Corporate Social Responsibility and Employee Misconduct

Vanessa Burbano, Columbia Business School*

&

Bennett Chiles, Columbia Business School*

Working Paper, October 2018

We study corporate social responsibility (CSR) as a potential governance tool for reducing employee misconduct. A baseline study of full-time U.S. workers indicates that CSR reduces participants' stated intentions to lie to their employer and that this result is driven by an increased sense of moral obligation: individuals appear to feel worse about behaving badly when they perceive their employer to be doing good. To examine the causal effect of CSR on actual employee misconduct, we implement a randomized field experiment on Amazon Mechanical Turk in which workers are hired to complete a job designed to elicit workers' willingness to lie to their employer. We manipulate whether the workers' employer engages in CSR, and, as a point of comparison, whether the worker is required to sign an honor code pledge. We find CSR to decrease employee misreporting substantially, with effects similar in magnitude to those of an honor code pledge. Given the challenges of measuring employee misconduct in practice, a notable contribution of this study is the way in which our experimental design allows us to cleanly observe the extent to which employees lie to their employer in a real work context and thus estimate the treatment effect of CSR on employee misconduct.

*Equal authorship.

Keywords: corporate social responsibility, unethical behavior, employee misconduct, employee governance, honor code, moral obligation, moral licensing, field experiment

INTRODUCTION

Employee misconduct is costly to organizations – shirking and time theft alone are estimated to cost U.S. firms as much as $400 billion annually (Shulman, 2007), and more egregious behavior such as direct financial theft and outright fraud may add up to an additional $200 billion (Murphy, 1993). As such, preventing unethical behavior among workers is a key concern for managers, especially in human capital intensive industries. Scholarly research on misconduct, however, has been limited by the fact that unethical behavior is, by its nature, typically concealed and thus difficult to observe. Existing work has focused primarily on the potential to reduce misconduct via increased monitoring/control (Becker, 1968; Hubbard, 2000; Nagin, Rebitzer, Sanders, and Taylor, 2002; Detert et al., 2007; Olken, 2007; DeHoratius and Raman, 2008; Pierce, Snow, and McAfee, 2015), financial incentives (Lazear and Rosen, 1981; Dye, 1984; Konrad, 2000; Duflo, Hanna, and Ryan, 2012; Flory, Leibbrant, and List, 2016; Balasubramanian, Bennett, and Pierce, 2017), and moral reminders such as honor codes (McCabe, Trevino and Butterfield, 1996; Shu et al., 2012). In this paper, we offer evidence of a substantially understudied tool that firms may utilize to discourage unethical behavior among workers: corporate social responsibility (CSR).

Recent empirical evidence suggests that firms may, indeed, be attempting to use CSR as an employee governance tool to reduce the possibility of adverse employee behavior. Flammer and Luo (2017) find that firms increase their engagement in employee-related CSR when unemployment insurance benefits for workers increase (thus reducing the cost of being unemployed and increasing employees' incentives to engage in misconduct). This finding seems sensible in light of other work showing that CSR boosts employee motivation (Rupp et al., 2006; Rupp et al. 2013) and productivity (Tonin and Vlassopoulous, 2014; Burbano, 2018). Yet while Flammer and Luo (2017) make a compelling case that firms appear to be *attempting* to use CSR to reduce misconduct, their data does not allow them to show that CSR is actually effective in serving this purpose. Our study fills this gap by examining the causal link between CSR and employee misconduct directly.

Theoretically, there are several mechanisms through which CSR might affect employees' propensity to behave unethically. A substantial body of research suggests that CSR increases employees' identification

with their organization (Ashforth and Mael 1989; Dutton and Dukerich, 1991; Mael and Ashforth, 1992; Dutton, Dukerich, and Celia, 1994; Turban and Greening, 1996; Greening and Turban, 2000; Brockner et al., 2013), which, in turn, should decrease their propensity to engage in misconduct. Another way in which CSR might reduce misconduct is by increasing employees' sense of moral obligation. Hansen et al. (2011), for example, suggest that individuals who perceive their employers as socially responsible will feel an increased obligation to behave ethically. Such obligation is also in line with the internal cost-benefit framework for dishonesty proposed by Mazar, Amir, and Ariely (2008) if workers feel worse about cheating a socially responsible employer (i.e., if CSR increases the "internal" costs of dishonesty). Finally, in contrast, a concurrent working paper from List and Momeni (2017) posits that CSR may sometimes actually *increase* unethical behavior on the job. The underlying theory for why this might occur is moral licensing: when "good" behavior in one domain causes individuals to feel license to behave "badly" in other domains (Benabou and Tirole, 2010; see Merritt, Effron and Monin 2010 for a review on experimental evidence of moral licensing). In this particular context, the idea is that if employees feel that the very act of working for a socially responsible firm is, in itself, a prosocial deed, then this may cause them to feel license to behave less ethically on the job.

In this paper, we present two studies designed to test for a causal relationship between CSR and employee misconduct; we also explore the relevance of the potential underlying mechanisms discussed above. Study 1 is a survey of full-time U.S. employees administered through Qualtrics Panel. Here, we presented employees with a hypothetical scenario in which they were tasked with completing market research phone calls for their employer. In this scenario, we emphasized to participants that their compliance with the assigned task would be difficult for their supervisor to observe – thus largely eliminating any perceived (external) "costs" associated with shirking. We manipulated whether participants received information about the social responsibility of the hypothetical employer (or not) and measured the impact of this CSR treatment on participants' stated likelihood of misreporting call attempts/outcomes. As a benchmark, we also included an honor code (HC) treatment, as honor codes have previously been shown

to decrease unethical behavior substantially (e.g., Shu et al., 2012). Finally, all participants were asked a series of questions to assess potential mechanisms for any observed treatment effects.

We found that CSR reduced the (stated) likelihood of misreporting, but only among individuals who believed that the socially responsible employer was "doing good." This is consistent with the notion that CSR must be perceived to be authentic and non-instrumental for it to induce positive behavioral results (e.g., Cassar and Meier, 2017). The magnitude of the CSR effect was similar to that of the HC treatment. Notably, our results also indicate that CSR and honor codes work via the same underlying mechanism: moral obligation. Both CSR and the implementation of an honor code caused individuals to feel badly about lying to their employer; this increased sense of moral obligation, in turn, mediated the effect of CSR and honor code signing on individuals' stated likelihood of behaving unethically. In contrast, we find only marginal evidence that CSR activates increased worker identification with the employer and no evidence that CSR elicits moral licensing (those in the CSR treatment group were not any more likely to feel that they themselves were "doing good").[1]

Study 2, our main study, mirrors Study 1 but elicits revealed (as opposed to merely stated) employee preference by utilizing gig workers on Amazon Mechanical Turk ("MTurk"). These workers were recruited to complete actual market research phone calls for payment and were told that they would receive additional bonuses for each survey response that they were able to obtain. While we did not directly state to these participants that their compliance with the assigned task would be unobservable, it was clear from the parameters of the job that, under typical conditions, it would be impossible for their employer to ascertain whether or not the calls were attempted as instructed and/or whether or not any survey responses were legitimately obtained. This job is, thus, one where workers might be expected to misreport quite frequently. Unbeknownst to the workers, however, we owned the phone numbers listed on their call sheets and were thus able to cleanly observe actual versus reported outcomes. This enabled us to quantify worker

[1]Moreover, contrary to List and Momeni's (2017) argument that, through moral licensing, CSR should cause individuals to misreport *more* frequently, we found that those individuals who did believe that they were doing good were less (not more) likely to say that they would misreport.

misreporting – an outcome that is usually impossible to effectively measure in field settings since workers have the incentive to conceal such behavior. We manipulated whether the workers received information about the social responsibility of the employer (or not), and whether workers were required to sign an honor code (or not), then tracked the effect of these treatments on misreporting.

Here, we found that both CSR and the implementation of an honor code reduced workers' propensity to lie to their employer about call completion but had no effect on their propensity to illegitimately claim bonuses. In our fully specified model, individuals who received the CSR treatment misreported 0.65 fewer calls on average (out of the five total voicemails they were tasked with completing); workers in the honor code treatment group misreported 0.54 fewer calls on average. These effects, however, were not fully additive – if the employer was engaged in CSR, there was little incremental benefit to implementing an honor code. Notably, we also find that individuals who reported that they volunteer with charities were more likely to behave unethically on the job, which is consistent with the theory of moral licensing.[2] This finding suggests that at the *individual* level, "good behavior" may, indeed, cause employees to feel license to behave badly. In contrast, our main findings (that CSR has a negative impact on misreporting) cast doubt on the notion that *organizational*-level "good behavior" elicits this sort of moral licensing among workers.

Taken together, our studies have important implications for the potential use of CSR as an employee governance tool. Our central finding – that CSR reduces misreporting – was empirically robust in both studies. Moreover, while the use of MTurk in Study 2 does come with the usual limitations (the employer-employee relationship in this setting is admittedly different from that in many organizations), we believe our unique field experimental design, which allows us to observe actual employee misconduct in a natural (if not "typical") context, is a noteworthy contribution given the difficulty of observing such behavior in practice. This limitation has substantially hindered research on employee misconduct to date, and existing work has been largely based on laboratory experiments or self-reported survey data (Edelman and Larkin, 2015; Pierce and Snyder, 2008; Pierce and Balasubramanian, 2015) rather than on behavioral field

[2] This is based on self-reported volunteer data. We discuss the limitations of such data for drawing conclusions about this particular finding in our writeup.

evidence. Finally, this paper also contributes to our understanding of another mechanism through which CSR may benefit the firm – by reducing unethical behavior – and sheds light on an important boundary condition of this effect: that workers must perceive employers engaged in CSR to be doing good.

MOTIVATION AND THEORETICAL FRAMEWORK

Drivers of Employee Unethical Behavior

Scholars have examined individual-level characteristics that influence employee misconduct, including but not limited to having unmet goals (Shweitzer, Ordonez, & Douma, 2017), cognitive moral development (Ford & Richardson, 1994; Kohlberg, 1969; Trevino and Youngblood, 1990), locus of control (Hegarty and Sims, 1978), rivalry (Kilduff, Galinksy, Gallo, and Reade, 2015), promotion focus (Gino and Morgolis, 2011), strength of moral identity (Mayer, Aquino, Greenbaum, and Kuenzi, 2012), and in some contexts, even prosocial motivation (Pierce & Balasubramanian, 2015).

In theory, firms could seek to attract and hire individuals of certain characteristics to reduce unethical behavior in the workplace. As many of these individual-level characteristics are difficult to observe prior to hiring, however, another avenue by which organizations can reduce employee misconduct is to implement policies and take on organization-level characteristics that reduce this type of detrimental behavior. Much of the work in this domain has focused on reducing misconduct via increased monitoring/control (Becker, 1968; Hubbard, 2000; Nagin, Rebitzer, Sanders, and Taylor, 2002; Detert et al., 2007; Olken, 2007; DeHoratius and Raman, 2008; Pierce, Snow, and McAfee, 2015), financial incentives (Lazear and Rosen, 1981; Dye, 1984; Konrad, 2000; Duflo, Hanna, and Ryan, 2012; Flory, Leibbrant, and List, 2016; Balasubramanian, Bennett, and Pierce, 2017), and moral reminders such as honor codes (McCabe, Trevino and Butterfield, 1996; Shu et al., 2012). Other work has examined the importance of employee relationships and reference groups (Jones, 1997; Umphress and Bingham, 2011; Brass, Butterfield, and Skaggs, 1998) and the organization's ethical climate (Victor and Cullen, 1988) as important determinants of unethical behavior.

Much of the research on employee misconduct has relied on survey and laboratory evidence due to the innate challenges of observing unethical behavior (Pierce & Balasubramanian, 2015), which, naturally, employees typically attempt to conceal. As such, there is an opportunity for behavioral field research on the topic to make an important contribution. Given the empirical challenge of holding constant the unobservable organization-level characteristics likely to influence employee misconduct, field experimental studies in which individuals are observed behaving in their natural work context are one important avenue to help us causally identify drivers of employee unethical behavior (Balafoutas, Beck, Kerschbamer, Sutter, 2013; List and Momeni, 2017).

Corporate Social Responsibility and Employee Unethical Behavior

One organizational-level characteristic whose influence on employee unethical behavior has been understudied is corporate social responsibility. Though there are many channels through which CSR has been shown to influence employee behavior (Bode et al., 2015; Burbano, 2016; Burbano, Mamer and Snyder, 2018; Carnahan et al., 2017), the relationship between CSR and unethical behavior has received relatively little empirical attention in the literature (List and Momeni, 2017). Furthermore, while there are some theoretical perspectives that suggest CSR should result in *less* unethical behavior, there are also theoretical arguments for why CSR might actually increase such misbehavior. We discuss three key perspectives below.

An Identification with the Employer Mechanism. One way in which CSR might reduce bad behavior on the job is by increasing employees' sense of identification with their employers. CSR has been purported to help satisfy an employee's need for a meaningful existence (Rupp et al., 2006; Rupp et al., 2013) and manifest as improved self-image and self-concept among employees. That is, when an employee of a socially responsible firm favorably compares his or her qualities—or those of his or her employer—to those of others, his or her self-image increases (Ashforth and Mael, 1989; Dutton and Dukerich, 1991; Greening and Turban, 2000; Turban et al., 1996; Brockner et al., 2013). Higher self-image and self-concept increase the attractiveness of categorizing oneself as part of an organization and thus, increase organizational identification (Ashforth abd Mael 1989; Brockner et al., 2013; Dutton and Dukerich, 1991; Dutton,

Dukerich, and Celia, 1994; Greening and Turban, 2000; Mael and Ashforth, 1992; Turban and Greening, 1996). Organizational identification has in turn been linked to prosocial organizational behaviors by employees (Bateman and Organ, 1983; Illies, Scott, and Judge, 2006; O'Reilly and Chatman, 1986; Organ and Ryan, 1995), that is, behavior which benefits, rather than harms, the organization (Brief and Motowidlo, 1986). As such, this mechanism would imply that CSR should result in less employee behavior that is harmful to the organization, including unethical behavior.

A Moral Obligation Mechanism. Another theoretical rationale for why CSR might reduce, rather than increase, employee misbehavior is by increasing employees' sense of moral obligation towards their employers. Utilizing a deontic justice framework, Hansen et al. (2011), for example, posit that individuals who perceive their employers as socially responsible will feel an increased obligation to "give back" in the form of good behavior. Such obligation is in line with the internal cost-benefit framework for dishonesty proposed by Mazar, Amir, and Ariely (2008). Here, in addition to weighing the *external* costs and benefits of dishonesty (e.g., potential punishment vs. potential financial gain), individuals also weigh the *internal* costs and benefits, specifically with regard to the impact that dishonesty might have on their own self-concept. Thus, if CSR increases the internal costs of dishonesty (i.e., if employees feel worse about cheating a socially responsible employer), then we would expect CSR to reduce misreporting in practice.

A Moral Licensing Mechanism. In contrast, moral licensing theory suggests that prosocial activities may cause *increased* prevalence of unethical behavior among employees. To the extent that working at a socially responsible organization makes employees more inclined to believe that they, themselves, are doing good, then these workers may actually feel increased license to behave badly (Benabou and Tirole, 2010; see Merritt, Effron and Monin 2010 for a review on experimental evidence of moral licensing). Indeed, a working paper from List and Momeni (2017) finds that, when provided with a CSR statement designed to elicit moral licensing, gig workers on MTurk shirk more frequently.[3]

[3] Notably, though, the authors' main results are statistically significant only when the pro-social act is framed in direct relation to the *individual employee*. Specifically, only when workers are told that, "We donate the equivalent

Empirical Study of the Effect of CSR on Employee Unethical Behavior

The dearth of existing empirical evidence on the relationship between CSR and unethical behavior is likely due in part to the challenges of observing such employee behavior in the first place, and, furthermore, of ascribing a causal relationship between CSR and misbehavior. Flammer and Luo (2017) leverage shocks to unemployment benefits to demonstrate a causal relationship between higher unemployment benefits (which should arguably reduce the costs of being unemployed and hence increase employees' incentives to engage in adverse behavior) and the utilization of employee-related CSR. They use this finding to make the case that companies use CSR as an employee governance tool to counter the likelihood of adverse behavior. They stop short, however, of demonstrating a causal relationship between CSR and actual adverse employee behavior. A more thorough understanding of this causal relationship is, indeed, critical, as a concurrent working paper from List and Momeni (2017) suggests that MTurk gig workers, when presented with certain types of CSR messaging, actually shirk *more*. While these findings are quite troubling if employers are, indeed, attempting to utilize CSR as a governance tool, there is substantial reason to cast doubt on their generalizability. List and Momeni's implementation of a CSR treatment is atypical of the sort of language that most real firms use to communicate their social initiatives, and statistically significant effects are limited to a particular wording explicitly crafted to elicit moral licensing.[4] There thus remains a compelling need to study the relationship between employee misconduct and CSR as more typically implemented and communicated in practice.

Honor Codes and Employee Unethical Behavior

of x% of our wage bill in cash *(on behalf of all workers who help us with this project)* to UNICEF Education Programs" do List and Momeni (2017) find a statistically significant effect on unethical behavior. When the same statement without the parenthetical is administered, key results are insignificant. One important open question, then, is whether or not moral licensing is elicited when CSR is framed in terms of the *organization* rather than the individual.

[4] Only when workers are told that, "We donate the equivalent of x% of our wage bill in cash *(on behalf of all workers who help us with this project)* to UNICEF Education Programs" do List and Momeni (2017) find a statistically significant effect on unethical behavior. When the same statement without the parenthetical is administered, key results are insignificant.

Additionally, to provide a reference point for the magnitude of the effect of CSR on employee unethical behavior, we empirically examine the effect of requiring an employee to sign an honor code on misreporting and shirking behavior. Ethics codes have become more common in corporations, in part due to stakeholder pressure to curb organizations' unethical behavior (Stevens et al., 2004). Based on self-reported survey data, Weeks and Nantel (1992) suggest that corporate codes of ethics are correlated with ethical sales force behavior; similarly, McCabe, Trevino, and Butterfield (1996) find that unethical behavior is lower at organizations with corporate codes of conduct. Shu et al. (2012) use lab and field experiments to demonstrate that increasing the salience of moral standards by signing an honor code reduces cheating behavior in the lab and reduces consumers' misreporting to an insurance company. We would expect a similar effect in the domain of employee behavior: that increasing the salience of moral standards by requiring employees to sign an honor code would reduce unethical behavior on the job.

STUDY 1: CSR AND STATED LIKELIHOOD OF MISCONDUCT AMONG FULL-TIME U.S. EMPLOYEES

Experimental Design and Sample Characteristics

In Study 1, we surveyed roughly 200 full-time U.S. employees from a wide range of industries via Qualtrics Panel.[5] Our two-by-two design included a social mission (CSR) treatment and an honor code (HC) treatment. All survey participants received the following job description (intended to mirror the situation that real MTurk workers faced in Study 2):

> *Imagine you begin a job working for a startup eyeglasses manufacturer. Your job is to make market research calls on behalf of this startup; your boss provides you with a list of phone numbers and indicates that you are to call each number and record answers to a series of questions about eyeglasses if someone answers the phone and is willing to participate.*
>
> *Your pay is structured as follows:*

[5] We instructed Qualtrics Panel to recruit 200 participants, and they exceeded this request with 210. We dropped 9 participants from our analysis for either failing to complete the study in full or failing to select the correct answer to a basic attention check question. IRB approval was obtained.

You receive a flat payment (agreed upon in advance) if you submit a report indicating that you have attempted all phone calls as instructed. You receive this payment regardless of call outcomes. You receive a bonus payment for each call in which you report speaking with a person and obtaining responses to the market research questions.

Your boss has no way of observing whether or not you actually attempt these calls or speak with potential market research survey participants.

All participants also received two sentences describing the employer's mission. Those in the CSR treatment group received the message, "This startup company's motto is '**buy one, give one**.**' For every pair of glasses that the company sells, it gives one pair of eyeglasses to a child in need in a developing country**." The remaining participants received the message, "This startup company's motto is that '**two pairs of glasses are better than one**.' **For every pair of glasses bought, the customer gets a second pair of eyeglasses free**." In addition, those in the HC treatment group received a message stating, "**Also imagine that you are asked to sign an honor code** when you start this job attesting that you pledge to report your call outcomes accurately and honestly." After reading this initial content, participants were asked two key questions:

1) *How likely do you think it is that you would report attempting some or all phone calls as instructed even if you hadn't actually attempted the calls?*

2) *How likely do you think it is that you would report successfully obtaining one or more market research survey responses (in order to claim a bonus payment) even if you hadn't actually obtained such responses?*

Potential responses were presented on a Likert scale with seven options ranging from "extremely unlikely" to "extremely likely." After answering these questions pertaining to potential misreporting, participants were asked to complete a brief demographic survey, which asked them about gender, income, and whether or not they volunteered and/or donated to charity. Finally, participants were asked several questions aimed at assessing potential mechanisms for any observed effect on stated likelihood of

misreporting. On a Likert scale with seven options ranging from "strongly disagree" to "strongly agree," respondents were asked to assess their level of (dis)agreement with the following statements:

1) *I feel that this employer is doing good.*
2) *I feel that I am doing good by working for this employer.*
3) *I identify with this employer.*
4) *I would feel bad lying to this employer.*
5) *I would feel bad cheating this employer out of money.*

As a baseline, the first statement is included to assess whether the CSR messages were, indeed, interpreted by the participants as prosocial behavior by the employer. The remaining statements are designed to test the relevance of several potential theories for why CSR and signing an honor code might influence misreporting. Specifically, the second statement was included to test for the presence of moral licensing in response to CSR, and the third to test whether CSR messaging caused the participants to identify more strongly with the employer. The fourth and fifth statements were included to assess the extent to which our treatments increase individuals' sense of moral obligation towards the employer, as manifested by feeling bad about lying (4) or cheating the employer out of money (5).

Table 1 reports sample characteristics by condition and indicates that randomization across observable characteristics was generally successful. There were more females in the "Only CSR" condition compared to the control group, those in the "Only HC" condition had slightly higher income, and those in the "Only CSR" condition were more likely to have donated in the past. Characteristics between the treatment groups and control group were otherwise statistically equivalent. We include control variables for all observable demographic characteristics in our regression specifications.

Insert Table 1 Here

Results

Figures 1-4 present the distributions of responses to our two key misreporting questions by treatment group. The first notable observation from Figures 1 and 2 is that the proportion of respondents who answer that they would be "extremely unlikely" to misreport the completion of any phone calls is substantially greater in both the CSR and HC treatment conditions. In Figure 2, those who received the honor code treatment were also less likely to state that they were "extremely likely" to misreport. In contrast, the effect of social responsibility in Figure 1 is more complicated. Of those respondents who received the CSR treatment, a greater proportion say that they would be "extremely unlikely" to misreport calls completed, but a greater proportion also say they would be "extremely *likely*" to misreport. In other words, the CSR treatment increased outcomes at both extremes. This is also true in Figure 3, which compares responses regarding the likelihood of (illegitimately) claiming a bonus by CSR treatment.

Insert Figures 1-4 Here

One explanation for why this might occur is that some workers may actually have *negative* reactions to our social mission treatment that subsequently cause them to be more (rather than less) likely to misreport. For example, it has been shown that if employees perceive CSR initiatives as being used instrumentally for the purpose of profiting the firm, CSR can backfire in the form of *reduced* effort provision (Casssar and Meier, 2017). Because we capture survey responses on whether individuals agree with the statement "I feel that this employer is doing good," we are able to explore this possibility in our empirical analysis in a fairly direct way.

Insert Tables 2 and 3 Here

In baseline regression results presented in Table 2 (Models 1-2 and 4-5), we observe a negative effect of honor code on stated likelihood of misreporting (as expected), but no effect of CSR. This is, perhaps, unsurprising given the heterogeneous way in which CSR seems to affect misreporting, as depicted in

Figures 1 and 3. We thus explore the possibility that an important determinant of the direction of CSR's effect is the extent to which an individual agrees with the social mission itself (as proxied by agreement with the statement "I feel that this employer is doing good"). Table 3 presents preliminary evidence that this may, indeed, be the case. Respondents in aggregate were more likely to agree that the socially conscious firm was "doing good," but among those who answered that they would be moderately or extremely likely to misreport outcomes, this relationship is reversed: within this subset, individuals in the CSR treatment group are actually *less* likely to agree that the firm is doing good. Accordingly, we introduce an interaction term in Models 3 and 6 of Table 2 between CSR and the variable *agrees_emp_good*, defined to indicate whether or not an individual agrees with the statement "I believe that this employer is doing good." Once this term is introduced, results look quite different: if a participant receives the CSR treatment but does not agree that the firm is doing good, she is substantially *more* likely to misreport (coefficients of 1.21 and 1.01, respectively, both significant at the 5% level), but if she receives the CSR treatment and does agree that the firm is doing good, she is *less* likely to misreport (coefficients of -1.65 and -1.29, significant at the 1% and 5% levels, respectively).[6] Notably, summing relevant coefficients yields a total effect of of -0.44 (in Model 3) when an individual receives the CSR treatment and agrees that the firm is doing good; this effect is similar in magnitude to the effect of the honor code.

It seems essential, then, that employees believe in the benevolence of a firm's stated prosocial endeavors for CSR to have a positive effect; among those who do not believe that the firm is doing good, misbehavior actually increases. In either case, though, a key question remains: what is the mechanism through which the change in behavior takes place? Results thus far are not consistent with a moral licensing mechanism. Those who believe in the employer's benevolence should also, arguably, be the same individuals who feel good about their own prosocial contributions – subsequently giving them license to misbehave. But here, these individuals are less likely to say that they would misreport. Nevertheless, we

[6] In the absence of a CSR treatment, a belief that the employer is doing good seems to have no bearing on misreporting.

test for evidence of moral licensing, in addition to other mechanisms, directly by examining the mediation of agreement with the Likert scale statements described on page 13.

Table 4 assesses the extent to which CSR (and other variables) affect outcomes for each of the potential mediators discussed above. In Models 1 and 2, results indicate no statistically significant relationship between CSR and the respondent agreeing with the statement, "I feel that I am doing good by working for this employer." Given this result, a moral licensing mechanism seems unlikely. In Models 3 and 4, CSR does seem to affect workers' identification with the employer (either positively or negatively, depending on agreement with the statement that the firm is doing good), but statistical significance is marginal. In Model 6, however, both CSR and HC move the needle significantly with regard to workers' propensity to feel bad about lying – making this channel a good potential candidate for mediation.

Insert Table 4 Here

Table 5 confirms that, indeed, the effects of both CSR and HC are mediated by the worker feeling bad about lying (and, to a lesser extent, cheating). Here, Models 1 and 4 replicate Models 3 and 6 from Table 2. Models 2 and 5 subsequently introduce all four potential mediators; only those pertaining to feeling bad about lying or cheating are significant. Moreover, when these terms are added to the model, the coefficient estimates on both CSR terms and the HC term fall in magnitude by roughly half and lose statistical significance, indicating partial mediation (Baron and Kenny, 1986). Models 3 and 6 repeat this specification with only the mediators relating to feeling bad about lying and/or cheating, and all coefficient estimates remain essentially unchanged. This provides suggestive evidence that both the CSR and HC treatment effects on unethical behavior are driven by workers feeling bad about lying to their employer, consistent with a moral obligation mechanism. We find no support that identification with the employer mediates the (stated) likelihood of misreporting, nor do we find any support for a moral licensing mechanism.

Insert Table 5 Here

STUDY 2: CSR AND ACTUAL MISCONDUCT AMONG MTURK GIG WORKERS

Experimental Design and Sample Characteristics

Study 2 was conducted on MTurk over three weeks during February and March 2018.[7] Acting as a hiring employer, we advertised a job requiring workers to make five market research phone calls and answer a short survey (see Figure 5 for a screenshot of the actual posting). To maximize employee recruitment, we utilized both automatic reposting (offered as a feature through MTurk) and manual reposting of the job every 1-3 hours daily. We also varied the advertised payment each week: $1 in the first week, $2 in the second week, and $3 in the third week. Employees were hired and completed the job on a rolling basis.

Insert Figure 5 Here

Figure 6 presents a flow chart illustrating the timing of all key events in the experiment, from job posting through payment. Workers interested in the job could click on the link included in our MTurk posting to receive further instructions before deciding whether or not to accept the contract. To accept and begin the job, workers were required to provide their MTurk ID number. Note that neither the initial MTurk posting nor the preliminary instructions on our external job site contained any mention of either social mission or honor code; these treatment conditions were assigned only after a worker committed to the job so as to prevent any potential selection effects. 100% of the 697 workers who accepted the job "completed" it, in that they filled in responses (either honestly or dishonestly) to all required fields. Whether or not employees actually completed the phone calls as instructed is a separate question, and one of the key dependent variables studied in this analysis; we paid all 697 workers in full, however, regardless of their completion of the task as instructed.

[7] IRB approval was obtained for this study. The study was also preregistered, and the preregistration is available here: https://aspredicted.org/blind.php?x=y6n2th.

Upon accepting the job, each employee was randomly assigned to one of four conditions in a two-by-two design: control (no honor code, no CSR), only honor code, only CSR, or honor code plus CSR. The two treatment dimensions were actualized as illustrated in Figure 6. An honor code ("HC") was implemented by having participants write their name below a statement reading: "I pledge to report call outcomes accurately and honestly." The CSR condition was incorporated both in an initial description of the employer (a fictional company that we will refer to as "XYZ Eyewear") and in the call script that workers were asked to read.[8] Specifically, the socially-oriented firm "believes in giving back: for every pair of glasses that we sell, we give one pair of eyeglasses to a child in need in a developing country." This language corresponds to a "buy-one give-one" model, which is commonly used in practice (Marquis and Park, 2014). In contrast, the non-social firm "believes that two pairs of glasses are better than one, so we offer customers a buy-one, get-one-free deal on all frames."

Insert Figure 6 Here

To complete their task, workers in all groups were provided with two scripts: one to read if a call recipient answered, and one to read if a call rang to voicemail (see Figure 6 for details). In practice, obtaining responses to the market research questions was impossible. We purchased all the phone numbers listed in the task instructions, and during the course of this experiment, ensured that no calls were actually answered.[9] Employees following instructions should have thus made five phone calls, left five voicemails, and reported their call outcomes as such. There are, however, at least two ways in which workers may be motivated to misreport in this setting. First, workers were unaware that their employer could monitor whether or not they actually attempted to make the calls, leave voicemails as instructed, etc. Under these

[8] The name of the fictional company is available from the authors upon request. It consisted of three letters followed by "Eyewear." If workers were to have googled the name of the company, they would have found a website that indicated that it was under construction.

[9] We also ensured that all voicemail boxes were properly configured and that the lines were never busy; the only potential truthful call outcome for a diligent employee to report, then, was that they had left a voicemail as instructed.

circumstances, workers had a fairly strong incentive to misreport by checking the box to indicate that they had attempted a call even when they actually hadn't. Secondly, employees had an incentive to report that they had obtained one or more responses to the market research survey (even though obtaining such responses was impossible), since they were awarded with a bonus of $0.25 for each response that they claimed. We examine both the prevalence and magnitude of these two types of misreporting (the first opportunity to misreport is arguably less egregious than the second) as our primary dependent variables.

Following completion of the market research phone calls but prior to receiving the code needed for payment, workers were also asked to fill out a short demographic survey asking them about gender, education, and income. The survey also asked workers about their primary motivation for working on MTurk (specifically, money versus other factors)[10] and whether they volunteer and/or donate to charity. Among all workers, the median time spent on the job was 10.4 minutes.

Table 6 provides a summary of sample characteristics by condition (note that of the 697 total workers hired, 6 were dropped from the study for failing to select the correct answer to a basic attention check question). As Table 6 illustrates, the sample is reasonably well-balanced, although the second treatment group (CSR, no honor code) has a proportion of female participants that is statistically different from the control group at the 5% level. All other differences between groups are either insignificant or only marginally significant. In regression analysis, controls are included for all variables listed in Table 6.

Insert Table 6 Here

As discussed above, a unique feature of our experimental setting is that we can cleanly observe instances of employee misreporting. In particular, we study instances of *voicemails misreported* and *bonuses misreported*. The latter was straightforward to quantify, as it could be obtained directly from the

[10] As indicated in Table 6, the majority of workers stated that their primary motivation was to earn money.

employees' call outcome report.[11] Quantifying the former involved a slightly more complex process. Upon hiring, in addition to being assigned a treatment condition, each employee was also assigned a unique random number. This number was inserted into their voicemail script as the extension number to which they were to direct potential survey respondents (see Figure 6). Accordingly, research assistants coding these voicemails could identify which employee had left each message. This allowed us to ascertain the number of voicemails that each employee actually completed as instructed.[12]

How do these two types of misreporting differ? From the employee's perspective, claiming a bonus is arguably the more egregious type of lie. This is because claiming a bonus requires concocting a complete set of responses to our market research questions. Of the three market research questions, the first two (pertaining to whether or not members of the household wear glasses and if so how much these glasses typically cost) were presented as multiple choice. The third question, however, regarding what brand(s) of frames members of the household wear, was presented simply as a blank text field that required an input – arguably a more difficult type of response to fabricate. In instances where employees did claim a bonus, their responses to this third question are rather interesting. Some simply enter responses like "n/a," but the majority of responses are fairly creative, and often quite detailed (e.g., one respondent wrote, "Whatever brand I am able to get for cheap. Kids break glasses so easily that I am not interested in spending a lot.").

Results

Table 7 presents the overall incidence of the two types of misreporting described in the preceding section. More than 60% of workers engaged in at least one instance of misreporting. This primarily involved misreporting voicemails (i.e., claiming that a task had been completed when it actually had not). Notably, among employees who misreported at least one voicemail, roughly half also claimed at least one bonus;

[11] Since, by design, we ensured that no one answered the calls, any reporting of survey responses completed (and thus, bonuses owed) were misreports.

[12] By matching call times with survey completion times, we are also able to infer instances where employees attempted a call but opted not to leave a voicemail, although this occurred fairly rarely (if a call was attempted, a voicemail was also left 93.6% of the time). In empirical analyses, we focus on voicemails left rather than calls attempted, as voicemails were the task that was assigned. All key results, however, are substantively robust if the dependent variable is based on calls attempted rather than voicemails left.

among employees who did not misreport any voicemails, virtually none claimed a bonus. In other words, conditional upon honestly completing the task as instructed, workers were extremely unlikely to falsely claim a bonus. That the two types of misreporting are highly correlated has several potential interpretations. The most plausible, perhaps, is that some workers are inherently "honest types," and are unlikely to engage in either type of misreporting, while "dishonest types" are more likely to engage in both types of misreporting.

Insert Table 7 Here

Figures 7-10 break out the incidence of both types of misreporting by CSR and HC treatment. Both the CSR and HC treatments reduce the number of voicemails misreported (Wilcoxon rank-sum tests yield p-values of 0.07 and 0.04, respectively), although Figures 7 and 8 illustrate that this effect manifests in different ways. Specifically, for CSR, the largest shift in the distribution occurs at the two extremes: when working for a firm with a social mission, fewer employees misreport all five voicemails, and more employees do not misreport at all. In contrast, for HC, the largest shift occurs at the low end of the distribution: when asked to sign an honor code, a large number of workers appear to move from misreporting one voicemail to misreporting none. Very little change occurs at the high end of the distribution in this case (i.e., roughly the same proportion of workers misreport all five voicemails in both groups).

Insert Figures 7-10 Here

Notably, the patterns in Figures 7 and 8 are not replicated in Figures 9 and 10, which illustrate each treatment's effect on the more egregious measure of misreporting, number of bonuses (illegitimately) claimed. For the CSR treatment, there is very little difference in the distributions. For the HC treatment, the number of bonuses claimed actually appears to *increase* slightly when workers are asked to sign an honor

code. We believe the most plausible explanation for this is that in this setting, our honor code pledge may increase the salience of the opportunity to misreport – thus increasing actual misreporting. This trend, however, may also just be noise, as the relationship does not turn out to be robustly significant in regression analysis.

Tables 8 and 9 present results for the effect of our social responsibility and honor code treatments on voicemails misreported. CSR does, indeed, reduce misreporting: specifically, workers misreport 0.37 fewer voicemails on average (significant at the 5% level once controls are added in Model 2 of Table 9) when their employer touts a social mission. The utilization of an honor code also reduces misreporting, by a slightly smaller amount – 0.28 voicemails on average (at a level close to statistical significance with a p-value of 0.11 in Model 2). However, as both Table 8 and Table 9's Model 3 suggest, these main-effect estimates may be somewhat understated, as estimates rise substantially in both magnitude and significance (to 0.65 and 0.54 fewer voicemails misreported for CSR and HC, respectively) when an interaction term (CSR x HC) is added.

Insert Tables 8 and 9 Here

A positive (and nearly significant; p-value of 0.12) coefficient estimate for this interaction term that is similar in magnitude to both main effect estimates indicates that the effects of CSR and HC are not additive – in other words, there is little benefit to employing both. Results are substantively similar in Models 4 and 5 when the dependent variable is a binary indicator for whether a worker misreported any voicemails. Here again, both CSR and HC reduce the probability that a worker will misreport, but when both treatments are employed simultaneously, the effects are not (fully) additive, as indicated by the positive and statistically significant coefficient on the interaction term. CSR and HC can thus be viewed as substitutes, since utilizing an honor code in addition to CSR offers minimal incremental benefit. In all of our specifications, the magnitude of CSR's effect on misreporting is similar to that of an honor code, and

both are substantial in terms of meaningful impact. For example, in Model 3, CSR decreases misreporting by 26% versus a baseline level of 2.5 voicemails misreported on average among all workers.

Several coefficient estimates on our set of control variables are also significant and worth mentioning. First, the strongest predictor of misreporting in our model appears to be gender, with females misreporting substantially less than males. This is consistent with other work which has shown women to report more ethical attitudes (see Borkowski and Ugras, 1998, for a meta-analysis), to report less favorable attitudes toward cheating (Whitley, Nelson and Jones, 1999), and to be less likely to lie for monetary benefit (Dreber and Johannsson, 2008). Education does not seem to predict misreporting. Lower incomes are correlated with less misreporting, and workers who state that the money they earn on MTurk is an important source of income are also less likely to misreport. This makes sense if we view these variables as proxies for how important the job is to the worker (and, correspondingly, how important a worker's MTurk reputation and employer approval ratings are to her).

The positive and significant coefficient on our indicator for whether a worker volunteers for non-profit / charitable organizations is notable, suggesting that those who volunteer are *more* likely to misreport. While at first this result may seem counterintuitive, it is in line with the view that employees engage in moral licensing: here, individuals who volunteer (i.e., behave prosocially in one domain) appear to feel increased authorization to misreport (i.e., to behave unethically in another domain). We emphasize that this result should be interpreted with caution, as there are other potential explanations. Most notably, our survey questions were administered *after* employees had submitted their outputs for the market research task. It is plausible, then, that causation might run in the opposite direction. Perhaps those who engaged in misreporting on the task itself may have been more likely to report that they volunteered if doing so made them feel better about having behaved unethically. (Alternatively, individuals who are more likely to lie about call outcomes may also be more likely to lie about other things such as volunteer work.)

Turning now to Table 10, we also examine the effects of our treatments on bonuses (illegitimately) claimed. Here, however, there appears to be no relationship between either CSR or HC and misreporting. Given that this is the more egregious form of misreporting from the employee's perspective, it is possible

that only certain "dishonest types" in the population are willing to engage in more egregious forms of misreporting and that treatments have little effect on these individuals.

Insert Table 10 Here

Several control variables are again worth discussing. Consistent with the results in Table 9, females are less likely to misreport, and those with higher incomes are more likely to misreport. Perhaps most interestingly, workers who donate to charity are *more* likely to illegitimately claim a bonus, though it is important to emphasize that the statistical significance is only marginal. This finding, however, mirrors that in Table 9 where workers who were also volunteers tended to be more likely to misreport. Here, workers who donate are more likely to falsely lay claim to bonuses – again potentially indicative of moral licensing. Moreover, in both cases, the moral licensing seems to take place along corresponding dimensions (in the first case, effort given then dishonestly shirked, and in the second case, money given then fraudulently claimed).

DISCUSSION OF RESULTS AND LIMITATIONS

Results from our field experiment on MTurk, where we are able to study misreporting via revealed preference/behavior, shed light on just how prevalent employee misconduct may be in settings where effort and outcomes are not (from the workers' perspective) observable to the employer. In aggregate, roughly 45% of workers on this job reported that they had completed all five calls as instructed without actually having attempted *any*. Even if rates of adverse behavior are lower in more traditional organizational settings, our results suggest that managers should exercise caution when assigning workers tasks where effort and outcomes are unverifiable. Additionally, our MTurk study indicates that there are important differences in the extent to which employers have the ability to mitigate various types of misreporting. Specifically, neither CSR nor the implementation of an honor code had any effect on more egregious misconduct (here, illegitimately claiming a bonus). One explanation for this finding is that there may simply be some "dishonest types" in the population who will engage in unethical behavior regardless of any

intervention. If this is the case, then being able to identify these individuals is critical for firms – as simply not hiring them in the first place may be the only way to avoid the consequences of their misconduct.

There are a few differences in findings between our two studies that are worth discussing. Perhaps most notably, some participants actually seem to react *negatively* to the CSR treatment in Study 1; we do not observe this in Study 2 on MTurk. Moreover, we do not observe any meaningful variation in results for the two different types of misreporting in Study 1 (whereas on MTurk, workers were much more likely to engage in the less egregious form of misreporting). These sorts of incongruities may be attributable to the fact that likely outcomes in Study 1 are merely stated via survey as opposed to observed naturally; they may also be due to differences in the two sample populations (full-time U.S. workers vs. MTurk workers).

Certainly our paper is not without limitations. As with all field experiments conducted on MTurk, the extent to which we can generalize findings to broader organizational contexts is uncertain. In combination with Study 1, however, where our sample consisted of full-time U.S. employees, we believe that the robustness of our most important finding – that CSR reduces misreporting by roughly the same amount as an honor code – is compelling. Moreover, given the challenge of observing employee misconduct in practice (as well as that of establishing a causal relationship between organization-level characteristics and individual-level misconduct), we believe our field experimental design, which elicits employees' unethical behavior in their natural MTurk work context, to be an important contribution.

CONCLUSION

Our paper sheds light on the effectiveness of CSR as a potential governance tool for guarding against shirking and other forms of adverse employee behavior. Despite empirical evidence that firms are, indeed, *attempting* to utilize CSR in this manner (Flammer and Luo, 2017), the direct relationship between CSR and employee misconduct has been understudied to date. We present robust causal evidence that CSR substantially reduces misreporting among employees. Moreover, we provide suggestive evidence that a moral obligation mechanism explains the reduction of employee unethical behavior in response to CSR, and we find no evidence that CSR elicits moral licensing (contrary to List and Momeni's working paper,

2017). We do, however, find some evidence that moral licensing may be relevant for other reasons. Specifically, our results suggest that individuals who (self-report that they) volunteer with charities are more likely to lie to their employers about effort expended on the job, and that employees who (self-report that they) donate money to charity are more likely to lie to their employers to falsely lay claim to monetary bonuses. This suggests that individual-level "good behavior" may indeed cause individuals to feel license to behave badly and that, furthermore, this moral licensing tends to take place along the same dimension as the preceding "good behavior." In contrast, our finding that CSR causes a decrease, rather than an increase, in employee misconduct suggests that organizational-level or employer-level "good behavior" does not cause individuals to feel license to behave badly in the same way that individual-level "good behavior" does.

We also highlight an important boundary condition of the benefits of CSR: workers must believe that the firm is, in fact, "doing good" for CSR to have a positive effect on employee behavior. This finding is in line with that of Cassar and Meier (2017), who demonstrate that CSR can have a demotivational effect on employees when it is framed as an activity intended to boost profits. Our work suggests that, even if not explicitly framed as profit-motivated, CSR may be perceived negatively by some employees – who are then *more* likely to engage in misconduct. Future work might explore the importance of perceived authenticity of CSR or individual-level characteristics that lead individuals to be more or less skeptical of CSR programs.

More generally, our main findings, taken in concert with those of List and Momeni's working paper (2017), suggest that there may be substantial variation with respect to the way in which different types of CSR affect employee misconduct (and likely employee behavior more broadly). The way in which CSR is implemented and communicated in practice can, seemingly, completely reverse the direction of its impact. A more thorough understanding of this nuance is critical for scholars, who have, to date, tended to lump many different types of prosocial activities together under the broader category of "CSR" (Burbano, Mamer, and Snyder, 2018). Clearly, though, not all CSR activities have the same impact. For example, our findings suggest that when it is the good that the *organization* is doing that is emphasized (as we argue to

be the case for the vast majority of CSR activity in practice), then CSR does not elicit moral licensing and tends to decrease misconduct. If, however, CSR is explicitly framed to make the *individual* feel as though she, herself, is doing good, then this may, indeed, cause the employee to feel license to behave unethically. Notable examples of this type of CSR include corporate volunteer programs and employee gift matching. While more research is needed to understand the complex relationship between CSR, moral licensing, and unethical behavior in the workplace, our results indicate that CSR is more likely to *decrease* (rather than increase) unethical behavior on the part of workers.

REFERENCES

Ashforth BE, Mael F (1989) Social identity theory and the organization. *Academy of Management Review* 14(1): 20–39.

Balafoutas L, Beck A, Kerschebamer R, Sutter M (2013) What drives taxi drivers? A field experiment on fraud in a market for credence goods. *Review of Economic Studies,* 80(3): 876-891.

Balasubramanian P, Bennett VM, Pierce L (2017) The wages of dishonesty: The supply of cheating under high-powered incentives. *Journal of Economic Behavior & Organization*, *137*, 428–444.

Baron RM, Kenny DA (1986) The moderator—mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51(6):1173–1182.

Bateman T, Organ DW (1983) Job satisfaction and the good soldier: the relationship between affect and employee 'citizenship.' *Academy of Management Journal* 26(4): 587–595.

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

Benabou R, Tirole J (2010) Individual and corporate social responsibility. *Economica* 77(305), 1-19.

Bode C, Singh J, Rogan M (2015) Corporate social initiatives and employee retention. *Organization Science* 26(6): 1702-1720.

Borkowski SC, Ugras YJ (1998) Business students and ethics: A meta-analysis. *Journal of Business Ethics*, 17, 1117-1127.

Brass, DJ, Butterfield JD, Skaggs BC (1998) Relationships and unethical behavior: a social network perspective. *The Academy of Management Review,* 23(1): 14-31.

Brief AP, Motowidlo SJ (1986) Prosocial organizational behaviors. *Academy of Management Review* 11(4): 710–725.

Brockner J, Senior D, Welch W (2013) Corporate volunteerism, the experience of self-integrity, and organizational commitment: evidence from the field. *Social Justice Research* 27(1):1–23.

Burbano V (2016) Social responsibility messages and worker wage requirements: field experimental evidence from online labor marketplaces. *Organization Science*, 27(4): 1010-1028.

Burbano V (2018) Getting gig workers to do well by doing good: Field experimental evidence from online labor marketplaces. Columbia Business School Working Paper.

Burbano V, Mamer J, Snyder J (2018) Pro bono as a human capital learning and screening mechanism: Evidence from law firms. *Strategic Management Journal* 39: 2899-2920.

Carnahan S, Kryscynski D, Olson D (2017) When does corporate social responsibility reduce employee turnover? Evidence from attorneys before and after 9/11. *Academy of Management Journal* 60(5): 1932-1962.

Cassar L, Meier S (2017) Intentions for doing good matter for doing well: The (negative) signaling value of prosocial incentives. NBER Working Paper 24109.
DeHoratius N, and Raman A (2008) Inventory record inaccuracy: An empirical analysis. *Management Science*, 54:627-641.
Detert JR, Trevino LK, Burris ER, Andiappan M (2007) Managerial modes of influence and counterproductivity in organizations: A longitudinal business-unit-level investigation. *Journal of Applied Psychology*, 92: 993-1005.
Dreber A, Johannesson M (2008) Gender differences in deception. *Economics Letters*, 99, 197-199.
Duflo E, Hanna R, Ryan S (2012) Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4): 1241-1278.
Dutton J, Dukerich J (1991) Keeping an eye on the mirror: image and identity in organizational adaptation. *The Academy of Management Journal* 34(3): 517–554.
Dutton J, Dukerich J, Harquail C, (1994) Organizational images and member identification. *Administrative Science Quarterly* 39(2): 239–263.
Dye R (1984) The trouble with tournaments. *Economic Inquiry*, 22(1), 147-149.
Edelman B, Larkin I (2015) Social comparisons and deception across workplace hierarchies: Field and experimental evidence. *Organization Science, 26*(1), 78–98.
Flammer C, Luo J (2015) Corporate social responsibility as an employee governance tool: evidence from a quasi-experiment. *Strategic Management Journal* 38:163-183.
Flory J, Leibbrandt A, List, J (2016) The effects of wage contracts on workplace misbehaviors: Evidence from a call center natural field experiment. NBER Working Paper 22342.
Ford R, Richardson W (1994) Ethical decision making: A review of the empirical literature. *Journal of Business Ethics*. 13: 205-221.
Gino F, Margolis J (2011) Bringing ethics into focus: how regulatory focus and risk preferences influence (un)ethical behavior. *Organizational Behavior and Human Decision Processes*.
Greening DW, Turban DB (2000) Corporate social performance as a competitive advantage in attracting a quality workforce. *Business & Society* 39(3): 254–280.
Hansen SD, Dunford BB, Boss AD, Boss RW, Angermeier I (2011) Corporate social responsibility and the benefits of employee trust: A cross-disciplinary perspective. *Journal of Business Ethics* 102(1): 29-45.
Hegarty WH, Sims HP (1978) Some determinants of unethical decision behavior: An experiment. *Journal of Applied Psychology*, 63: 451-457.
Hubbard TN (2000) The demand for monitoring technologies: The case of trucking. *Quarterly Journal of Economics*, 115:533-560.
Illies R, Scott B, Judge T (2006) The interactive effects of personal traits and experienced states on intraindividual patterns of citizenship behavior. *Academy of Management Journal,* 49(3): 561–575.
Jones T, Ryan LV (1997) The link between ethical judgement and action in organizations: A moral approbation approach. *Organization Science.* 663-680
Kilduff G, Galinsky A, Gallo E, Reade, J (2015) Whatever it takes to win: rivalry increases unethical behavior. *Academy of Management Journal,* 59(5): 1508-1534.
Kohlberg L,Turiel E (1973) Moralization: The cognitive developmental approach. New York: Holt, Rinehart & Winston.
Konrad, KA (2000) Sabotage in rent-seeking contests. *Journal of Law, Economics, and Organization*, 16(1), 155-165.
Lazear EP, Rosen S (1981) Rank-Order tournaments as optimum labor contracts. *The Journal of Political Economy*, 841-864.
List JA, Momeni F (2017) When corporate social responsibility backfires: theory and evidence from a natural field experiment. NBER Working Paper 24169.
Mael F, Ashforth BE (1992) Alumni and their alma mater: A partial test of the reformulated model organizational identification. *Journal of Organizational Behavior,* 13(2): 103-123.

Marquis C, Park A (2014) Inside the Buy-One Give-One Model. Stanford Social Innovation Review, Winter Edition, 28-33.
Mayer D, Aquino K, Greenbaum R, Kuenzi M (2013) Who displays ethical leadership, and why does it matter? An examination of antecedents and consequences of ethical leadership. *The Academy of Management Journal* 55(1): 151-171.
Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* 45(6): 633-644.
McCabe D, Trevino LK, Butterfield, KD (1996) The influence of collegiate and corporate codes of conduct on ethics-related behavior in the workplace. *Business Ethics Quarterly*, 6: 461-476.
Merritt A, Effron D, Monin B (2010) Moral self-licensing: When being good frees us to be bad empirical demonstrations of moral self-licensing. *Social and Personality Psychology Compass*, 4(5), 344-357.
Murphy KR (1993) *Honesty in the workplace*. Belmont, CA: Brooks/Cole.
Nagin D, Rebitzer J, Sanders S, Taylow, L (2002) Monitoring, motivation, and management: the determinants of opportunistic behavior in a field experiment. *American Economic Review*, 92 (4):850-873.
Olken B (2007) Monitoring corrupton: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115: 200-249.
Organ DW, Ryan, K (1995) A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology,* 48(4): 775–802.
O'Reilly C, Chatman J (1986) Organizational commitment and psychological attachment: the effects of compliance, identification and internationalization of prosocial behavior. *Journal of Applied Psychology,* 71(3):492-499.
Pierce L, Balasubramanian P (2015) Behavioral field evidence on psychological and social factors in dishonesty and misconduct. *Current Opinion in Psychology*, *6*, 70–76.
Pierce L, Snow D, and McAfee A (2015) Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science, 61*(10), 2299-2319.
Pierce L, Snyder, J (2008) Ethical spillovers in firms: Evidence from vehicle emissions testing. *Management Science, 54*(11), 1891–1903.
Rupp DE, Ganapathi J, Aguilera RV, Williams, CA (2006) Employee reactions to corporate social responsibility: an organizational justice framework. *Journal of Organizational Behavior*, *27*(4), 537–543.
Rupp DE, Shao R, Thornton MA, Skarlicki, DP (2013) Applicants' and employees' reactions to corporate social responsibility: the moderating effects of first-party justice perceptions and moral identity. *Personnel Psychology* 66(4): 895–933.
Schweitzer M, Ordóñez L, Douma B (2017) Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47(3).
Shu LL, Mazar N, Gino F, Ariely D, Bazerman, MH (2012) Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences*, *109*(38), 15197–15200.
Shulman T (2007) Biting the hand that feeds: The employee theft epidemic. The Kentucky CPA Journal, Fall Issue.
Stevens JM, Steensma HK, Harrison DA, Cochran PL (2004) Symbolic or substantive document? The influence of ethics codes on financial executives' decisions. *Strategic Management Journal*, *26*(2), 181–195.
Tonin M, Vlassopoulos M (2014) Corporate philanthropy and productivity: evidence from an online real effort experiment. *Management Science,* 61(8): 1795–1811.
Trevino LK, Youngblood SA (1990) Bad apples in bad barrels: A causal analysis of ethical decision-making behavior. *Journal of Applied Psychology*, 75: 378-385.
Turban DB, Greening DW (1997) Corporate social performance and organizational attractiveness to prospective employees. *Academy of Management Journal* 40(3):

658–672.
Umpress E, Bingham J (2011) When employees do bad things for good reasons: Examining unethical pro-social organizational behaviors. *Organization Science* 22(3):621-640.
Victor B, Cullen JB (1988) The organizational bases of ethical work climates. *Administrative Science Quarterly*, 33: 101-125.
Weeks WA, Nantel J (1992) Corporate codes of ethics and sales force behavior: a case study. *Journal of Business Ethics* 11:753-760.
Whitley BE Jr, Nelson AB, Jones CJ (1999) Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis. Sex Roles, 41, 657-680.

1 Tables and Figures

1.1 Tables

Table 1: Sample Characteristics by Condition
(Study 1: Full-Time U.S. Employees)

	Full Sample	**Control**	**Only HC**	**Only CSR**	**HC + CSR**
N	*201*	*51*	*47*	*50*	*53*
female	0.69	0.63	0.72	**0.84*** **	0.58
	(0.46)	(0.49)	(0.45)	(0.37)	(0.50)
income	67.17	60.45	**74.24***	73.37	61.53
	(50.94)	(40.33)	(56.47)	(63.83)	(39.95)
volunteer	0.22	0.25	0.19	0.20	0.25
	(0.42)	(0.44)	(0.40)	(0.40)	(0.43)
donate	0.40	0.33	0.40	**0.52****	0.36
	(0.49)	(0.48)	(0.50)	(0.50)	(0.48)

*Note: All variables are binary except **income**, which is expressed in thousands of dollars. Standard deviations reported in parentheses. Bold figures indicate sample means that are statistically different from the control group. (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$)*

Table 2: Effect of Social Mission and Honor Code on Perceived Likelihood of Misreporting

	Calls			**Bonuses**		
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
csr	−0.05	−0.06	1.21 ∗ ∗	0.02	−0.00	1.01 ∗ ∗
	(0.31)	(0.31)	(0.49)	(0.31)	(0.30)	(0.50)
csr * agrees_emp_good			−1.65∗∗∗			−1.29 ∗ ∗
			(0.62)			(0.63)
hc	−0.52∗	−0.54∗	−0.53∗	−0.51∗	−0.51∗	−0.48
	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)
female		−0.43	−0.48		−0.35	−0.38
		(0.32)	(0.31)		(0.31)	(0.31)
income		−0.00	−0.00		−0.00	−0.00
		(0.00)	(0.00)		(0.00)	(0.00)
volunteer		0.36	0.27		0.16	0.07
		(0.40)	(0.37)		(0.38)	(0.37)
donate		0.37	0.41		0.60∗	0.64∗
		(0.35)	(0.34)		(0.35)	(0.34)
agrees_emp_good			−0.02			−0.23
			(0.43)			(0.43)
Constant	3.65∗∗∗	3.88∗∗∗	3.89∗∗∗	3.53∗∗∗	3.69∗∗∗	3.82∗∗∗
	(0.25)	(0.38)	(0.39)	(0.26)	(0.38)	(0.42)
Observations	201	201	201	201	201	201
R-squared	0.01	0.03	0.09	0.01	0.04	0.08

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

*Note: In all regressions, the dependent variable is expressed on a scale from 1-7, where 1="Extremely **unlikely** to misreport" and 7="Extremely **likely** to misreport"*

Table 3: Percent of Respondents Who Agree That "This Employer is Doing Good"

	social = 0	*social* = 1	Overall
*Among Respondents "Moderately" or "Extremely" **Likely** to Misreport*	72.2%	59.1%	**65.0%**
*Among Respondents "Moderately" or "Extremely" **Unikely** to Misreport*	70.0%	89.6%	**80.7%**
All Respondents	**64.3%**	**76.7%**	**70.6%**

Table 4: Analysis of Potential Mechanisms

	"I feel that I am doing good by working for this employer"		**"I identify with this employer"**		**"I would feel bad lying to this employer"**		**"I would feel bad cheating this employer out of money"**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS
csr	0.27	−0.30	0.09	−0.70*	0.34	−0.59	0.10	−0.76
	(0.23)	(0.43)	(0.24)	(0.39)	(0.23)	(0.48)	(0.22)	(0.52)
csr * agrees_emp_good		0.44		0.79*		1.10**		1.00*
		(0.47)		(0.45)		(0.53)		(0.57)
hc	0.45*	0.22	0.09	−0.09	0.52**	0.43**	0.45**	0.35
	(0.24)	(0.19)	(0.24)	(0.20)	(0.23)	(0.22)	(0.22)	(0.21)
female	0.10	0.09	0.00	0.00	0.50*	0.52**	0.15	0.16
	(0.26)	(0.20)	(0.26)	(0.22)	(0.28)	(0.26)	(0.23)	(0.24)
income	−0.00	−0.00	0.00	−0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
volunteer	−0.09	0.08	−0.37	−0.21	−0.13	−0.02	−0.13	−0.02
	(0.32)	(0.24)	(0.32)	(0.26)	(0.31)	(0.26)	(0.30)	(0.24)
donate	0.16	0.02	0.22	0.10	−0.10	−0.17	−0.16	−0.23
	(0.27)	(0.20)	(0.27)	(0.23)	(0.27)	(0.24)	(0.27)	(0.24)
agrees_emp_good		2.09***		1.67***		0.78**		0.85**
		(0.32)		(0.31)		(0.38)		(0.34)
Constant	4.86***	3.68***	4.81***	3.87***	4.98***	4.54***	5.48***	5.01***
	(0.32)	(0.29)	(0.29)	(0.29)	(0.36)	(0.38)	(0.33)	(0.35)
Observations	201	201	201	201	201	201	201	201
R-squared	0.03	0.42	0.01	0.32	0.06	0.21	0.03	0.19

Robust standard errors in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: In all regressions, the dependent variable is expressed on a scale from 1-7, where 1="Strongly disagree" and 7="Strongly agree"

Table 5: Mediation of Treatment Effects

	Calls			**Bonuses**		
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
csr	1.21 * *	0.74*	0.74*	1.01 * *	0.67	0.64
	(0.49)	(0.43)	(0.43)	(0.50)	(0.46)	(0.46)
csr * agrees_emp_good	−1.65***	−0.88	−0.87	−1.29 * *	−0.66	−0.62
	(0.62)	(0.57)	(0.56)	(0.63)	(0.59)	(0.59)
hc	−0.53*	−0.22	−0.23	−0.48	−0.22	−0.22
	(0.31)	(0.29)	(0.29)	(0.31)	(0.30)	(0.30)
female	−0.48	−0.18	−0.18	−0.38	−0.08	−0.08
	(0.31)	(0.28)	(0.28)	(0.31)	(0.29)	(0.29)
income	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
volunteer	0.27	0.26	0.26	0.07	0.06	0.06
	(0.37)	(0.33)	(0.33)	(0.37)	(0.34)	(0.34)
donate	0.41	0.27	0.27	0.64*	0.49	0.53
	(0.34)	(0.32)	(0.32)	(0.34)	(0.32)	(0.32)
agrees_emp_good	−0.02	0.58	0.57	−0.23	−0.19	0.25
	(0.43)	(0.36)	(0.36)	(0.43)	(0.42)	(0.39)
feel_doing_good		−0.02			0.19	
		(0.12)			(0.12)	
identify		0.01			0.08	
		(0.11)			(0.11)	
feel_bad_lying		−0.49***	−0.49***		−0.57***	−0.55***
		(0.15)	(0.15)		(0.16)	(0.15)
feel_bad_cheating		−0.25*	−0.25*		−0.15	−0.06
		(0.15)	(0.15)		(0.16)	(0.15)
Constant	3.89***	7.34***	7.34***	3.82***	6.18***	6.63***
	(0.39)	(0.71)	(0.62)	(0.42)	(0.70)	(0.63)
Observations	201	201	201	201	201	201
R-squared	0.09	0.29	0.29	0.08	0.25	0.24

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

*Note: In all regressions, the dependent variable is expressed on a scale from 1-7, where 1="Extremely **unlikely** to misreport" and 7="Extremely **likely** to misreport"*

Table 6: Sample Characteristics by Condition
(Study 2: Gig Workers on MTurk)

	Full Sample	**Control**	**Only HC**	**Only CSR**	**HC + CSR**
N	*691*	*178*	*172*	*169*	*172*
Female	0.61	0.58	0.60	**0.68****	0.57
	(0.49)	(0.50)	(0.49)	(0.47)	(0.50)
College	0.39	0.38	0.40	0.39	0.40
	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)
Income > $40K	0.53	0.47	0.52	**0.56***	**0.56***
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Income is Motivation for MTurk Work	0.68	0.70	0.64	0.70	0.68
	(0.47)	(0.46)	(0.48)	(0.46)	(0.47)
Volunteer	0.31	0.34	0.28	0.30	0.34
	(0.46)	(0.47)	(0.45)	(0.46)	(0.47)
Donate	0.44	0.46	0.45	**0.39***	0.45
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)
Payment	2.08	2.04	2.13	**2.16***	1.99
	(0.79)	(0.82)	(0.78)	(0.77)	(0.79)

*Note: All variables are binary except **payment**, which is equal to either 1, 2, or 3. Standard deviations reported in parentheses. Bold figures indicate sample means that are statistically different from the control group. (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$)*

Table 7: Aggregate Incidence of Misreporting at the Worker Level

	No Bonuses Misreported	**At Least One Bonus Misreported**	**Total**
No VMs Misreported	267 *38.6%*	3 *0.4%*	**270** ***39.1%***
At Least One VM Misreported	215 *31.1%*	206 *29.8%*	**421** ***60.9%***
Total	**482** ***69.8%***	**209** ***30.2%***	**691** ***100.0%***

Table 8: Voicemails Misreported by Treatment Group

	csr=0	**csr=1**
hc=0	2.94 *(0.17)*	**2.30***** *(0.17)*
hc=1	**2.44**** *(0.18)*	**2.42**** *(0.18)*

*Standard deviations in parentheses. Bold figures indicate sample means that are statistically different from the control group (top left). (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$)*

Table 9: Effect of Social Mission and Honor Code on Voicemails Misreported

	Voicemails Misreported (Count 0-5)			**Any Voicemails Misreported? (Binary)**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Logit
csr	−0.33∗	−0.37 ∗ ∗	−0.65∗∗∗	−0.17∗∗∗	−0.86∗∗∗
	(0.18)	(0.18)	(0.24)	(0.05)	(0.25)
hc	−0.19	−0.28	−0.54 ∗ ∗	−0.23∗∗∗	−1.12∗∗∗
	(0.18)	(0.18)	(0.24)	(0.05)	(0.25)
csr * hc			0.54	0.16 ∗ ∗	0.82 ∗ ∗
			(0.35)	(0.07)	(0.34)
female		−0.73∗∗∗	−0.71∗∗∗	−0.11∗∗∗	−0.53∗∗∗
		(0.18)	(0.18)	(0.04)	(0.18)
college		0.00	0.01	−0.02	−0.07
		(0.19)	(0.19)	(0.04)	(0.19)
income_gt40k		0.53∗∗∗	0.54∗∗∗	0.07∗	0.31∗
		(0.19)	(0.19)	(0.04)	(0.18)
amt_income_important		−0.48 ∗ ∗	−0.48∗∗∗	−0.07∗	−0.33∗
		(0.19)	(0.19)	(0.04)	(0.19)
volunteer		0.37∗	0.36∗	0.11∗∗∗	0.55∗∗∗
		(0.21)	(0.21)	(0.04)	(0.21)
donate		0.11	0.11	0.03	0.13
		(0.20)	(0.20)	(0.04)	(0.19)
payment		0.82	0.85	0.22 ∗ ∗	1.11∗
		(0.60)	(0.60)	(0.11)	(0.58)
Date Fixed Effects	No	Yes	Yes	Yes	Yes
Constant	2.79∗∗∗	1.46	1.52	0.34	0.83
	(0.15)	(1.30)	(1.29)	(0.25)	(0.77)
Observations	691	691	691	691	691
R-squared	0.01	0.10	0.11	0.13	NA

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10: Effect of Social Mission and Honor Code on Bonuses (Illegitimately) Claimed

	Bonuses Claimed (Count 0-5)			**Any Bonuses Claimed? (Binary)**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Logit
csr	0.05	0.07	−0.11	−0.02	−0.11
	(0.11)	(0.11)	(0.14)	(0.05)	(0.25)
hc	0.17	0.14	−0.04	0.04	0.19
	(0.11)	(0.11)	(0.14)	(0.05)	(0.25)
csr * hc			0.36∗	0.05	0.23
			(0.22)	(0.07)	(0.35)
female		−0.26 ∗ ∗	−0.25 ∗ ∗	−0.11∗∗∗	−0.53∗∗∗
		(0.11)	(0.11)	(0.04)	(0.18)
college		−0.17	−0.17	0.00	0.02
		(0.11)	(0.11)	(0.04)	(0.19)
income_gt40k		0.24 ∗ ∗	0.25 ∗ ∗	0.07 ∗ ∗	0.36 ∗ ∗
		(0.11)	(0.11)	(0.04)	(0.18)
amt_income_important		−0.17	−0.17	−0.06	−0.30
		(0.12)	(0.12)	(0.04)	(0.19)
volunteer		0.08	0.07	0.03	0.13
		(0.14)	(0.14)	(0.04)	(0.20)
donate		0.21∗	0.21∗	0.02	0.13
		(0.13)	(0.13)	(0.04)	(0.19)
payment		0.39	0.41	0.08	0.52
		(0.37)	(0.37)	(0.14)	(0.71)
Date Fixed Effects	No	Yes	Yes	Yes	Yes
Constant	0.66∗∗∗	−0.04	−0.00	0.16	−0.55
	(0.09)	(0.82)	(0.80)	(0.30)	(0.86)
Observations	691	691	691	691	691
R-squared	0.00	0.06	0.06	0.06	NA

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

1.2 Figures

Figure 1: Perceived Likelihood of Misreporting Calls by Social Mission Condition

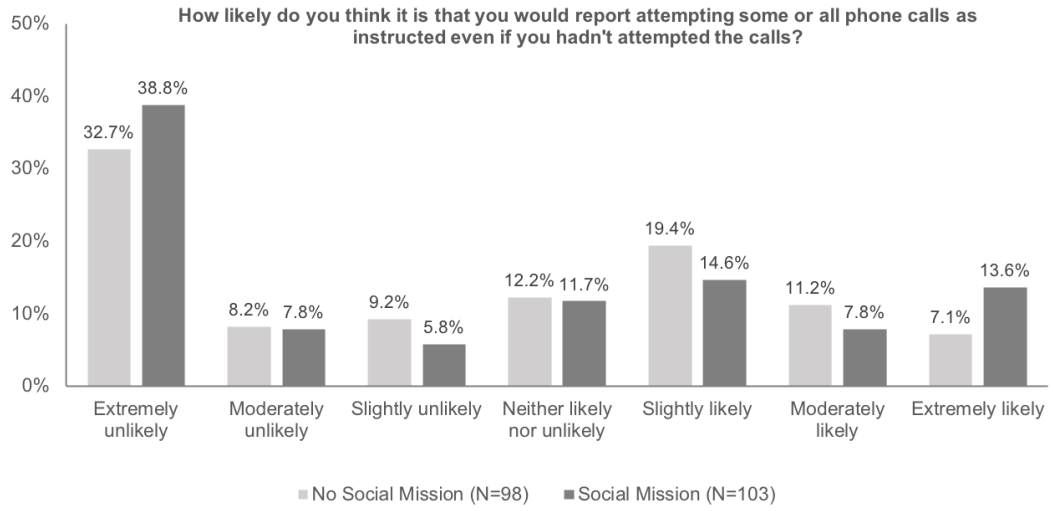

Figure 2: Perceived Likelihood of Misreporting Calls by Honor Code Condition

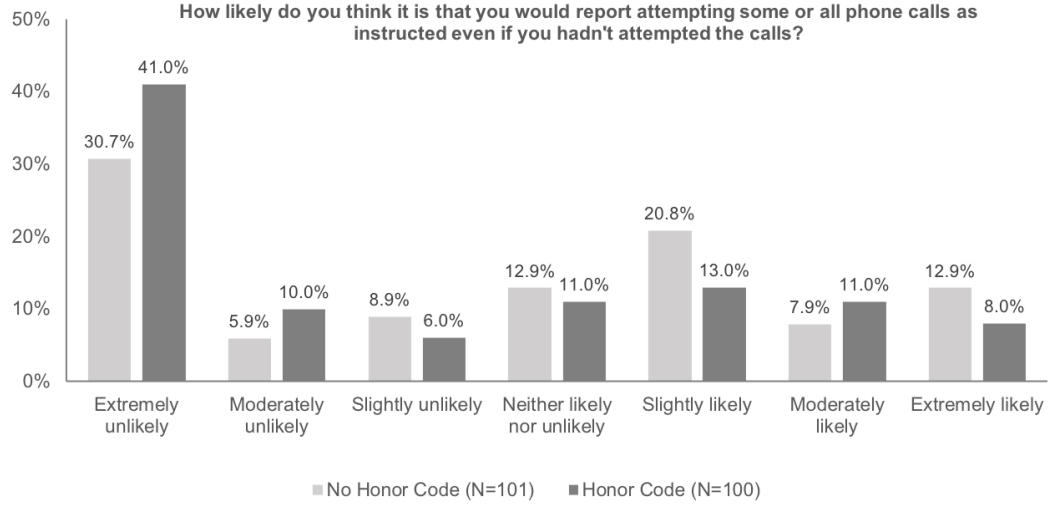

Figure 3: Perceived Likelihood of Claiming a Bonus by Social Mission Condition

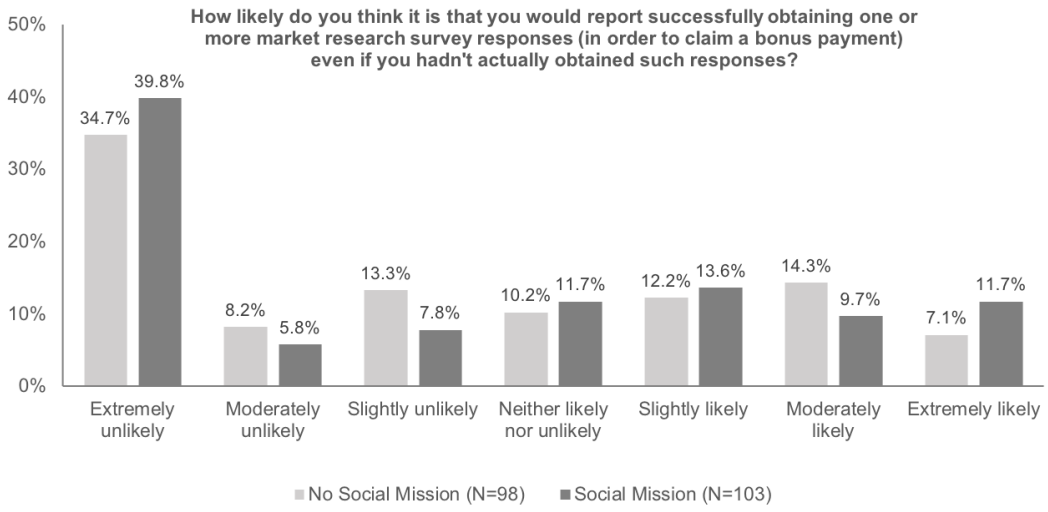

Figure 4: Perceived Likelihood of Claiming a Bonus by Honor Code Condition

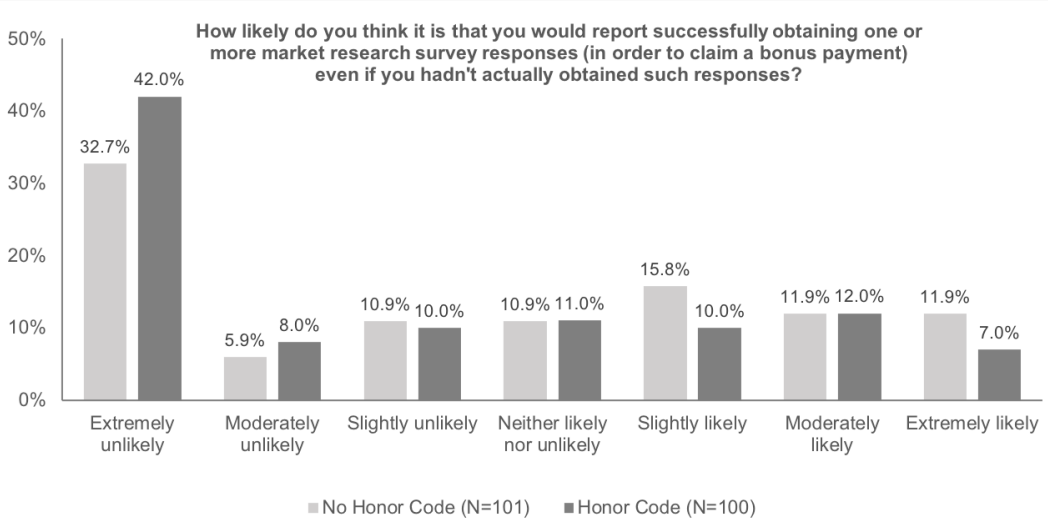

Figure 5: MTurk Job Description

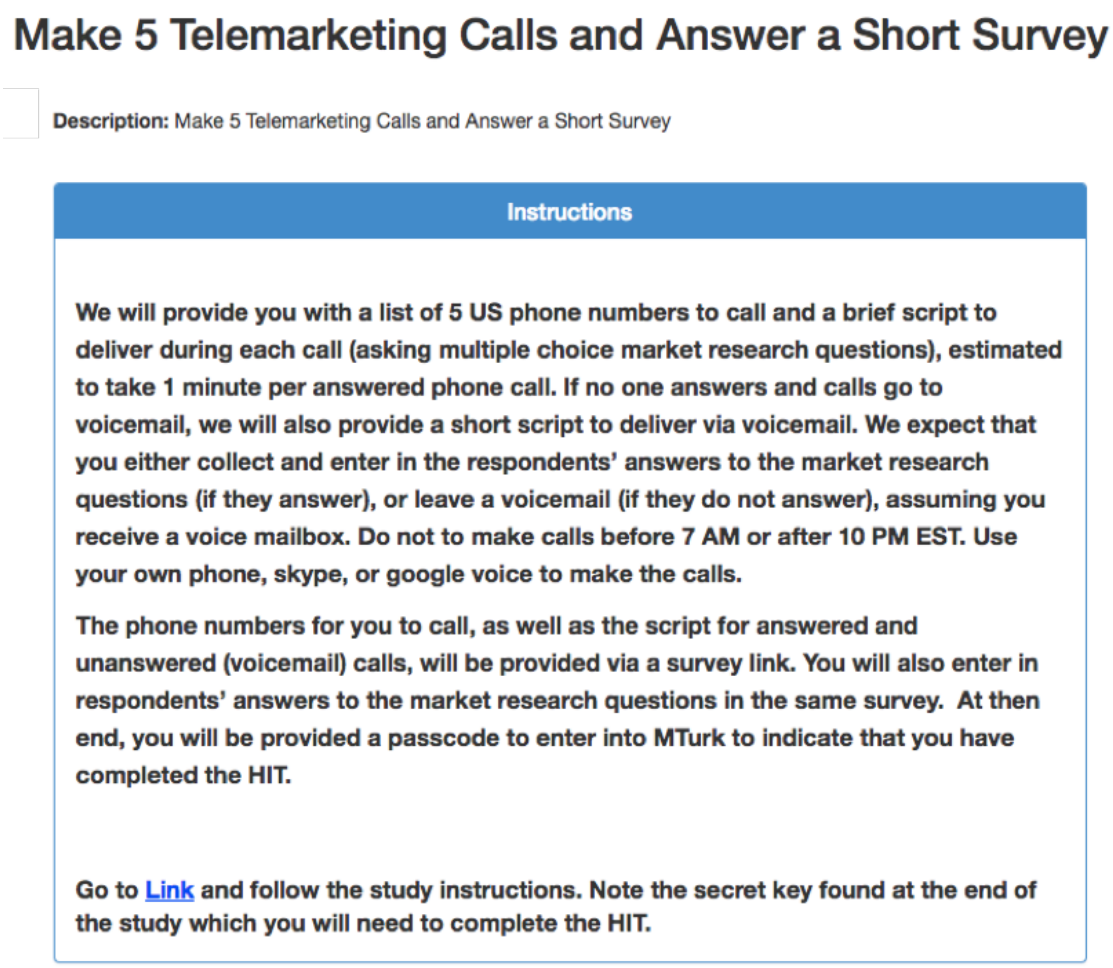

Figure 6: Experiment Design Flow Chart

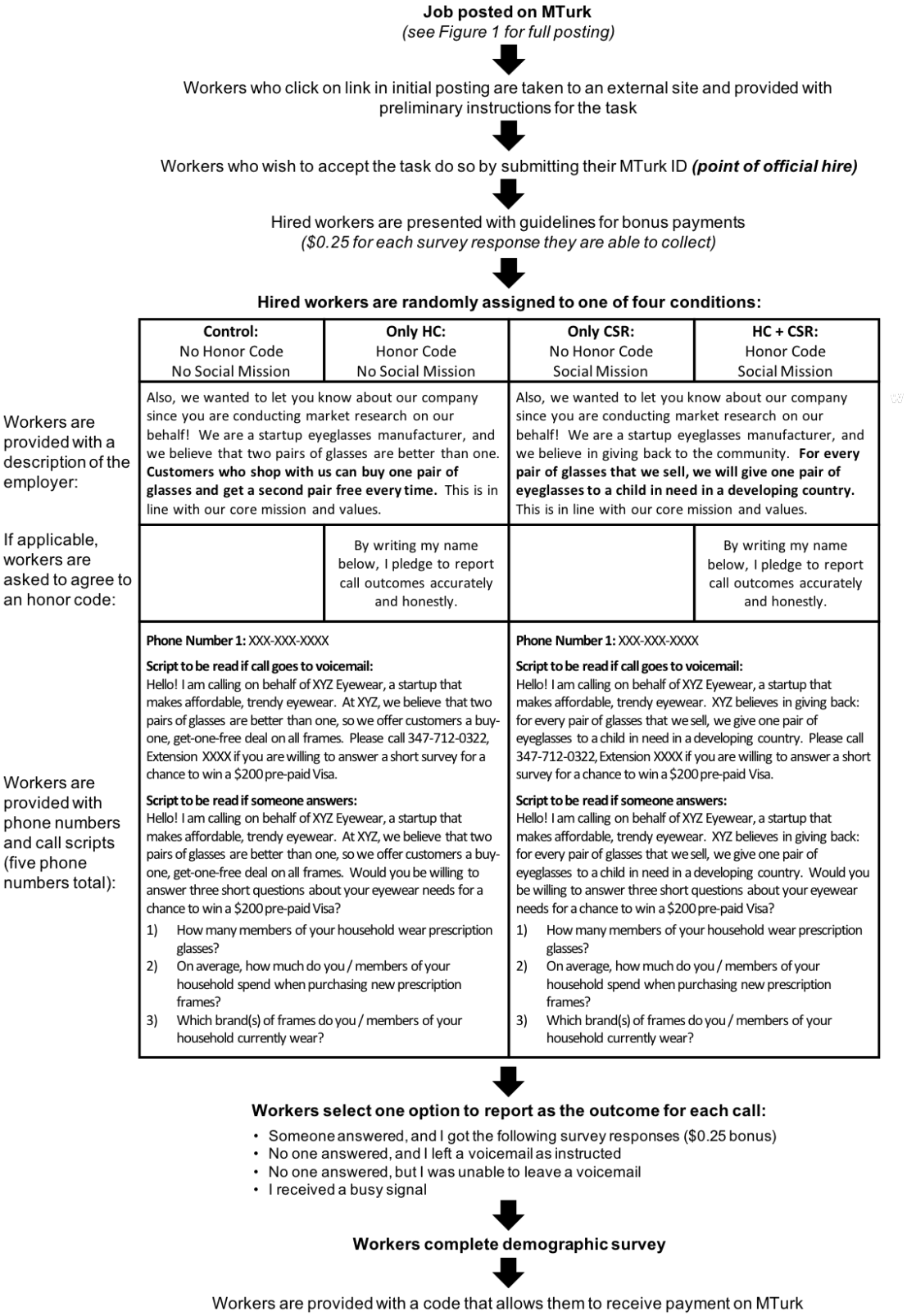

Figure 7: Voicemails Misreported by Social Mission Condition

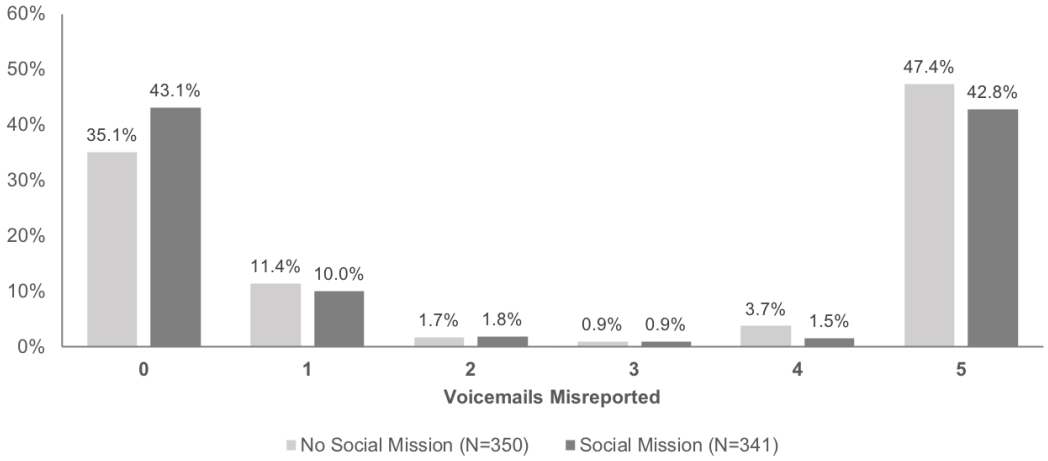

Figure 8: Voicemails Misreported by Honor Code Condition

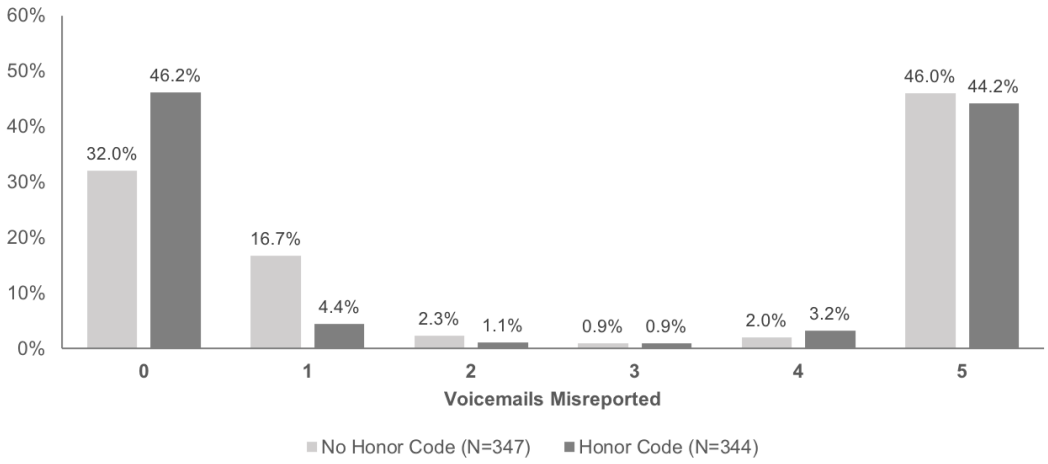

Figure 9: Bonuses (Illegitimately) Claimed by Social Mission Condition

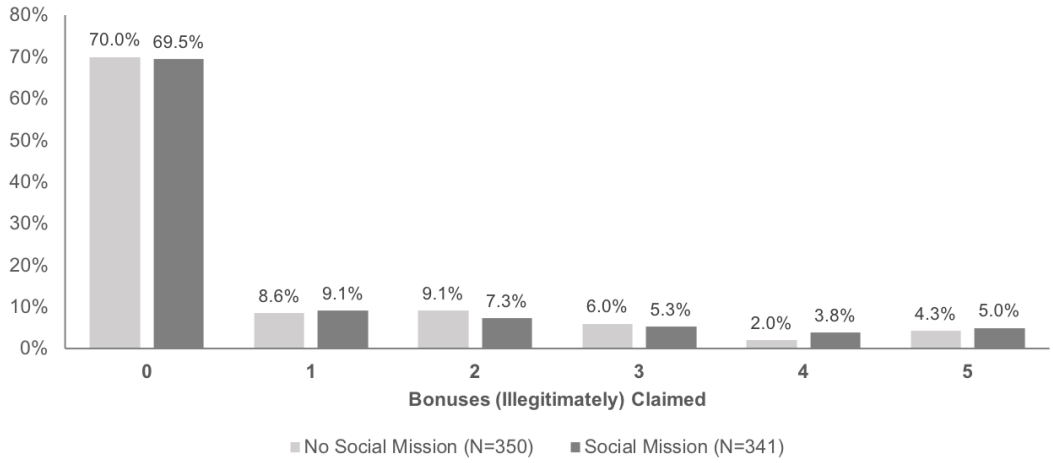

Figure 10: Bonuses (Illegitimately) Claimed by Honor Code Condition