

The Impact of Authority on Reporting Behavior, Rationalization and Affect

PAMELA R. MURPHY

Assistant Professor and E. Marie Shantz Fellow in Accounting

Queen's School of Business

Queen's University

Goodes Hall

Kingston, Ontario, Canada K7L 3N6

pmurphy@business.queensu.ca

613 533-3260

BRIAN W. MAYHEW

Associate Professor of Accounting

Accounting and Information Systems Department

University of Wisconsin-Madison

975 University Avenue

Madison, Wisconsin, USA 53706

bmayhew@bus.wisc.edu

January 11, 2013

Pam Murphy wishes to thank Queen's School of Business for its generous financial support and Fred Davis for his superb job as "boss" and for creating the lecture. The authors thank Susan Krische (Editor), two anonymous reviewers, Naman Desai, Karla Johnstone, Steve Salterio, Gregory Trompeter, workshop participants at the University of Central Florida, reviewers for the 2010 AAA annual meeting, and reviewers for the 2011 AAA audit mid-year meeting for their helpful suggestions.

Abstract

We examine reporting choices, rationalizations and emotional responses when an authority figure directs participants to misreport the results of their performance for financial gain. Our research is motivated by the assertions of several individuals involved in major accounting scandals that an authority figure instructed them to perpetrate fraudulent financial reporting. We employ a laboratory experiment where a “boss” instructs participants to misreport for financial benefit. We find that, when instructed to misreport: (1) more participants misreport, (2) they rationalize their behavior primarily by displacing responsibility, and (3) they do not feel as badly as they do when they misreport on their own volition. We find that displacing responsibility mediates the relation between being told to misreport and the act of misreporting, resulting in lower levels of negative affect. Our research addresses calls to better understand the role of rationalizations in fraudulent reporting (Hermanson 2009) and lays the groundwork necessary to explore interventions that reduce fraudulent financial reporting (AICPA 2002, Wells 2004).

Keywords: obedience, misreporting, rationalization, negative affect, moral disengagement.

1. Introduction

We conduct an experiment to examine reporting choices, rationalizations and emotional responses when an authority figure directs participants to misreport. Several accounting scandals reportedly involved an authority figure instructing subordinates to perpetrate fraudulent financial reporting. For example, Scott Sullivan, WorldCom's chief financial officer (CFO), testified in court that he was told by Bernie Ebbers, the chief executive officer (CEO), "we have to hit the numbers" (Latour and Young 2005). Subordinates of HealthSouth's Richard Scrushy claimed that he threatened their jobs if they did not alter financial information (Stuart 2005). These fraud examples are consistent with Stanley Milgram's (1974) controversial research on obedience to authority in which more than 60% of participants followed an unethical authority figure's directions.

While prior research using the theory of moral disengagement examines mechanisms individuals use to rationalize, or morally disengage from, performing an unethical act (Bandura 1991, 1999), we utilize the fraud triangle in our research setting to connect the theory of moral disengagement to the rationalization leg of the triangle. The fraud triangle suggests that three elements are necessary for fraud to occur: opportunity, motivation, and attitude/rationalization (AICPA 2002, PCAOB 2005). Our experimental reporting setting provides opportunity by allowing participants to report any income within the range of possible income. It also provides motivation by paying participants the income they report rather than the income they earn. We manipulate the setting to allow for ease of rationalization, by using an authority figure who instructs participants to misreport. When told to misreport, participants can rationalize their misreporting

behavior by displacing responsibility (Murphy and Dacin 2011). In a mixed design, we vary when the authority figure instructs participants to misreport, either on the first or second of two reporting opportunities. This design allows us to open the “black box” of whether and how an easily accessible rationalization impacts behavior and resulting emotion (affect)¹, while answering calls for research focusing on actual moral conduct and not simply moral thought (Bandura 1991). Our research addresses calls to better understand the role of rationalizations in fraudulent reporting (Hermanson 2009) and lay the groundwork necessary to explore interventions that reduce fraudulent financial reporting (AICPA 2002, Wells 2004).

In our experiment, we first establish that when participants are told to misreport, they are significantly more likely to do so. We further hypothesize and find that misreporters rationalize their decision by displacing responsibility when they are told to misreport. Specifically, misreporters in the authority treatments displace responsibility by saying “I was told to” at significantly higher rates. The use of displacing responsibility fully mediates the relation between misreport instructions and misreporting behavior, consistent with Bandura’s theory of moral disengagement. These results contribute to prior accounting literature by providing detailed insight into the cognitive processes related to honesty in reporting (Church et al. 2012, Evans et al. 2001, Hannan et al. 2006, Mayhew and Murphy 2009, Murphy 2012).

We also find that misreporters feel increased negative affect, consistent with Murphy (2012). However, negative affect is significantly *lower* for misreporters who are

¹ We use the term emotion interchangeably with affect, and we explore both negative and positive affect. Emotions are distinct from moods. While moods can impact a decision or behavior (Chung et al. 2011), emotions are the results of a decision or behavior.

instructed to misreport than those who misreport on their own volition. Thus, our research extends the theory of moral disengagement which only focuses on the anticipated affective response to behaving unethically. Our rationalization and affect results also support untested conjectures of Milgram's (1974) seminal work such as participants' experience of negative affect when following instructions.²

Finally, these findings enhance the extant understanding of the fraud triangle's less understood rationalization leg by demonstrating the impact of a specific rationalization on fraudulent reporting. We provide clear evidence that a rationalization can increase the level of misreporting, through its mediating role between misreport instructions and misreporting behavior.

The paper proceeds as follows. In the next section, we discuss relevant literature in obedience, affect, and rationalization. We then describe our experimental methodology. The results section provides an analysis of the data, and the conclusion discusses limitations and opportunities for future research.

2. Theory and Hypotheses

We first establish our expectations over changes in misreporting behavior resulting from an authority figure's instructions to misreport. Milgram (1974) finds that more than 60 percent of individuals of differing ages, genders, socioeconomic statuses, and countries of origin will administer a maximum-level electric shock to another individual when told to do so by the researcher. In Milgram's sessions, the researcher,

² Milgram did not measure affect in his original research. A recent neuroscience study uses virtual reality to replicate Milgram's paradigm in order to decipher whether the negative emotions are self-oriented or empathetic (Cheetham et al. 2009), but the study does not address affect following the behavior nor does it examine settings in which an unethical act is available with or without authority instructions.

wearing a white lab coat, serves as the authority figure. He claims to be studying the effects of punishment on learning. The participant, in the teacher role, reads each question to the learner, in an adjoining room, through a microphone. If the learner selects the wrong answer, the teacher administers an electric shock while announcing the correct answer. Each wrong answer requires a shock at the next highest shock level, in increasing increments of 15 volts up to 450 volts (which is labeled “danger, severe shock”) and a last switch that is labeled “XXX.” At the 300-volt level, the learner, who is a confederate and not actually shocked, begins to complain loudly and asks that the experiment stop. His complaints increase until eventually he stops responding altogether. If asked at any time during the experiment, the researcher indicates that the experiment must continue. The experiment ends when the participant refuses to continue or reaches the highest shock level. At that point, the researcher debriefs the participant and introduces the confederate learner.

Despite much debate about the ethics of Milgram’s methodology, the evidence clearly indicates that individuals tend to do what they are told, even if they believe it may not be the right thing to do. Milgram (1974, p. 205) summarizes his findings as follows:

the social psychology of this century reveals a major lesson: Often, it is not so much the kind of person a man is as the kind of situation in which he finds himself that determines how he will act.

Our design parallels essential elements of Milgram’s research in a corporate reporting setting in which an individual (e.g. a CFO) is told to misreport by a boss (e.g. a CEO). However, three distinct differences exist between the corporate reporting environment and Milgram’s setting, all of which tend to *increase* the likelihood that participants will follow the authority figure’s directions to misreport. First, Milgram’s

participants had little motivation to do what they were told and did not share common incentives with the authority figure. They were paid a set fee for participation and could quit at any time while keeping the money. In contrast, a CFO can extract financial gains from misreporting and shares similar incentives with his authority figure – the CEO. Both the CEO and CFO benefit from favorable reported earnings through bonuses, additional stock options/grants, and/or increases in share prices on company stock they already own. In addition, it is arguably more difficult for a CFO to walk away if s/he feels uncomfortable with unethical authoritative instructions because many of his/her financial rewards are tied to continued employment with the company. Second, the misreporting victim is usually not as physically close to participants as s/he was in Milgram's experiments. Milgram (1974) found diminishing obedience as the victim was placed closer to the participant; however, the victim was at least within earshot of the participant in all his experimental sessions. The victim is generally physically distant in a corporate reporting setting. Finally, we argue that the stress levels are lower in a corporate reporting setting compared to Milgram's shock setting. We do not dismiss the tremendous stress and fear that can result from threats to losing one's job for not following misreport orders; however, we posit it is less stressful than following orders to physically harm another individual. This lower stress level enables us to study obedience to authority in a manner acceptable to human subject internal review boards, whereas replications of Milgram's methods are essentially banned.

The above discussion suggests that the effects of authority on obedience will persist in the corporate reporting environment, essentially replicating it, which leads to our first hypothesis:

Hypothesis 1a: A higher proportion of individuals will misreport when an authority figure instructs them to misreport than when not instructed to misreport.

We also expect that once individuals receive instructions to misreport, they are more likely to misreport again, even without instructions. This is consistent with research indicating that once individuals commit fraud or fraudulent reporting, they are likely to continue (Murphy and Dacin 2011, Albrecht et al. 2012), leading to the second part of our hypothesis:

Hypothesis 1b: A higher proportion of individuals will misreport when an authority figure previously instructed them to misreport than when never instructed to misreport.

Rationalization and Negative Affect

Next, we examine the effect of authority on rationalizations and emotions. The theory of moral disengagement (Bandura 1991) provides the theoretical framework. According to Bandura, individuals learn moral principles through socialization. These principles guide behavior through a self-sanctioning regulatory mechanism. If an individual upholds society's moral principles, it brings self-worth; if an individual violates them, it brings self-condemnation and guilt. Before individuals can behave contrary to these principles, they must morally disengage from that behavior.

Moral disengagement occurs when an individual, faced with a situational inducement to behave unethically, disengages from his/her normal self-sanctioning. Broadly speaking, there are three ways in which individuals can disengage: 1) by reconstructing the act itself as being ethical or even righteous, 2) by minimizing or

disregarding the consequences of the act, and 3) by throwing responsibility elsewhere. These broad categories are essentially categories of rationalizations (Murphy and Dacin 2011).

Our experiment contains two situational inducements to misreport. First, all participants receive a financial benefit for misreporting. Second and more importantly, our treatment provides a rationalization: displacing responsibility. When individuals displace responsibility, they are throwing responsibility elsewhere. In our experiment, misreporters can argue they are not responsible because they are told to misreport. According to Bandura et al. (1996), people who displace responsibility view their actions as “springing from the social pressures or dictates of others... Because they are not the actual agents of their actions, they are spared self-censuring reactions” (Bandura et al., 1996, p. 365).

More important than hypothesizing that individuals who are told to misreport indeed misreport and displace responsibility, we examine whether this particular rationalization mediates the relation between being instructed to misreport and the act of misreporting. Though Bandura’s theory suggests this relation, much of his work has centered on identifying specific categories of disengagement and the pathways that begin with moral disengagement and end with the related behavior (Bandura et al. 1996, 2001). We are unaware of any experiments that examine situational triggers that lead to moral disengagement and, ultimately, to the related unethical behavior.³ This leads to our second hypothesis:

³ Using interview and survey methodology, Bandura and others have examined mechanisms of moral disengagement stemming from specific situational inducements, such as 9/11 (McAlister et al, 2006) and

Hypothesis 2: The association between an authority figure's instructions to misreport and participants' misreporting behavior will be mediated by the "displacing responsibility" rationalization.

Negative affect plays an important role in Bandura's theory. He refers to the regulatory system that monitors behavior as an "affective self-regulatory mechanism" as opposed to "dispassionate abstract reasoning" (Bandura 2002, p. 102). In a similar vein, breakthroughs in neuroscience over the past decade reveal a strong correlation between ethical decision-making and the part of the brain associated with negative emotion.

Damasio (2003, 2007) describes his findings as follows:

Following neurological damage to specific sectors of the prefrontal cortex, previously healthy adults who are well acquainted with social conventions and ethical rules and who are known to have behaved according to such conventions and rules in the past, now fail to observe them. They deviate, more or less significantly, from broadly accepted norms (Damasio 2007, p. 3).

These patients showed no other behavioral or cognitive deficiencies, yet they were no longer able to make ethical decisions. Other research using functional magnetic resonance imaging (fMRI) finds that this same part of the brain lights up when participants are asked to make ethical judgments (Greene et al. 2001, Hotz 2007, Koenigs et al. 2007).⁴

Based on these studies, we expect that the anticipation of negative affect plays a central role in a decision to act ethically when facing temptation. This discussion also suggests that individuals who decide to act unethically will actually experience the

executing prisoners (Osofsky et al, 2005). We are unaware of experiments that include inducements and allow participants the choice to act as they wish.

⁴ One such study correlates participants' answers to "How do you feel right now?" to the same brain images (Schaefer et al. 2002). In other words, participants reported feeling negative emotion at the same time this part of the brain lit up during ethical decision making.

negative affect they anticipated. Prior research identifies two types of negative affect associated with unethical decisions: guilt, a moral emotion (Bandura 1991, 1999; Plant and Devine 1998; Haidt 2009) and discomfort, a feeling caused by performing a counter-attitudinal behavior (Festinger 1957). Murphy (2012) finds that individuals who misreport when provided the opportunity and motive feel significantly increased levels of both guilt and discomfort afterwards.

Bandura has not studied negative affect *after* an unethical behavior. Moral disengagement theory implies that individuals either act ethically or morally disengage to act unethically, with no discernible difference in negative affect between the two afterwards. We believe it is more likely that individuals experience negative affect on a continuum where the decision to misbehave does not depend on reducing ex post negative affect to zero but to a tolerable threshold level. We argue individuals anticipate how badly they will feel if they act unethically, and will do so when that negative emotion is sufficiently low. In essence, we argue that individuals can tolerate some negative affect without imposing self-sanctions. Under this formulation, individuals who act unethically will still experience negative affect, even if they morally disengage. Because individuals seek to minimize negative affect, this formulation also leaves open the possibility that affect-based interventions that interfere with the moral disengagement process could deter misreporting.

To summarize this discussion in a reporting framework, we argue a continuum of negative affect after making a reporting decision which depends upon the individuals' decisions and whether they are told to misreport, as follows:

Hypothesis 3a: Misreporting individuals who are instructed to misreport will feel more guilt and discomfort than honest reporters.

Hypothesis 3b: Misreporting individuals who are instructed to misreport will feel less guilt and discomfort than misreporters who were not instructed to do so.

These hypotheses extend Bandura's theory by examining residual negative affect following an unethical act.

3. Research Method

We conduct an experiment using 88 participants recruited from an intermediate accounting class at a North American university. The average age of participants is 21 years. The experiment provides the opportunity and financial motivation to misreport. Participants listen to a lecture on measuring wealth and complete two different multiple-choice quizzes on the topic. Each quiz culminates in an earned income based on the participant's answers. Upon completing each quiz, participants learn their earned income and are then asked to report an income. They are paid what they report.

We use student participants because our experiment does not require task-specific knowledge, unlike more complicated accounting decisions (Arnold et al. 2004). The theories we use (i.e. moral disengagement and obedience) are both general in nature, applying to all individuals regardless of experience or expertise. Moreover, academic literature finds little or no evidence of differences in ethical attitudes or reasoning between students and professionals. Emerson et al. (2007) find that professional

accountants' ethical attitudes are not significantly different than students'.⁵ Scofield et al. (2004) find that levels of moral reasoning are not significantly different by position within accounting firms. Staff members, who are fresh out of college and therefore similar to our students, have similar moral reasoning ability as partners.

We operationalize the authority figure through role-play using a confederate "boss" who observes each participant's reported income. The boss is a Ph.D. student unknown to the participants. He wears a business suit and acts in a detached but friendly manner during all experimental sessions. He attempts to build a positive rapport with participants by providing some hints for improving performance on the quizzes.

The experiment proceeds as follows:

1. Participants enter the lab, where they sign consent forms. The researcher introduces the Ph.D. student to the participants and tells them to follow the Ph.D. student's instructions and think of him as their "boss."⁶ The researcher leaves the room, explaining that she will return to pay the participants at the end.
2. In writing and verbally, the boss explains the rules and procedures for the first part of the experiment. He explains that all participants' actions will remain confidential. For the first part of the experiment, he will show a lecture (previously recorded by the boss) on measuring wealth. After the lecture is completed, participants will take two timed quizzes that culminate in earned incomes. Each quiz consists of 10 multiple choice questions. For each question, the correct answer automatically results in a draw from a uniformly distributed income lottery with a range from \$0.50 to \$2.00. An incorrect answer automatically draws from a uniformly distributed income lottery with a range from -\$0.50 to +\$1.00 (see Appendix A).
3. The boss shows the lecture. When it is completed, he provides some hints for how participants can perform well on the quiz and instructs them to turn to the networked computer in front of them and take the first quiz.
4. Participants each turn to a networked computer and answer the timed questions. Questions are randomly drawn from a test bank, without replacement. When

⁵ However, several factors have been found to be associated with moral reasoning ability in other studies, such as gender, political orientation, education level, work experience, or age. We thank an anonymous reviewer for this insight. While we examine gender differences in our experiment (see footnote 12), we call for future research to examine the other factors.

⁶ All experimental sessions use the same individuals in the roles of researcher and boss. None of the participants were students of the researcher at the time.

- participants complete the quiz but before they learn their income, the computer instructs them to turn back to the boss for more instructions.
5. The boss explains, in writing and verbally, the process for the next part of the experiment and how everyone is paid. First, the computer will inform each of them, privately, of their income earned from the quiz. After that, each participant is to *report* an income to the boss and researcher. The computer accepts any income within the range of possibilities, from 0 to \$20.00. ***The researcher pays each participant the amount s/he reports, and pays the boss the average of all the participants' payments.*** The boss sits at the server and can see each participant's reported income only. Once the boss sees that everyone has completed their reporting task, he prints a payment report containing each participant's name and reported income only, and hands the report to the researcher in an adjoining room. The researcher prepares the payments, places them in sealed envelopes, and returns them to the boss for distribution (see Appendix B for these instructions).
 6. Participants turn back to their computers, where they answer two questions designed to ensure they understand the payment scheme.
 7. Participants are informed, privately via computer, of the income they earned from the quiz.
 8. Participants report an income through the networked computer to the boss and researcher, after which they complete a set of affect measures, with items randomized to eliminate order or demand effects.
 9. The computer reminds each participant of his/her actual and reported income, followed by "Why did you report more than you earned?" for misreporters, to gather rationalizations.
 10. After ensuring that all participants completed the questions on the computer, the boss distributes the envelopes and tells the participants that they can open the envelopes. He says the second quiz will proceed the same way as the first.
 11. Participants take the second quiz, are informed of their income, and report an income. After this, they complete the same affect measures randomized to reduce order effects and answer the question of why they reported more than they earned if they misreported.
 12. Participants complete manipulation and comprehension check questions.
 13. The researcher enters the room with the second set of sealed payment envelopes. While the boss distributes the envelopes, the researcher explains that she had instructed the boss to tell participants to misreport. She explains that the researchers are interested in what people do when told to misreport, and how they feel about it, but in a safe and confidential setting. She ensures that participants are given time to ask questions or provide comments and concerns.⁷

⁷ No participants expressed any concern about the experiment. The institutional research board and ethics board at both authors' institutions approved this protocol. Although a confederate boss is employed, all statements and instructions accurately reflect the payment scheme used in the experiment.

We randomly assign participants to one of two conditions. In condition A, the boss provides no instruction to misreport on the first quiz. On the second quiz, he instructs participants to report more than they earned (i.e., misreport), at the end of step 10. In condition B, the boss instructs participants to misreport prior to the first quiz, at the end of step 5. He makes no further mention about misreporting prior to quiz 2. In both conditions, the boss says, “I would like you to report more than you earned,” explaining that both he and the participants will be paid more. These instructions are not in writing; the boss does not use notes and maintains a neutral expression when instructing participants to misreport.

Two experimental design elements deserve explication. First, we operationalize the perceived power of the boss in the participants’ eyes. When introducing the boss, the researcher asks participants to think of him as the boss and follow his instructions. The boss himself wears a business suit and possesses the knowledge behind the wealth creation lecture. We believe these features lead participants to view the boss as a legitimate authority figure.⁸ Second, we tailor the way the boss delivers the misreport instructions to make it appear that he is not just implementing the researcher’s will but is making a rational appeal to the participants to misreport for everyone’s mutual benefit. He looks away from the written instructions when he tells them to misreport, making it appear that he is saying this on his own rather than following instructions. He makes a rational appeal for misreporting—that everyone, including him, makes more money if

⁸ Blass (2000) identifies six categories of power, many of which are thought to have played role in Milgram’s research: reward, coercive, legitimate, reference, expert and informational. We directly operationalize legitimate, expert and informational power; and we indirectly operationalize reward power. For ethical reasons, we do not operationalize coercive power, though we acknowledge it likely plays a role in a corporate misreporting context.

they misreport. We intentionally operationalize his instructions in this manner to make it appear to participants that his instructions are not part of the planned experiment, and yet the instructions make sense in terms of benefiting both participants and the boss.

Variables and Analyses

Our variables include misreporting instruction treatments, reporting behavior, negative and positive affect, and rationalization. We code the misreporting instruction treatments into three categories: 1) NOT TOLD, coded 1 for the treatment in which participants are not told to misreport, 2) TOLD, coded 1 for treatments in which participants are told to misreport, and 3) NOT RETOLD, coded 1 for the treatment in which participants were told to misreport previously, but are not retold. NOT TOLD captures observations from the first quiz of condition A (n = 45), TOLD captures observations from the second quiz of condition A and the first quiz of condition B (n = 88), while NOT RETOLD captures observations from the second quiz of condition B (n = 43).

Reporting behavior is measured in two ways. First, we code MIS as 1 if a participant's reported income is greater than his/her earned income, 0 otherwise. Second, DIFF captures the magnitude of misreporting, or the proportion of misreporting relative to the total amount possible (Evans et al. 2001), using the following formula:

$$\text{DIFF} = [(\text{reported} - \text{earned}) / 20.00 - \text{earned}] * 100.$$

The range of this variable is from 0 (honest) to 100 (misreporting to the greatest extent possible).

We measure three distinct types of affect: guilt, psychological discomfort, and positive self. All three types of emotions have been linked to unethical or counter-attitudinal behavior (Murphy and Dacin 2011, Plant and Devine 1988). We follow prior literature on administering and measuring affect (Devine et al. 1991, Monteith et al. 1993, Murphy 2012, Plant and Devine 1998) by asking participants to respond to a series of words and short phrases. When presented with a word or phrase, the computer instructs participants to answer how they feel right now, using a 7-point Likert scale from 1 (*does not apply at all*) to 7 (*applies very much*). Six words or phrases capture GUILT: annoyed at myself, disappointed with myself, guilty, regretful, self-critical, and shame (Chronbach's α of 0.86 after quiz 1 and 0.89 after quiz 2). Six words or phrases capture DISCOMFORT: bothered, distressed, negative, tense, uncomfortable, and uneasy (Chronbach's α of 0.88 after quiz 1 and 0.91 after quiz 2). The positive affect measure includes five words that capture POSITIVE SELF: friendly, content, good, happy, and optimistic (Chronbach's α of 0.84 after quiz 1 and 0.91 after quiz 2).

To identify rationalizations, we analyze and code answers to the question: "Why did you report more than you earned?". Two coders, including one of the authors, each code answers independently by rationalization category or no rationalization.⁹ Initial inter-rater reliability is 0.843. Any differences are discussed and recoded. Rationalization categories include ADVAN for advantageous comparison, MORAL for moral justification, MIN for minimizing or misconstruing the consequences, DISPLACE for displacing responsibility, and DIFFUSE for diffusing responsibility. We merge a few other rationalizations into an OTHER category.

⁹ Answers that are not rationalizations are usually either opportunity (e.g., "because I could") or motivation (e.g., "because I want the money").

To examine the first set of hypotheses, we use both logistic and linear regressions, the first with MIS as the dependent variable and the second with DIFF as the dependent variable. The models are as follows:

$$\text{MIS (or DIFF)} = \beta_0 + \beta_1 \text{ TOLD} + \beta_2 \text{ NOT RETOLD} + \beta_3 \text{ EARNED} + \varepsilon$$

Where:

MIS = 1 is misreported (reported more than earned), 0 otherwise

DIFF = the proportion of the total possible misreporting amount, given by the following equation: $[(\text{reported} - \text{earned}) / 20.00 - \text{earned}] * 100$. Range from 0 (honest) to 100 (misreporting by the greatest amount possible).

TOLD = 1 if participants are told to misreport (second quiz of condition A, first quiz of condition B), 0 otherwise

NOT RETOLD = 1 if participants are not told to misreport after being told to misreport once before (second quiz of condition B), 0 otherwise

EARNED = the amount the participant actually earned based on quiz responses.

We include EARNED as a control variable, asserting that individuals are less likely to misreport if they earn more money from the quiz. Hypothesis 1a (1b) is tested by examining the TOLD (NOT RETOLD) variable in each of the two analyses. A significant positive coefficient indicates each hypothesis is supported.

We analyze the second hypothesis by following the three-step process suggested by Baron and Kenney (1986) using logistic regression. In the first step, we regress the mediator (DISPLACE) on the independent variable (TOLD). TOLD should be significant. Second, we regress the dependent variable (MIS) on the independent variable (TOLD). TOLD should be significant, consistent with hypothesis 1a. Third, we add DISPLACE to the second equation. If DISPLACE is significant and TOLD becomes insignificant, then this suggests full mediation.

To analyze hypothesis 3a, we perform a t-test of GUILT and DISCOMFORT on misreporting individuals versus honest individuals within the TOLD condition (n = 88).

We address hypothesis 3b by analyzing the sample of misreports (n = 119) using the following equation:

$$\text{GUILT (or DISCOMFORT)} = \beta_0 + \beta_1 \text{ TOLD} + \beta_2 \text{ NOT RETOLD} + \beta_3 \text{ DIFF} + \beta_4 \text{ RATIONALIZE} + \varepsilon$$

Where:

GUILT = sum of six words or phrases, each on a 7-point Likert scale: annoyed at myself, disappointed with myself, guilty, regretful, self-critical, and shame. Range from 6 to 42; higher = more guilt.

DISCOMFORT = sum of six words, each on a 7-point Likert scale: bothered, distressed, negative, tense, uncomfortable, and uneasy. Range from 6 to 42; higher = more discomfort.

TOLD = 1 if participants are told to misreport (second quiz of condition A, first quiz of condition B), 0 otherwise

NOT RETOLD = 1 if participants are not told to misreport after being told to misreport once before (second quiz of condition B), 0 otherwise

DIFF = proportion of the total possible misreporting amount, given by the following equation: $[(\text{reported} - \text{earned}) / 20.00 - \text{earned}] * 100$. Range from 0 (honest) to 100 (misreporting by the greatest amount possible).

RATIONALIZE = 1 if participant used any rationalization category, 0 otherwise.

We include DIFF to test whether those who misreport by greater amounts feel more negative affect, and we include RATIONALIZE to test whether the use of any rationalization reduces negative affect after misreporting. A significant positive coefficient on the TOLD variable provides support for hypothesis 3b.

4. Results

Comprehension and Manipulation Checks

Before participants report their income, they answer a two-question test to determine whether they understand how they and the boss are paid. Three of the 88 participants answered both questions incorrectly. The computer explains the correct answer, so it is not clear whether these three participants understand the payment scheme.

As a robustness check, we eliminate these three participants from the dataset and rerun all tests. The results do not change.

Additionally, six debriefing statements are administered at the end of the experiment, with answers on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The first states that the participants viewed the research assistant as their boss in the experiment. The mean response was 5.22 (standard deviation of 0.503), significantly higher than the midpoint of 4 ($p = 0.000$, one-tailed), indicating participants viewed the research assistant as their boss. The second statement, “I was given the opportunity to report as much as \$20.00 on each quiz, and be paid that amount, regardless of how much I actually earned,” has a mean response of 6.68 (standard deviation of 1.034), also significantly higher than the midpoint of 4 ($p = 0.000$, one-tailed), indicating they understood there was an opportunity to misreport. The third statement says that participants were given the answers to the quiz. The mean response is 4.36, marginally significantly different than the midpoint ($p = 0.072$ one-tailed), indicating slight agreement. We view this response as acceptable since the boss provides hints on how to perform better on the quiz. The fourth and fifth statements are manipulation checks of the boss’s instructions to misreport, as follows: “I felt pressured by the boss to misreport my income on the first (second) quiz.” Mean responses to the fourth statement are 4.67 (1.71) in condition B (A); mean responses for the fifth statement are 2.79 (5.98) in condition B (A). Responses are significantly different across conditions ($p = 0.000$ for each, one-tailed) and significantly different from the midpoint ($p \leq 0.014$ for each, one-tailed) as expected. The final statement said, “Another student told me what this experiment was about before I entered this room” as a check that students are not told by prior

participants how to behave in the experiment. The mean response of 1.17, significantly different than the midpoint ($p = 0.000$ one-tailed), indicates they are not aware of the experimental setting prior to entering the room.

Descriptive Statistics

Table 1 reports descriptive statistics broken down by condition, quiz and reporting behavior. Recall that in condition A, misreport instructions are given prior to the second quiz; while in condition B, misreport instructions are provided prior to the first quiz. Table 1 reports that the lowest levels of misreporting (49 percent) occur in the absence of instructions whereas misreport instructions lead to greater levels of misreporting (76 percent on the second quiz of condition A and 72 percent on the first quiz of condition B). Misreporting levels remain relatively high, at 74 percent, when participants are not retold to misreport. The magnitude of misreporting (DIFF) is also relatively high, between 77 and 88 percent of the total possible level of misreporting. The levels of negative affect (GUILT and DISCOMFORT) differ as expected, with higher levels among misreporters than honest reports. POSITIVE SELF is also as expected, with higher levels among honest reporters and lower levels among misreporters.

[Insert Table 1 about here]

Table 2 Panel A reports the number of rationalizations used by misreporters in each of six categories: (1) advantageous comparison; (2) moral justification; (3) minimize, ignore, or misconstrue the consequences of the act; (4) displace responsibility; (5) diffuse responsibility; and (6) other. Specific examples of each rationalization category are provided in the variable definition, while Panel B reports examples of non-

rationalizations. First, a majority (between 55 percent and 91 percent) of misreporters across conditions and quizzes use rationalizations rather than other answers such as motivation (“because I wanted the money”) and opportunity (“because I could”). Second, displacing responsibility is used more than any other rationalization category when misreport instructions are given. Finally, moral justification (“I am helping the boss”) is the second most common and is the most common absent instructions to misreport.

[Insert Table 2 about here]

Figure 1 provides insight into the overall dynamics of misreporting across the two quizzes and conditions. Twenty-one participants misreport on both quizzes in condition A and 28 in condition B, suggesting that many participants misreport regardless of instructions from the authority figure. A total of eighteen individuals report honestly on both quizzes across both conditions. This result shows some participants are honest even when instructed to misreport.¹⁰ In addition, most who misreport on quiz 1 continue to misreport on quiz 2. Finally, condition A contains the largest change in reporting behavior where 13 participants, who report honestly in quiz 1, misreport on quiz 2 after authority figure instructions.

[Insert Figure 1 about here]

¹⁰ Three reports were less than \$1.00 more than earned: 0.15, 0.41, and 0.51, respectively. All three occurred in Condition B, one on the first quiz and two on the second quiz. We code these three reports as honest. In all cases, the participants indicated they wished to avoid excessive coinage, and their reported income was an even dollar amount. This explanation can be viewed as a rationalization. We believe it is more conservative to code these minor misreports as honest but note that if we did not, our results would be moderately strengthened.

Results of Hypothesis Tests

Hypothesis 1a and 1b address the frequency of misreporting when instructed. Table 3 reports the results, which have the same inference whether using logistic (Panel A) or linear regression (Panel B). TOLD is significantly positively associated with misreporting ($p = 0.002$ and 0.005 in Panel A and B, respectively).¹¹ This finding supports hypothesis 1a. NOT RETOLD is also significantly positively related to misreporting ($p = 0.006$ in Panel A and B), supporting hypothesis 1b. Results remain significant when excluding EARNED from the analysis.¹² Not only do we replicate Milgram's authority instructions, but these instructions appear to carry over even when not re-delivered.¹³

[Insert Table 3 about here]

Hypothesis 2 examines whether displacing responsibility mediates the relation between misreport instructions and misreporting behavior. Table 4 documents the three step mediation analysis where, in each step, the variables are significant in the direction posited. In step three, we document that DISPLACE is significant but TOLD is not, suggesting that it completely mediates TOLD. This result supports hypothesis 2. As an additional mediation analysis, we examine whether displacing responsibility mediates the relation between NOT RETOLD and misreporting behavior. We find that it does not,

¹¹ All our significance tests are one-tailed unless specified otherwise.

¹² We perform three additional robustness checks with no changes to our inferences. First, our experimental design allows us to compare within- and between-participant effects (Schepanski et al. 1992). We use t-tests to compare observations from the first and second quiz of condition A for within-participant tests and observations from the first quiz of condition A and first quiz of condition B for the between-participant tests. Results are consistent. Second, we examine the data for gender differences and find none. Third, we control for the repeated measures by clustering on participants (each subject provides two data points, one from each quiz). In all analyses, our inferences regarding reporting behavior, affect and rationalization do not change.

¹³ We can't disentangle whether this effect results from the participant already misreporting on quiz one or because the participant previously received instructions to misreport.

suggesting that the effect of displacing responsibility does not carry over when participants have been told to misreport once but not retold.

[Insert Table 4 about here]

We examine the impact of misreporting and being told to misreport on negative affect in hypotheses 3a and 3b. Table 5 documents the results. First, we examine whether misreporting individuals who are instructed to misreport feel more guilt and discomfort than honest reporters by performing t-tests. We examine misreports versus honest reports in the TOLD condition and, as reported in panel A, we find significantly more guilt and discomfort. Hypothesis 3a is supported. To examine differences among misreports ($n = 119$), we perform two linear regressions, one with GUILT as the dependent variable and one with DISCOMFORT as the dependent variable. As reported in Panels B and C, TOLD is significantly negatively associated with GUILT and DISCOMFORT ($p = 0.005$ and 0.009 respectively), suggesting that being told to misreport significantly reduces negative affect compared to no misreport instructions. This supports H3b.¹⁴ We also note that NOT RETOLD is significantly negatively associated with DISCOMFORT ($p = 0.032$) but not with GUILT ($p = 0.120$). This suggests that those told to misreport previously feel significantly less discomfort when misreporting again but do not feel significantly less guilt when misreporting again. This result is also consistent with the notion that guilt and discomfort are distinct but related constructs (Devine et al. 1991) having different associations with other factors (Murphy 2012). Finally, we note that RATIONALIZE is not significant for either GUILT ($p = 0.204$) or DISCOMFORT ($p =$

¹⁴ As a robustness check, we consider whether honest reporters feel more negative affect as a result of not following misreport instructions. We compare honest reporters' affect based upon whether they are instructed to misreport or not, and find no significant differences in any affect measure.

0.427), suggesting that among misreporters rationalization does not generate differences in negative affect after misreporting.¹⁵

[Insert Table 5 about here]

Discussion of Rationalization and Negative Affect

First, in addition to the displacing responsibility rationalization used most frequently by participants in our experiment, we find another authority-related rationalization not present in prior misreporting research: moral justification (Murphy 2012).¹⁶ We find that misreporters use moral justification by stating that they are helping the boss. This justification is used regardless of whether the boss provides instructions to misreport (see Table 2 Panel A). Because the boss intentionally tries to develop a neutral-to-positive relation with participants, we believe this rapport triggers participants to think of moral justifications like “I am helping the boss” for misreporting behavior. This finding has implications for corporate governance in situations where employees view their superior(s) positively or are loyal to them. For example, subordinates at WorldCom cited a sense of loyalty to their boss, Scott Sullivan, for following orders to book fraudulent accounting entries (Pulliam 2005).

Second, to further discuss the notion of a continuum of negative affect, Figure 2 visually displays negative affect based upon reporting behavior and treatment. This graph

¹⁵ As an additional test of the impact of rationalization on affect, we replace RATIONALIZE with DISPLACE in our analyses for both GUILT and DISCOMFORT, and find no significant association.

¹⁶ Our use of a “boss” across all our experimental treatments led some participants to employ a moral justification rationalization, arguing that they are “helping the boss.” Prior research found very little use of moral justification in similar settings that did not include an authority figure (Murphy 2012). Murphy (2012) finds only four instances of moral justification across three treatments, with three of those instances occurring in a single treatment that limited the potential rationalizations. We identify 22 such instances across our treatments.

is consistent with a negative affect continuum from honest reports to misreports without instructions. We see that being told to misreport reduces negative affect, consistent with hypothesis 3b, and not being re-told has the same residual effect as misreporting behavior – levels are in-between being told and not being told.

[Insert Figure 2 about here]

Third, regarding the impact of rationalization on ex post negative affect, Table 5 suggests that negative affect among misreporting participants does not differ significantly based on whether they rationalize or not. On the surface, this finding appears inconsistent with Murphy (2012) who finds that rationalizations reduce negative affect among misreporters. We believe that rationalization does play a role in reducing negative affect, but that we cannot observe its full effects due to our inability to measure the *anticipated* negative affect of a participant deciding whether or not to misreport. Recall that we argue there is a threshold of anticipated negative affect below which an individual can misreport and tolerate the negative affect without self-sanction. Either reporters may not anticipate much negative affect in the first place, and therefore opt to misreport; or they may rationalize so that the level of negative affect is reduced to a tolerable level, and then misreport. In either case, these individuals' negative affect will not look much different from other misreporters; they misreport and experience some negative affect. Murphy (2012) reports results consistent with our anticipated affect assertions. She examines reporting behavior after reminding participants of flaws in rationalizations, and finds significantly reduced misreporting. Her manipulation removes or mitigates the role of rationalization in misreporting. The results suggest that, without rationalization, participants anticipate feeling too much negative affect and instead opt to report honestly.

Together, these results expand on Bandura's initial theory. Bandura states that individuals who act unethically can experience *some* negative affect (Bandura 1999, p. 203). We document his claim of some negative affect after misreporting and provide evidence that rationalization facilitates the willingness of participants to misreport. We point to future research to address the important question of what can be done to push individuals' anticipated negative affect above their threshold, steering them to report honestly.

5. Conclusion, Limitations, and Future Research

We examine both the emotional and cognitive processing of individuals facing authority instructions to misreport, answering calls for research on rationalizations in fraudulent reporting (Hermanson 2009) and research illuminating potential interventions that might reduce fraudulent reporting (Wells 2004). Our methodology also addresses Bandura's call for research into moral actions rather than just moral thought (Bandura 2002).

We find the presence of authority instructions increases misreporting behavior, which is fully mediated by rationalizing about the boss. Misreporters tend to use a specific category of rationalization: displacing responsibility ("I was told to"). Misreporting instructions also lower negative affect. These results increase our understanding of how an authority figure's directions to misreport impact the rationalizations and resulting reporting decisions. We show that the situational inducement of being told to misreport enables individuals to mitigate the emotional cost of misreporting.

Our results are consistent with Bandura's theory of moral disengagement (Bandura 1991, 1999) while also extending it. Our finding of the mediating role of displacing responsibility supports his theory. We extend Bandura's theory by examining negative affect after the reporting decision and showing that negative affect is not eliminated for those who misreport, even when they rationalize.

Our results imply that situational inducements can be a powerful force on behavior and affect. We argue this is both good news and bad news for those charged with governance. The bad news is that an authority figure's misreport instructions allow individuals to misreport with reduced negative affect. In effect, this inducement allows individuals to avoid self-sanctioning. However, the good news is that the environment can also prevent individuals from misreporting. Efforts to intercede in the moral disengagement process could deter individuals from misreporting. We quoted Milgram earlier as saying that it is often not so much the kind of person but the kind of situation the person finds him/herself in that determines behavior (Milgram 1974, p. 205). This emphasizes the critical importance of understanding the situational pressures employees face in their daily jobs. Consistent with literature on the importance of "tone at the top" (CAQ 2010, Hunton et al. 2011, PCAOB 2007), we believe the culture and leadership of the organization can go a long way toward inducing or prohibiting unethical behavior.

We acknowledge the inherent limitations of experiments. Two limitations with respect to generalizability deserve discussion: 1) social desirability bias, and 2) whether we capture the most important elements of the reporting environment. First, any setting containing an ethical component can suffer from the social desirability bias. The social desirability bias is the tendency to overestimate (underestimate) the likelihood of

performing a desirable (undesirable) action (Chung and Monroe 2003). Many studies involving fraud or unethical behavior ask respondents what they (or someone else) would do in certain situations. Though some participants in our study may report honestly because they think it is the “right” thing to do, we mitigate this problem by allowing participants’ actions to be confidential. We argue that observing participants’ *actions* is far superior to asking them to report what they hypothetically would do. Mazar et al. (2008) find significant differences between what individuals think others will do versus what individuals actually do. Importantly, we are not trying to generalize the levels of misreporting from our experiment to the workplace. We argue that our findings of *differences* in reporting behavior, rationalization and emotion generalize to other settings.

Second, because we use students as participants in our experiment, there may be important elements of the corporate world or reporting environment we ignore. We do not operationalize several features that may play a role in whether an individual misreports when told to in the workplace. A professional holding a professional designation such as a Certified Public Accountant (CPA) or Chartered Accountant (CA) may act differently in similar situations, due to the code of professional conduct and high standards of the profession. The socialization process of becoming a CPA/CA and identifying with the associated ethics could produce a strong self-regulation mechanism within an individual when faced with a misreporting decision. Perhaps individuals holding such designations would be less likely to misreport and feel greater negative affect if they did. Work experience, age or concern for one’s long-term reputation can also cause individuals to behave differently. Implications of omitting these professional features are unclear. While we might expect individuals holding professional

designations or those concerned for their long-term reputation to misreport less frequently in general, it is not clear that they would act, rationalize or experience negative affect any differently than other individuals in similar situations.

Perhaps fear is the most important missing element in our experiment that is likely to be present in real-world situations. Many stories of misreporting under authority conditions imply that the subordinate was fearful of the superior or fearful of what the superior might do. It is difficult to operationalize fear in an experiment due to ethical human subject concerns. To some degree the lack of fear strengthens our design in that by excluding it, we are able to hold it constant across conditions and focus specifically on the authority construct in isolation. We also note that fear was not necessary to generate misreporting in our setting. The authority figure's simple suggestion to misreport for mutual gain was sufficient to enable the participants to morally disengage.

Nevertheless, considering how the individual fits into a larger, real-world organization remains a fruitful avenue for further research. For example, recent research suggests that not only individuals employ rationalization, but organizations use the language of rationalization as well. White et al. (2009) examine internal documents of several industries known for unethical practices, such as tobacco, lead and other industries that produce toxic substances. They found rationalization language in virtually all of them. This underscores the importance of thinking beyond the individual to group and organizational dynamics when examining unethical or fraudulent behavior.

Key aspects of the reporting setting likely deserve further research. Differences in fear, comfort or loyalty levels might impact individuals' reactions to authority instructions. In our setting, the boss is designed to be neutral or positive, since he

provides clues to help participants perform better on the quizzes. To examine this issue, researchers can manipulate the degree of comfort or level of loyalty individuals have toward their superior. Additionally, factors such as organizational commitment (Cohen-Charash and Spector 2001, Lord and DeZoort 2001, Schminke et al. 2005), ethical work climate (Victor and Cullen 1988), or normalization (Ashforth and Anand 2003) could all be considered. Future research could also examine reporting behavior along with the moral justification rationalization when the subordinate does not receive a financial benefit, as is often the case when lower level accountants are told to book inappropriate entries.

While the current research opens the door into the world of *how* individuals process, both cognitively and emotionally, a decision to misreport under authority pressure, we believe the ultimate goal of future research should be the development of interventions or de-biasing techniques that deter misreporting. We document the apparent ease with which misreporting individuals rationalize misreporting behavior in the presence of authority. The more we know about the causes, the better equipped we are to confront the problem. Future research should address whether the use of these rationalizations could be impeded, thus increasing anticipated negative affect and reengaging the individual's self-regulatory processes.

Appendix A: Instructions, Part I

[Read aloud by the boss after handing out hard copies to all participants.]

Thank you for your participation. This experiment should last approximately one hour. As I read through these instructions, please follow along and listen. If you have any questions, please hold them until later; you will have an opportunity to ask me privately. These instructions accurately portray what will happen in this experiment.

The researcher (Dr. Murphy) will collect all data related to your answers and decisions *after* you have left the room. She keeps those decisions and actions confidential.

I (the boss) will take 15 minutes explaining a valuation model for determining the value of publicly traded stocks. You may take notes during this lecture. When I am finished, each of you will sit at one of the networked computers in this room and take two different quizzes on the material. You will earn an income from each quiz. Each quiz has ten questions, each with three choices. The *correct* choice will result in an automatic draw from an income lottery with a range from \$0.50 to \$2.00, in increments of \$.01. An incorrect choice will result in a draw from a lower income lottery with a range from -\$0.50 to \$1.00, in increments of \$.01. The distribution of outcomes in each income lottery is a normal distribution. This means that for every *correct* answer, the most likely draw is close to \$1.25 (from a range of \$0.50 to \$2.00) and for every incorrect answer, the most likely draw is close to \$0.25 (from a range of -\$0.50 to \$1.00).

The range of possible incomes from *each* quiz is from -\$5.00 to \$20.00, with the most likely income close to \$12.50 if you get all the questions right, or close to \$2.50 if you get all the questions wrong. A negative total, though unlikely, will result in an income of zero. The range of incomes from *both* quizzes – that is, the amount of money you will walk out with – is between 0 to \$40.00.

For each quiz, you will have 4 minutes in which to complete all ten questions. A clock on the computer screen will assist you with time management. Any questions left unanswered at the end of 4 minutes will automatically draw from the lower income lottery. At the end of each quiz, you will receive further instructions.

The experiment concludes when you have completed both quizzes and have been paid.

From this point forward, please *do not communicate* with anyone other than through the computers as directed. Unfortunately, you cannot ask me a question during the lecture, as it could distract others. Please, no talking! You may keep these instructions for reference.

Appendix B: Instructions Part II

[Read aloud by the boss after handing out hard copies to all participants.]

For this part of the experiment, the computer will report to you, privately, the income you earned from the quiz, based on your answers and draws from the corresponding income lotteries. Then you will be asked to report your income. The computer accepts any number within the range of possible incomes, between 0 and \$20.00. *The researcher, Dr. Murphy, pays each participant the income s/he reports, and pays the boss, Fred, the average of all reported incomes in this group.*

Following is an example of seven participants in a group:

Participant #	1	2	3	4	5	6	7
ReportedIncome	\$6.49	\$15.16	\$20.00	\$12.45	\$0	\$8.90	\$12.25

The “Reported Income” is what each participant reported as his/her income. The following table provides more information, using the same example:

Participant #	1	2	3	4	5	6	7
Reported Income	\$6.49	\$15.16	\$20.00	\$12.45	\$0	\$8.90	\$12.25
Earned Income	\$6.49	\$10.16	\$8.98	\$12.45	\$-0.20	\$5.98	\$20.00
Payment	\$6.49	\$15.16	\$20.00	\$12.45	\$0	\$8.90	\$12.25

The “Earned Income” is the amount each participant actually earned from the quiz, based on his/her quiz answers and corresponding draws from the income lotteries. The “Payment” is the amount of money the researcher pays the participant. As you can see, each participant is paid the amount *s/he reports*.

In the above example, the *boss* would be paid the average of all reported incomes, or:

$$(6.49 + 15.16 + 20.00 + 12.45 + 0 + 8.90 + 12.25) / 7 = \mathbf{\$10.75}$$

When you turn back to your computer to continue, please follow the instructions and answer the questions as they appear on the screen. I (the boss) will sit at the server and wait for a file containing *only* participant names and reported incomes. I will deliver this information to the researcher who will pay each participant via a sealed envelope. The researcher gives the payment envelopes to me for distribution.

Remember that these instructions are accurate; they represent exactly how this experiment works. You may keep these instructions for reference. ***Please remember; no talking.***

References

- Albrecht, W. S., C. C. Albrecht, C. O Albrecht, and M. F. Zimbelman, 2012. *Fraud Examination, 4th edition*. South-Western Cengage Learning, Mason, OH.
- American Institute of Certified Public Accountants (AICPA). 2002. *SAS No. 99: Consideration of Fraud in a Financial Statement Audit*. October 2002. New York, NY.
- Arnold, V., P. A. Collier, S. A. Leech, and S. G. Sutton. 2004. Impact of Intelligent Decision Aids on Expert and Novice Decision-makers' Judgments. *Accounting and Finance* 44: 1-26.
- Ashforth, B. E. and V. Anand. 2003, The Normalization of Corruption in Organizations, *Research in Organizational Behavior* 25: 1-52.
- Bandura, A. 1991. Social Cognitive Theory of Moral Thought and Action, in W. M. Kurtines & J. L. Gewirtz (eds.), *Handbook of Moral Behavior and Development: Vol. 1 Theory*. Hillsdale, NJ: Erlbaum.
- Bandura, A., C. Barbaranelli, G. V. Caprara, and C. Pastorelli. 1996. Mechanisms of Moral Disengagement in the Exercise of Moral Agency. *Journal of Personality and Social Psychology* 71 (2): 364-374.
- Bandura, A. 1999. Moral Disengagement in the Perpetration of Inhumanities. *Personality and Social Psychology Review* 3: 193-209.
- Bandura, A., G. V. Caprara, C. Barbaranelli, C. Pastorelli, and C. Regalia. 2001. Sociocognitive Self-Regulatory Mechanisms Governing Transgressive Behavior. *Journal of Personality and Social Psychology* 80 (1): 125-135.
- Bandura, A. 2002. Selective Moral Disengagement in the Exercise of Moral Agency. *Journal of Moral Education* 31 (2): 101-119.
- Baron, R. M. and D. A. Kenny. 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.
- Blass, T. 2000. *Obedience to Authority: Current Perspectives on the Milgram Paradigm*. Mahwah, NJ: Lawrence Erlbaum.
- Center for Audit Quality (CAQ). 2010. *Deterring and Detecting Financial Reporting Fraud: A Platform for Action*. Washington DC: CAQ.

- Cheetham, M., A. F. Pedroni, A. Antley, M. Slater, and L. Jancke. 2009. Virtual Milgram: Empathic Concern or Personal Distress? Evidence from Functional MRI and Dispositional Measures. *Frontiers in Human Neuroscience* 3 (29): 1-13.
- Chung, J., J. Cohen, and G. S. Monroe. 2011. The Effect of Moods on Auditors' Inventory Valuation Decisions. Working Paper, York University, Boston College and Australian National University.
- Chung, J. and G. S. Monroe. 2003. Exploring Social Desirability Bias. *Journal of Business Ethics* 44: 291-302.
- Church, B. K., R. L. Hannan, and X. Kuang. 2012. Shared Interest and Honesty in Budget Reporting. *Accounting, Organizations and Society* 37: 155-167.
- Cohen-Charash, Y. and P. E. Spector. 2001. The Role of Justice in Organizations: A Meta-Analysis. *Organizational Behavior and Human Decision Processes* 86 (2): 278-321.
- Damasio, A. 2003. *Looking for Spinaza*. Orlando, FL: Harcourt.
- Damasio, A. 2007. Neuroscience and Ethics: Intersections. *American Journal of Bioethics* 7 (1): 3-7.
- Devine, P. G., M. J. Monteith, J. R. Zuwerink, and A. J. Elliot. 1991. Prejudice With and Without Compunction. *Journal of Personality and Social Psychology* 60 (6): 817-830.
- Emerson, T. L. N., S. J. Conroy and C. W. Stanley. 2007. Ethical Attitudes of Accountants: Recent Evidence from a Practitioners' Survey. *Journal of Business Ethics* 71: 73-87.
- Evans III, J. H., R. L. Hannan, R. Krishnan, and D. V. Moser. 2001. Honesty in Managerial Reporting. *Accounting Review* 76 (4): 537-559.
- Festinger, L. A. 1957. *A Theory of Cognitive Dissonance*. Evanston, IL: Peterson.
- Greene, J. D., R. B. Sommerville, L. E. Nystrom, J. M. Darley, and J. D. Cohen. 2001. An fMRI Investigation of Emotional Engagement in Moral Judgment, *Science* 293: 2105-2108.
- Haidt, J. 2009. *The Moral Emotions*. In R. J. Davidson, K. R. Scherer and H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. New York: Oxford University Press.

- Hannan, R. L., F. W. Rankin, and K. L. Towry. 2006. The Effect of Information Systems on Honesty in Managerial Reporting: A Behavioral Perspective. *Contemporary Accounting Review* 23 (4): 885-918.
- Hermanson, D. R. 2009. Fraudulent Financial Reporting. Speech delivered at the American Accounting Association Audit Midyear Meeting, January 17.
- Hotz, R. L. 2007. Scientists Draw Link Between Morality and Brain's Wiring. *Wall Street Journal*, May 11.
- Hunton, J. E., R. Hoitash, and J. C. Thibodeau. 2011. The Relationship Between Perceived Tone at the Top and Earnings Quality. *Contemporary Accounting Research* 28 (4): 1190-1224.
- Koenigs, M., L. Young, R. Adolphs, D. Tranel, F. Cushman, and M. Hauser. 2007. Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments. *Nature* 446: 908-911.
- Latour, A. and S. Young. 2005. WorldCom's Sullivan Says He Told CEO of Problems. *The Wall Street Journal*, February 9, 2005.
- Lord, A. T., and F. T. DeZoort. 2001. The Impact of Commitment and Moral Reasoning on Auditors' Responses to Social Influence Pressure. *Accounting, Organizations and Society* 26: 215-235.
- Mayhew, B. W., and P. R. Murphy. 2009. The Impact of Ethics Education on Reporting Behavior. *Journal of Business Ethics* 86: 397-416.
- Mazar, N., O. Amir, and D. Ariely. 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research* 45 (6): 633-644.
- McAlister, A. L., A. Bandura, and S. V. Owen. 2006. Mechanisms of Moral Disengagement in Support of Military Force: The Impact of Sept. 11. *Journal of Social and Clinical Psychology* 25 (2): 141-165.
- Milgram, S. 1974. *Obedience to Authority: An Experimental View*. New York: Harper & Row.
- Monteith, M. J., P. G. Devine, and J. R. Zuwerink. 1993. Self-Directed Versus Other-Directed Affect as a Consequence of Prejudice-Related Discrepancies. *Journal of Personality and Social Psychology* 64 (2): 198-210.

- Murphy, P. R. 2012. Attitude, Machiavellianism and the Rationalization of Misreporting. *Accounting, Organizations and Society* 37: 242-259.
- Murphy, P. R., and M. T. Dacin. 2011. Psychological Pathways to Fraud: Understanding and Preventing Fraud in Organizations, *Journal of Business Ethics* 101: 601-618.
- Osofsky, M. J., A. Bandura, and P. G. Zimbardo. 2005. The Role of Moral Disengagement in the Execution Process. *Law and Human Behavior* 29 (4): 371-393.
- Plant, E. A., and P. G. Devine. 1998. Internal and External Motivation to Respond Without Prejudice. *Journal of Personality and Social Psychology* 75 (3): 811-832.
- Public Company Accounting Oversight Board (PCAOB). 2005. *Consideration of Fraud in a Financial Statement Audit*. AU Section 316, www.pcaobus.org.
- Public Company Accounting Oversight Board (PCAOB). 2007. *Auditing Standard No. 5: An Audit of Internal Control Over Financial Reporting that is Integrated with an Audit of Financial Statements and Related Independence Rule and Conforming Amendments*. Washington, DC: PCAOB.
- Pulliam, S. 2005. Crossing the Line: At Center of Fraud, WorldCom Official Sees Life Unravel; David Myers Fell Into Despair, Then Aided Prosecutors; Planning for Prison Time; A BMW Ride at 115 MPH, *The Wall Street Journal*. March 24, 2005.
- Schaefer, S. M., D. C. Jackson, R. J. Davidson, G. K. Aguirre, D. Y. Kimberg, and S. L. Thompson-Schill. 2002. Modulation of Amygdalar Activity by the Conscious Regulation of Negative Emotion. *Journal of Cognitive Neuroscience* 14(6): 913-921.
- Schepanski, A., R. M. Tubbs, and R. A. Grimlund. 1992. Issues of Concern Regarding Within- and Between-subjects Designs in Behavioral Accounting Research. *Journal of Accounting Literature* 11: 121-149.
- Schminke, M., M. L. Ambrose and D. O. Neubaum. 2005. The Effect of Leader Moral Development on Ethical Climate and Employee Attitudes. *Organizational Behavior and Human Decision Processes* 97: 135-151.
- Scofield, S. B., T. J. Phillips Jr., and C. D. Bailey. 2004. An Empirical Reanalysis of the Selection-Socialization Hypothesis: A Research Note. *Accounting, Organizations and Society* 29: 543-563.

Stuart, A. N. 2005. Keeping Secrets. How Five CFOs Cooked the Books of HealthSouth. *CFOMagazine*, June 1, 2005.

Victor, B. and J.B. Cullen. 1988. The Organizational Bases of Ethical Work Climates, *Administrative Science Quarterly* 33: 101-125.

Wells, J. T. 2004. New Approaches for Fraud Deterrence. *Journal of Accountancy*, February 2004.

White, J., A. Bandura, and L. A. Bero. 2009. Moral Disengagement in the Corporate World. *Accountability in Research* 16: 41-74.

Table 1: Descriptive Statistics

	Quiz 1		Quiz 2	
	Misreporters	Honest Reporters	Misreporters	Honest Reporters
Condition A (<i>n</i> = 45)	(<i>n</i> = 22)	(<i>n</i> = 23)	(<i>n</i> = 34)	(<i>n</i> = 11)
% who misreport (report honestly)	49%	51%	76%	24%
DIFF	86.52 (24.70)	0 (0)	88.83 (22.14)	0 (0)
GUILT	25.05 (6.81)	13.43 (7.78)	18.85 (7.98)	11.09 (5.21)
DISCOMFORT	25.05 (8.50)	14.35 (8.64)	19.12 (8.62)	15.45 (6.27)
POSITIVE SELF	19.64 (5.46)	24.74 (6.14)	22.24 (6.54)	27.55 (5.45)
Condition B (<i>n</i> = 43)	(<i>n</i> = 31)	(<i>n</i> = 12)	(<i>n</i> = 32)	(<i>n</i> = 11)
% who misreport (report honestly)	72%	28%	74%	26%
DIFF ⁽¹⁾	77.99 (28.69)	0.15 (0.53)	87.68 (25.03)	0.94 (2.13)
GUILT	21.97 (7.34)	10.75 (4.12)	22.63 (7.78)	9.55 (4.06)
DISCOMFORT	22.00 (7.45)	16.25 (7.18)	21.22 (8.22)	10.45 (4.91)
POSITIVE SELF	23.23 (6.96)	27.00 (5.41)	22.78 (8.19)	25.45 (6.76)

Notes: The figures reported above, except for % who misreport, are the mean (standard deviation) in each cell.

⁽¹⁾ We code as honest three participants in condition B who reported less than \$1.00 more than earned, in order to be conservative and because these participants indicated they wished to avoid small change. This accounts for the small mean of DIFF in these conditions.

Variable definitions:

Condition A: instructions to misreport are provided prior to the second quiz, not the first.

Condition B: instructions to misreport are provided prior to the first quiz and not the second.

% who misreport (report honestly) = the proportion of participants in each condition and quiz who misreported versus report honestly. We code participants as misreporting when their reported income exceeds their earned income by at least \$1.00.

DIFF = the proportion of the total possible misreporting amount, given by the following equation: $[(\text{reported} - \text{earned}) / 20.00 - \text{earned}] * 100$. Range from 0 (honest) to 100 (misreporting by the greatest amount possible).

GUILT = the sum of six words or phrases, each on a 7-point Likert scale: annoyed at myself, disappointed with myself, guilty, regretful, self-critical, and shame. Range is from 6 to 42; higher = more guilt.

DISCOMFORT = sum of six words or phrases, each on a 7-point Likert scale: bothered, distressed, negative, tense, uncomfortable, and uneasy. Range is from 6 to 42; higher = more discomfort.

POSITIVE SELF = sum of five words or phrases, each on a 7-point Likert scale: friendly, content, good, happy, and optimistic. Range from 5 to 35; higher = more positive about oneself.

Table 2: Reasons Provided for Misreporting

Panel A: Descriptive Summary of Reasons Provided for Misreporting (n = 119)

	Quiz 1						Quiz 2					
A	22 misreporters 10 non-rationalizations (2 motivation, 7 opportunity, and 1 neither) 12 (55%) provided rationalization(s)						34 misreporters 3 non-rationalizations (2 opportunity, 1 neither) 31 (91%) provided rationalization(s)					
	ADVAN	MORAL	MIN	DISPLACE	DIFFUSE	OTHER	ADVAN	MORAL	MIN	DISPLACE	DIFFUSE	OTHER
	3 (20%)	6 (40%)	0 (0%)	3 (20%)	1 (7%)	2 (13%)	2 (5%)	6 (15%)	2 (5%)	26 (65%)	3 (8%)	1 (2%)
	15 rationalizations from 12 misreporters						40 rationalizations from 31 misreporters					
B	31 misreporters. 7 non-rationalizations (4 motivation, 2 opportunity, 1 both) 24 (77%) provided rationalization(s)						32 misreporters. 12 non-rationalizations (3 motivation, 7 opportunity, 2 both) 20 (63%) provided rationalization(s)					
	ADVAN	MORAL	MIN	DISPLACE	DIFFUSE	OTHER	ADVAN	MORAL	MIN	DISPLACE	DIFFUSE	OTHER
	4 (11%)	6 (17%)	4 (11%)	17 (47%)	3 (8%)	2 (6%)	3 (11%)	4 (15%)	4 (15%)	10 (37%)	3 (11%)	3 (11%)
	36 rationalizations from 24 misreporters						27 rationalizations from 20 misreporters					

Panel B: Examples of Reasons Provided that are not Rationalizations:

Non-rationalization Category	Definition	Examples from This Experiment
Motivation	A reason to misreport	“Because I wanted to make more money”; “I want to receive as much money as possible”; “To earn more money”
Opportunity	The perceived ability to misreport and not get caught	“Because I will be paid for what I report”; “The amount reported is the amount received”

Notes: Each cell of Panel A contains the number (percentage) of answers in each rationalization category. The total number of rationalizations is greater than the number of misreporters who provided rationalizations because many participants provided more than one rationalization.

Table 2: Reasons Provided for Misreporting (continued)

Variable Definitions:

Rationalization categories in Panel A, along with the definition of each and a few examples:

ADVAN = advantageous comparison. By comparing the wrongful act against a much more flagrant act, the original act looks better (e.g.: “I didn’t go for \$20 because I would feel guilty”).

MORAL = moral justification. Reprehensible acts are reconstrued as socially worthy or having a moral purpose (e.g.: “I helped my boss earn more money”).

MIN = minimize, ignore or misconstrue consequences of the act. Ignore any consequences of the act, make the consequences appear less than they are, or misconstrue the consequences (e.g.: “There are no consequences for me lying”).

DISPLACE = displace responsibility by placing it with someone else (e.g.: “The boss told us to”).

DIFFUSE = diffuse responsibility by sharing it with others (e.g.: “I assumed everyone else would report more than they earned”).

OTHER = other rationalizations, including euphemistic labeling (using convoluted verbiage to make a wrongful act sound better) and two additional categories identified in this experiment: disbelief of the experimental instructions and entitlement.

Table 3: Tests of Hypothesis 1:

H1a: A higher proportion of individuals will misreport when an authority figure instructs them to misreport than when not instructed to misreport.

H1b: A higher proportion of individuals will misreport when an authority figure previously instructed them to misreport than when never instructed to misreport.

Panel A: Logistic Regression of Misreporting on Condition

$$\text{MIS} = \beta_0 + \beta_1 \text{TOLD} + \beta_2 \text{NOT RETOLD} + \beta_3 \text{EARNED} + \varepsilon$$

Variable	Predicted Relation	Beta	Wald Coefficient	Exp	Significance (one-tailed)
Intercept	?	1.156	2.846	3.177	0.046
TOLD	+	1.179	8.946	3.253	0.002
NOT RETOLD	+	1.187	6.429	3.278	0.006
EARNED	-	-0.140	3.826	0.870	0.025

Nagelkerke R2 = 0.102

Number of observations: 176

Panel B: Linear Regression of Misreporting on Condition

$$\text{DIFF} = \beta_0 + \beta_1 \text{TOLD} + \beta_2 \text{NOT RETOLD} + \beta_3 \text{EARNED} + \varepsilon$$

Variable	Predicted Relation	Standardized Coefficient	t statistic	Significance (one-tailed)
Intercept	?		5.035	0.000
TOLD	+	0.233	2.594	0.005
NOT RETOLD	+	0.231	2.577	0.006
EARNED	-	-0.156	-2.112	0.018

Adjusted R2 = 0.050 F = 4.048 (significant 0.008) Number of observations: 176

Significance indicators:

Bold = significant at 0.05 level (one-tailed)

Variable definitions:

MIS = 1 is misreported (reported more than earned), 0 otherwise

DIFF = the proportion of the total possible misreporting amount, given by the following equation: $[(\text{reported} - \text{earned}) / (20.00 - \text{earned})] * 100$. Range from 0 (honest) to 100 (misreporting by the greatest amount possible).

TOLD = 1 if participants are told to misreport (second quiz of condition A, first quiz of condition B), 0 otherwise

NOT RETOLD = 1 if participants are not told to misreport after being told to misreport once before (second quiz of condition B), 0 otherwise

EARNED = the amount the participant actually earned based on quiz responses.

Table 4: Test of Hypothesis 2:

H2: The association between an authority figure’s instructions to misreport and participants’ misreporting behavior will be mediated by the “displacing responsibility” rationalization.

Using the entire sample (n = 176), we follow Baron and Kenny’s (1986) 3 step process:

Step 1: Regress mediator on the independent variable, using logistic regression

$$\text{DISPLACE} = \beta_0 + \beta_1 \text{TOLD} + \varepsilon$$

Variable	Predicted Relation	Beta	Wald Coefficient	Exp	Significance (one-tailed)
TOLD	+	1.707	21.469	5.513	0.000

Nagelkerke $R^2 = 0.182$

Step 2: Regress dependent variable on independent variable, using logistic regression

$$\text{MIS} = \beta_0 + \beta_1 \text{TOLD} + \varepsilon$$

Variable	Predicted Relation	Beta	Wald Coefficient	Exp	Significance (one-tailed)
TOLD	+	0.576	3.110	1.779	0.039

Nagelkerke $R^2 = 0.025$

Step 3: Regress dependent variable on the independent variable and mediator, using logistic regression

$$\text{MIS} = \beta_0 + \beta_1 \text{TOLD} + \beta_2 \text{DISPLACE} + \varepsilon$$

Variable	Predicted Relation	Beta	Wald Coefficient	Exp	Significance (one-tailed)
TOLD	+	-0.172	0.211	0.842	0.323
DISPLACE	+	3.944	14.449	51.601	0.000

Nagelkerke $R^2 = 0.321$

Significance indicators:

Bold = significant at 0.05 level (one-tailed)

Variable definitions:

TOLD = independent variable. 1 if participants are told to misreport, 0 otherwise.

DISPLACE = mediator. 1 if participant displaces responsibility, 0 otherwise.

MIS = dependent variable. 1 if misreported, 0 otherwise.

Table 5: Tests of Hypothesis 3:

H3a: Misreporting individuals who are instructed to misreport will feel more guilt and discomfort than honest reporters.

H3b: Misreporting individuals who are instructed to misreport will feel less guilt and discomfort than misreporters who were not instructed to do so.

Panel A: T-Tests of Differences in Negative Affect

Misreports versus honest reports in TOLD condition only

Variable	Predicted Relation	df	t statistic	Significance (one-tailed)
GUILT	+	86	5.475	0.000
DISCOMFORT	+	86	2.448	0.005

Panel B: Linear Regression of Guilt on Condition (misreporters only)

$GUILT = \beta_0 + \beta_1 TOLD + \beta_2 NOT\ RETOLD + \beta_3 DIFF + \beta_4 RATIONALIZE + \varepsilon$

Variable	Predicted Relation	Standardized Coefficient	t statistic	Significance (one-tailed)
Intercept	?		8.590	0.000
TOLD	-	-0.333	-2.640	0.005
NOT RETOLD	-	-0.144	-1.184	0.120
DIFF	+	-0.048	-0.527	0.300
RATIONALIZE	-	0.080	0.833	0.204

Adjusted R2 = 0.029 F = 1.894 (significant 0.116) Number of observations: 119

Panel C: Linear Regression of Discomfort on Condition (misreporters only)

$DISCOMFORT = \beta_0 + \beta_1 TOLD + \beta_2 NOT\ RETOLD + \beta_3 DIFF + \beta_4 RATIONALIZE + \varepsilon$

Variable	Predicted Relation	Standardized Coefficient	t statistic	Significance (one-tailed)
Intercept	?		7.927	0.000
TOLD	-	-0.306	-2.417	0.009
NOT RETOLD	-	-0.229	-1.870	0.032
DIFF	+	-0.011	-0.118	0.460
RATIONALIZE		0.020	0.210	0.427

Adjusted R2 = 0.018 F = 1.533 (significant 0.197) Number of observations: 119

Significance indicators:

Bold = significant at 0.05 level (one-tailed)

Table 5: Tests of Hypothesis 3 (continued)

Variable definitions:

GUILT = sum of six words or phrases, each on a 7-point Likert scale: annoyed at myself, disappointed with myself, guilty, regretful, self-critical, and shame. Range from 6 to 42; higher = more guilt.

DISCOMFORT = sum of six words, each on a 7-point Likert scale: bothered, distressed, negative, tense, uncomfortable, and uneasy. Range from 6 to 42; higher = more discomfort.

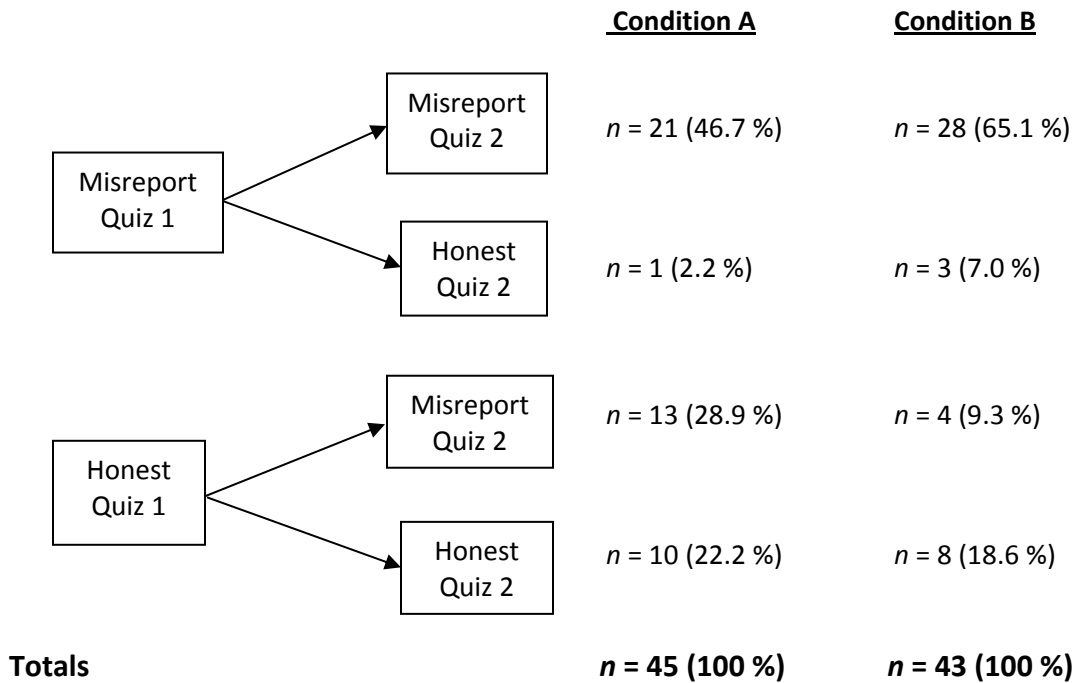
TOLD = 1 if participants are told to misreport (second quiz of condition A, first quiz of condition B), 0 otherwise

NOT RETOLD = 1 if participants are not told to misreport after being told to misreport once before (second quiz of condition B), 0 otherwise

DIFF = proportion of the total possible misreporting amount, given by the following equation: $[(\text{reported} - \text{earned}) / 20.00 - \text{earned}] * 100$. Range from 0 (honest) to 100 (misreporting by the greatest amount possible).

RATIONALIZE = 1 if participant used any rationalization category, 0 otherwise.

Figure 1: Reporting Behavior in Both Conditions



Notes:

Misreport Quiz 1 = participant misreported in the first quiz

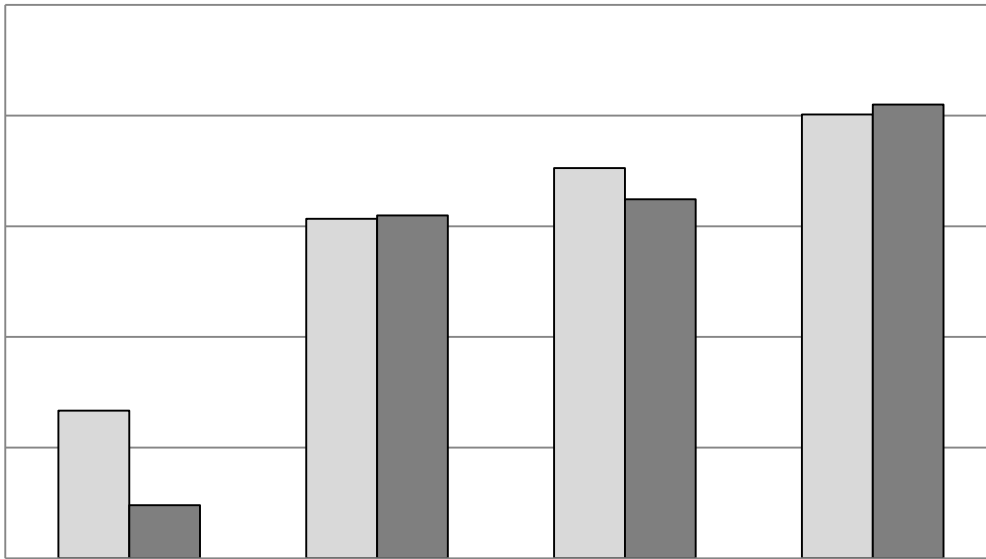
Misreport Quiz 2 = participant misreported in the second quiz, and so on.

Condition A: instructions to misreport are provided prior to the second quiz, not the first.

Condition B: instructions to misreport are provided prior to the first quiz and not the second.

The last two columns report the total number and percentage of participants on each path in each condition. For example, 21 (46.7 percent) of participants in condition A misreported in both quizzes while 28 (65.1 percent) of participants in condition B misreported in both quizzes.

Figure 2: Negative Affect Levels by Reporting Behavior and Treatment



Variable definitions:

GUILT = Level of guilt following a report decision; range from 6 to 42. Higher = more guilt.

DISCOMFORT = Level of discomfort following a report decision; range from 6 to 42. Higher = more discomfort.

Honest Reports = 1 if participant reported the same income as earned.

The following variables apply to only those participants who misreported:

Misreport: TOLD = 1 if participants are told to misreport (second quiz of condition A, first quiz of condition B), 0 otherwise

Misreport: NOT RETOLD = 1 if participants are not told to misreport after being told to misreport once before (second quiz of condition B), 0 otherwise

Misreport: NOT TOLD = 1 if participants are not told to misreport (first quiz of condition A), 0 otherwise.