, $p < 0.001$). In addition, the average number of approvals was not significantly different from zero in the abrupt-change condition (t $[32] = 1.97$, $p = 0.06$), but it was in the slippery-slope condition (t $[31] = 3.19$, $p < 0.01$).

Table 7
Rate of approval in Rounds 11–16 for each condition, Study 4.

Round	Slippery-slope condition (%)	Abrupt-change condition (%)	p-Values
11	13	0	0.036
12	19	3	0.041
13	16	3	0.079
14	22	3	0.021
15	19	0	0.009
16	25	6	0.034

Discussion

The results of Study 4 replicate the findings of our first three studies and provide further support for *Hypothesis* 1, which predicted that participants would be more likely to accept the unethical behavior of others if the unethical behavior develops gradually over time than if the unethical behavior occurs abruptly. Using a fraud-detection task, we found that participants in the slippery-slope condition were still more likely to approve inflated estimates than participants in the abrupt-change condition even if they were incentivized to detect fraud. These findings provide added support also for *Hypothesis* 2, which predicted that implicit biases would partly explain the slippery-slope effect.

General discussion and conclusions

Where is the line between not accepting any exception to ethicality and ignoring another person's unethical behavior? What does it take for ordinary people to slide across this line? We like to believe that only a few bad apples cross to the other side. In fact, under certain conditions, most of us can be expected to engage in unethical behavior. Here, our goal was to explore one condition that can lead ordinary people to cross that line. Namely, we hypothesized that people are more likely to accept the unethical behavior of others if the unethical behavior develops gradually, over time, than if the unethical behavior occurs abruptly. Providing support for this hypothesis, our results show that study participants were more likely to accept the unethical behavior of others if the unethical behavior developed gradually over time than if the unethical behavior occurred abruptly. It is worth noting the change in rate of approvals between Round 11 (when the abrupt change occurred) and the following rounds across both the slippery-slope and abrupt-change conditions. We observed this pattern of results in all of our studies except Study 4, in which approvers had the incentive to detect fraud. These results suggest that participants might have rewarded estimators when they crossed ethical boundaries to a small extent. Alternatively, participants might have become less sensitive to moral violations when such violations occurred repeatedly over time.

We also predicted and demonstrated that the slippery-slope effect is due, in part, to implicit biases that result in greater acceptance of others' unethical behavior if ethical erosion occurs through small changes. A useful metaphor for how implicit processes lead to unethical behavior in organizations is provided by the "boiling frog syndrome" (Senge, 1994). According to this old folk tale, if you throw a frog into boiling water, it will quickly jump out. But if you put a frog in a pan of warm water and raise the temperature very slowly, the gradual warming will make the frog doze happily. The frog will eventually cook to death due to its failure to sense the gradual increase in water temperature. The message of this tale is that, because its environment changes so gradually, the frog is never stimulated to take bold action to save its life. Similarly, this paper suggested and demonstrated that third parties (such as accountants working for clients involved in business scandals) may not react to a gradual deterioration of their clients' ethicality. Indeed, third parties might grow accustomed to unethical behavior *if* the latter develops gradually, and, as a result, accept such behavior.

Our studies show that when the unethical behavior of others develops gradually, over time, instead of occurring abruptly, people are more likely to accept this behavior. Other situational factors might facilitate the crossing of the line into unethical behavior; these could be the focus of further studies. For instance, a situational factor worthy of exploration in future research is diffusion of responsibility. When multiple people are responsible for a decision, rather than just one person, they are less likely to take responsibility for their unethical acts, thereby leading to an increase in unethical behavior (Bandura, 1999). Bandura (1999) notes that division of labor, or the subdivision of job responsibilities, is one mechanism that can lead to diffusion of responsibility. The task implemented in our studies may be an example of how diffusion of responsibility operates. Our task was subdivided: one person estimates, another accepts or rejects those estimates (the participant), and another checks those estimates 10% of the time. It is possible that, in both the slippery-slope and the abrupt-change conditions, participants were more willing to accept incorrect estimates because they felt they may not be the "final say" as to whether or not the estimates were correct. Specifically, participants may have been more likely to agree with incorrect estimates because they knew that, 10% of the time, someone else would double-check the estimate. Also, participants may have felt even less responsibility because they themselves did not make the unethical estimates. Therefore, the segmentation of the task could have led to an increase in acceptance of unethical estimates.

Future research could also go a step further than the studies presented here and investigate the inner workings of implicit biases in explaining the slippery-slope effect. We argued and demonstrated that, at least in part, the slippery-slope effect occurs beyond one's own intentionality or awareness. But what exactly are the consequences of implicit biases in producing the slippery-slope effect? On one hand, these biases might lead individuals to fail to notice shifts in others' behavior toward unethicality. In this case, implicit biases would affect individuals' perception processes. On the other hand, these implicit biases might lead individuals to fail to interpret the behavior of others as unethical. In this case, implicit biases would affect individuals' inferential or attributional processes. While the studies presented here do not allow us to clearly distinguish between these two processes, future research addressing this issue could provide important insights into the boundary conditions of the slippery-slope effect, as well as further evidence of the mechanisms explaining it.

Related to this issue, future research could further examine the role of conscious and unconscious processes in driving the slippery-slope effect demonstrated here, and could explore in more detail the multiply determined nature of such effect. In Study 3, we found a significant effect for incentives, suggesting the possibility that at least part of the slippery-slope effect may be driven by participants consciously responding to monetary incentives. Further studies manipulating both incentives and erosion in others' unethical behavior will increase our understanding of the main drivers of the slippery-slope effect.

Finally, future research could investigate potential ways to reduce or eliminate the slippery-slope effect demonstrated here. For instance, does creating an environment in which people feel pressured to justify their actions or decisions to others lead people to behave unethically? One strategy that can help people cope with such pressure is the "acceptability heuristics", through which people "adopt positions likely to gain the favor of those to whom they feel accountable" (Tetlock, 1992: 340). At times, people might use acceptability heuristics when coping with ethical dilemmas. For instance, Pennington and Schlenker (1999) show that the audience to whom one is accountable can shape one's views and behaviors regarding an ethical decision. Related studies by Brief, Dukerich, and Doran (1991) support such "conformity effects": when making decisions, people care less about the content of the decisions (i.e., whether or not the decisions are ethical) than about the potential acceptance of such decisions by those to whom they are accountable. Similar effects of compliance were earlier shown by Milgram's studies (1974) on obedience to authority. Building on this work, future research could investigate whether accountability reduces or even eliminates the slippery-slope effect in ethical judgment.

We recommend that researchers investigate these situational factors as well as potential mediators of our demonstrated effect. For individuals in the position of watchdog either within or outside an organizational setting, such research might provide invaluable guidance in judging others' behavior. Similarly, organizations interested in creating or maintaining an ethically healthy environment might use these findings to identify factors that can correct for individual shortcomings. We close by stressing that we believe and support research that identifies and attempts to reduce intentional unethical behavior. However, we believe that the unintentional aspects of ethical misconduct are also worthy of study. Accordingly, we have offered an additional perspective for understanding the psychology of this ordinary unethical behavior.

Acknowledgments

The authors gratefully acknowledge the financial support of the Harvard Business School Division of Faculty Research and Development and the support of the staff and facilities of the Computer Laboratory for Experimental Research, the Center for Behavioral Decision Research at Carnegie Mellon University, and the Center for Decision Research at the University of North Carolina at Chapel Hill. We are especially grateful to Daylian Cain, Howard Kunreuther, Don Moore, Lamar Pierce, Jan Rivkin, Maurice Schweitzer, and Sandra Sucher for their insightful feedback on earlier drafts. We also thank the three anonymous reviewers and the associate editor for their comments and guidance during the revision process.

References

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1*, 323–338.

Andrews, K. R. (1989). Ethics in practice. *Harvard Business Review*(September–October), 99–104.

Ashforth, B. E., & Anand, V. (2003). The normalization of corruption in organizations. *Research in Organizational Behavior, 25*, 1–52.

Banaji, M. R., Bazerman, M. H., & Chugh, D. (2003). How (un)ethical are you? *Harvard Business Review, 81*(12), 56–64.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality & Social Psychology Review, 3*, 193–209.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Hillsdale, NJ: Lawrence Erlbaum.

Bassili, J. N., & Smith, M. C. (1986). On the spontaneity of trait attribution: Converging evidence for the role of cognitive strategy. *Journal of Personality and Social Psychology, 50*, 239–245.

Bazerman, M. H., & Banaji, M. R. (2004). The social psychology of ordinary ethical failures. *Social Justice Research, 17*(2), 111–115.

Bazerman, M. H., & Samuelson, W. F. (1983). I won the auction but don't want the prize. *Journal of Conflict Resolution, 27*, 618–634.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*, 169–217.

Blair, I. V. (2001). Implicit stereotypes and prejudice. In Gordon B. Moskowitz (Ed.), *Cognitive social psychology. The Princeton symposium on the legacy and future of social cognition*. Psychology Press.

Brass, D. J., Butterfield, K. D., & Skaggs, B. C. (1998). Relationships and unethical behavior: A social network perspective. *Academy of Management Review, 23*(1), 14–31.

Brief, A. P., Buttram, R. T., & Dukerich, J. M. (2001). Collective corruption in the corporate world: Toward a process model. In M. E. Turner (Ed.), *Groups at work: Theory and research. Applied social research* (pp. 471–499). Mahwah, NJ: Lawrence Erlbaum.

Brief, A. P., Dukerich, J. M., & Doran, L. I. (1991). Resolving ethical dilemmas in management: Experimental investigations of values, accountability and choice. *Journal of Applied Social Psychology, 17*, 68–79.

Carr, A. Z. (1968). Is business bluffing unethical? *Harvard Business Review, 46*(1), 143–153.

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.), *Conflict of interest: Challenges and solutions in business, law, medicine, and public policy* (pp. 74–95). New York: Cambridge University Press.

Gellerman, S. W. (1986). Why 'good' managers make bad ethical choices. *Harvard Business Review*(July–August), 85–90.
Greenberger, D. B., Miceli, M. P., & Cohen, D. J. (1987). Oppositionists and group norms: The reciprocal influence of whistle-blowers and co-workers. *Journal of Business Ethics, 6*, 527–542.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480.
Gundlach, M. J., Douglas, S. C., & Martinko, M. J. (2003). The decision to blow the whistle: A social information processing framework. *Academy of Management Review, 28*(1), 107–123.
Hegarty, W. H., & Sims, H. P. (1978). Some determinants of unethical behavior: An experiment. *Journal of Applied Psychology, 63*, 451–457.
Hill, J., & Kochendorfer, R. A. (1969). Knowledge of peer success and risk of detection as determinants of cheating. *Developmental Psychology, 1*, 231–238.
Hosmer, L. T. (1987). The Institutionalization of unethical behavior. *Journal of Business Ethics, 6*(6), 439–447.
Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue contingent model. *Academy of Management Review, 16*, 366–395.
Kelman, H. C., & Hamilton, L. (1989). *Crimes of obedience*. New Haven, CT: Yale University Press.
Kern, M. C., & Chugh, D. (2009). Bounded ethicality: The perils of loss framing. *Psychological Science, 20*(3), 378–384.
King, G. (1997). The effects of interpersonal closeness and issue seriousness on blowing the whistle. *Journal of Business Communication, 34*, 419–436.
Leming, J. S. (1980). Cheating behavior, subject variables, and components of the internal–external scale under high and low risk conditions. *Journal of Educational Research, 74*, 83–87.
Levitt, A., & Dwyer, P. (2002). *Take on the street: What Wall Street and corporate America don't want you to know, and what you can do to fight back*. New York: Pantheon.
Longnecker, J. G. (1985). Management priorities and management ethics. *Journal of Business Ethics, 4*, 65–70.
Messick, D. M., & Bazerman, M. H. (1996). Ethics for the 21st century: A decision making approach. *Sloan Management Review, 37*, 9–22.
Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology, 21*, 480–500.
Miceli, M. P., Dozier, J. B., & Near, I. P. (1991). Blowing the whistle on data fudging: A controlled field experiment. *Journal of Applied Social Psychology, 21*, 271–295.
Miceli, M. P., & Near, J. P. (1992). *Blowing the whistle*. Riverside, NJ: Macmillan Publishing Company.
Miceli, M. P., & Near, J. P. (1988). Individual and situational correlates of whistle-blowing. *Personnel Psychology, 41*, 267–281.
Miceli, M. P., & Near, J. P. (1984). The relationship among beliefs, organizational position, and whistle-blowing status: A discriminant analysis. *Academy of Management Journal, 27*(4), 687–705.
Miceli, M. P., Near, J. P., & Schwenk, C. R. (1991). Who blows the whistle and why? *Industrial and Labor Relations Review, 45*(1), 113–130.
Michaels, J. W., & Miethe, T. D. (1989). Applying theories of deviance to academic cheating. *Social Science Quarterly, 70*, 870–885.
Milgram, S. (1974). *Obedience to authority: An experimental view*. New York, NY: Harper & Row.
Molander, E. (1987). A paradigm for design, promulgation and enforcement of ethical codes. *Journal of Business Ethics, 6*, 619–631.
Mollenkamp, C., Levy, A., Menn, J., & Rothfeder, J. (1998). *The people vs. big tobacco*. Princeton, New Jersey: Bloomberg Press.
Moore, D. A., Tetlock, P. E., Tanlu, L., & Bazerman, M. H. (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review, 31*(1), 10–29.
Near, J. P., Dworkin, T. M., & Miceli, M. P. (1993). Explaining the whistle-blowing process: Suggestions from power theory and justice theory. *Organization Science, 4*, 393–411.
Near, J. P., & Miceli, M. P. (1996). Whistle-blowing: Myth and reality. *Journal of Management, 22*(3), 507–526.
Near, J. P., & Miceli, M. P. (1995). Effective whistleblowing. *Academy of Management Review, 20*, 679–708.
Near, J. P., & Miceli, M. P. (1986). Retaliation against whistleblowers: Predictors and effects. *Journal of Applied Psychology, 71*, 137–145.
Near, J. P., & Miceli, M. P. (1985). Organizational dissidence: The case of whistle-blowing. *Journal of Business Ethics, 4*, 1–16.
Pennington, J., & Schlenker, B. R. (1999). Accountability for consequential decisions: Justifying ethical judgments to audiences. *Personality and Social Psychology Bulletin, 25*, 1067–1081.
Pitt, L. F., & Abratt, R. (1986). Corruption in business—are management attitudes right? *Journal of Business Ethics, 5*, 39–44.
Reynolds, S. J. (2006). Moral awareness and ethical predispositions: Investigating the role of individual differences in the recognition of moral issues. *Journal of Applied Psychology, 91*(1), 233–243.
Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology, 37*, 322–336.
Sugarman, D. B. (1986). Active versus passive euthanasia: An attributional analysis. *Journal of Applied Social Psychology, 16*, 60–76.
Senge, P. M. (1994). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Currency Doubleday.
Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition, 7*(1–3), 1–15.
Sims, R. L., & Keenan, J. P. (1998). Predictors of external whistleblowing: Organizational and intrapersonal variables. *Journal of Business Ethics, 17*(4), 411–421.
Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*, 76–105.
Steininger, M., Johnson, R. E., & Kirts, D. K. (1964). Cheating on college examinations as a function of situationally aroused anxiety and hostility. *Journal of Educational Psychology, 55*, 317–324.
Taylor, S. E. (1989). *Positive illusions*. New York: Basic Books.
Tenbrunsel, A. E. (1998). Misrepresentation and expectations of misrepresentation in an ethical dilemma: The role of incentives and temptation. *Academy of Management Journal, 41*(3), 330–339.
Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research, 17*(2), 223–236.
Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 331–376). New York, NY: Academic Press.
Tittle, C. R., & Rowe, A. R. (1973). Moral appeal, sanction threat and deviance: An experimental test. *Social Problems, 20*, 488–498.
Treviño, L. K., & Youngblood, S. A. (1990). Bad apples in bad barrels: A causal analysis of ethical decision-making behavior. *Journal of Applied Psychology, 75*, 378–385.
Treviño, L. K., Weaver, R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management, 32*(6), 951–990.
Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 336–342.
Vitro, F. T., & Schoer, L. A. (1972). The effects of probability of test success, test importance, and risk of detection on the incidence of cheating. *Journal of School Psychology, 10*, 269–277.