The Dissertation Committee for Michael Duane Buhrmester Certifies that this is the approved version of the following dissertation:

UNDERSTANDING THE COGNITIVE AND AFFECTIVE UNDERPINNINGS OF WHISTLEBLOWING

Committee:

William B. Swann, Jr., Supervisor

James W. Pennebaker

Samuel D. Gosling

Caroline Bartel

Harvey Whitehouse

UNDERSTANDING THE COGNITIVE AND AFFECTIVE UNDERPINNINGS OF WHISTLEBLOWING

by

Michael Duane Buhrmester, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2013

Dedication

To my mom and dad – your strength, love, and support guide me each day.

Acknowledgements

Many individuals have made an impact on this work. Above all, I owe a tremendous debt of gratitude to my advisor, Bill Swann. Without any hesitation I can proudly claim that Bill has been the very best mentor I could have imagined: a sage thinker, an inspiring teacher, an enthusiastic motivator, an extraordinarily hard-working collaborator, a tenacious racquetball player, and above all, a friend. Bill, thank you for your steady guidance over so many years.

I am also deeply appreciative to my dissertation committee members: Jamie, Sam, Caroline, and Harvey. Thank you for your thoughtful guidance and collaboration in the past, present, and hopefully the future. I also extend my gratitude to my fellow Swann Lab-mates both old and new; Matt Brooks, Tracy Kwang, and Bo Fraser, contributed important insights, time, and effort at multiple stages of the dissertation process. And to all of the research assistants who dedicated endless hours of careful transcription, coding and fine acting as confederates – a huge thanks to each of you.

Last but not least, I am forever indebted to the love of my life, Allison Zaby, for the positive support showered on me every step of the way.

UNDERSTANDING THE COGNITIVE AND AFFECTIVE UNDERPINNINGS OF WHISTLEBLOWING

Michael Duane Buhrmester, Ph.D.

The University of Texas at Austin, 2013

Supervisor: William B. Swann, Jr.

Enron, Pfizer, UBS, Halliburton: In recent years, organizational wrongdoing has cost taxpayers and stakeholders billions of dollars. Whistleblowers, organizational insiders who witness and report wrongdoing with the intent of effecting an organizational response, play a major role as combatants to such corruption. What motivates whistleblowers versus silent witnesses of wrongdoing? And what cognitive and emotional patterns underlie their actions? Here I construe whistleblowing as a personally costly but pro-organizational action (Miceli, Near, & Dworkin, 2008). As such, whistleblowing represents a novel type of extreme pro-group behavior that identity fusion theory seeks to explain (Swann, Jetten, Gomez, Whitehouse, & Bastian, 2012).

The identity fusion approach posits that some people experience a visceral feeling of "oneness" with a group, a feeling that motivates a range of extreme pro-group actions. Across four preliminary studies, I first establish that fusion with one's organization (i.e., work or university) parallels fusion with other groups (e.g., country, political party). In addition, Preliminary Study 4 shows that fusion and whistleblowing are associated in retrospective accounts of workplace behavior. Given this initial support, a controlled lab

experiment was conducted to address two major questions. First, to what extent is identity fusion with one's university associated with initial and formal whistleblowing behaviors? Second, in what ways, if any, do strongly vs. weakly fused individuals' cognitive and emotional experiences differ in response to witnessing organizational wrongdoing?

As hypothesized, fusion with one's university predicted spontaneous reporting of an in-group transgressor. Strongly fused students' actions were associated with several cognitive and emotional factors, and cross-method evidence indicated that active negative emotions (e.g., anger) coupled with a heightened sense of personal responsibility drove strongly fused persons to spontaneously blow the whistle. Furthermore, strongly fused students were also especially likely to formally (as compared to spontaneously) report the transgressor. Evidence from participants' debriefing responses suggested that while weakly fused students diffused formal reporting responsibility to others, strongly fused students felt personally responsible to follow-through with a formal report. Overall, these results suggest that identity fusion is a promising perspective for understanding motives underlying personally costly pro-group behaviors.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The popular media has portrayed whistleblowing as filled with drama and intrigue. Famous whistleblowers such as Mark Felt (aka "Deep Throat") of Watergate, Daniel Ellsburg, who leaked the Pentagon Papers, and Frank Serpico of the NYPD represent gripping narratives worthy of Hollywood scripts. However, high-profile cases reflect only a fraction of the diverse actions that make up whistleblowing. As evidence, take the formal, accepted definition of whistleblowing used by researchers in management and business ethics fields: Whistleblowing is "the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action" (Near & Miceli, 1985, p.4). This definition is broad, encompassing high-profile reports of political and sexual scandal, widespread company corruption, and low-profile acts, such as role-specified auditor reporting of accounting inaccuracies or wasteful spending.

Communication of perceived wrongdoing of an illegal, immoral, or illegitimate nature is central to the whistleblowing phenomenon. Perceived harassment, discrimination, unsafe or noncompliant working conditions, mismanagement, waste, fraud, cheating, and theft are common types of organizational wrongdoing that whistleblowers report (Near, Van Scotter, Rehg, & Miceli, 2004). Workplace complaints and disagreements (i.e., reports not about unethical wrongdoing) represent "employee voice," but not whistleblowing (Withey & Cooper, 1989).

Another key element of whistleblowing is that the whistleblower must intend to cause someone to address the wrongdoing. Chatting about a workplace incident with a spouse or friend who has no authority or ability to respond directly to the wrongdoer does not constitute whistleblowing. Individuals capable of addressing the wrongdoing range from organization supervisors, auditors, human resource managers, and ombudsmen to law enforcement, governmental agencies, media, or other investigative outlets outside of the organization (Miceli et al., 2008). Note that unlike most popular media treatments that

focus on whistleblower reports to channels external to their organization, the vast majority of whistleblowers actually report via internal channels, at least initially (Miceli & Near, 1992).

Whistleblowing shares some features with other psychological phenomena such as obedience (Milgram, 1974), criminal "snitching" (Swanner, Beike, & Cole, 2010), and eyewitness testimony (Wells & Olson, 2003), but it most closely overlaps with two other constructs: punitive altruism (Fehr & Gachter, 2002) and the bystander intervention phenomenon (Darley & Latane, 1968). Arguing that nepotistic and selfish motives for cooperation fail to account for a wide array of acts of pro-group cooperation, Fehr and colleagues have posited that cooperation evolved in part because some group members are willing to incur personal costs to punish anti-social group members (Fehr & Gachter, 2002). Based on experimental economic game paradigms, researchers have shown that under certain conditions some individuals in groups will sacrifice short-term personal gains in favor of punishing others who treat fellow group members unfairly (Fehr & Fischbacher, 2005). Whistleblowers, therefore, can be considered punitive altruists, albeit their intent to punish wrongdoers is usually mediated through channels of authority as opposed to having a direct punitive effect on wrongdoers.

Whistleblowing is also conceptually similar to bystander intervention phenomena. Borrowing heavily from the decision-making framework developed to explain why bystanders often fail to intervene in emergency situations, Dozier and Miceli (1985) proposed a model of whistleblowing that draws upon the bystander intervention literature. According to their model, potential whistleblowers make a series of judgments regarding the wrongdoing itself, responsibility to act, and the implications of different possible responses. Also in line with research on bystander effects, much whistleblowing research has focused on how situational factors influence judgments and reporting actions. In support of the model, studies have shown that whistleblowing acts are associated with multiple situational variables such as wrongdoing severity (Bergman, Langhout, Cortina, & Fitzgerald, 2002), wrongdoer power or status (Lee, Heilmann, & Near, 2004), and wrongdoing type (Miethe, 1999).

Studies of whistleblowing and bystander intervention are also parallel in that both are associated with widely varying rates of reporting. In one study, only 4% of federal employees who had experienced sexual harassment in the workplace said they had blown the whistle (Lee et al., 2004). In another study of auditors, 90% said they had reported financial wrongdoing as soon as they discovered it (Miceli, Near, & Schwenk, 1991). The wide range in reporting is similar to the range found in the bystander intervention literature (Latane & Darley, 1969).

Despite similarities with other constructs, whistleblowing is unique in large part because it is embedded within an organizational context. Unlike bystander and informant scenarios that usually occur outside of a strong group context (e.g., witnessing a crime on the street with other strangers), wrongdoing within organizations presents the opportunity for group-level concerns to influence whistleblowing outcomes. For a potential whistleblower, many concerns may arise: How will the wrongdoing affect the organization itself, its members, and me? How would reporting the incident impact the wrongdoer, the organization, my relationships, and me both at the workplace and personally? Why should *I* rather than other members blow the whistle? These and other questions, while making the phenomenon especially rich to study, have received little sustained attention by researchers due presumably to the difficulty of conducting such work; recruiting representative and large numbers of whistleblowers and studying whistleblowing experiences, ideally in real-time, present multiple challenges.

If there is one theme that has emerged consistently in the whistleblowing literature, it is that whistleblowers are typically torn between a desire to help the organization they love and a fear that they will end up hurting it, themselves, or both. Consider Sherron Watkins, the whistleblower in the Enron scandal of 2001. In a few short months, Enron went from a seemingly thriving energy company and darling of Wall Street to bankrupt due to years of widespread accounting malpractice (Healy & Palepu, 2003). According to Watkins, then vice president of the company, this outcome was exactly the opposite of what she intended (Swartz & Watkins, 2003). Instead, her initial intentions and actions upon discovering wrongdoing at the company were similar to those

of other individuals in similar cases: she simply wanted to fix the problems internally. Only after months of evidence gathering, internal memos, and meetings with other executives intended to help right Enron's financial ship did she see the writing on the wall for the company.

Is Watkins' case representative of whistleblowers' intentions more generally? Past research suggests so. Other proposed motives for whistleblowing, such as reporting based on principled moral grounds or for anti-social reasons, have received little sustained support (Graham, 1986; Miceli & Near, 1997). This is not to say that other reasons for whistleblowing are illegitimate or unfounded, but simply that most recent researchers construe whistleblowing as stemming from a desire to help or improve the organization (Brief & Motowidlo, 1986; Near & Miceli, 1996). Such pro-organization intentions are often identified as the underlying mechanism of various effects in the literature. For example, whistleblowing is more likely when employees perceive strong organizational support for reporting (Sims & Keenan, 1998; King, 1997), when the work organization is a highly ethical environment (Zhang et al., 2009), or when whistleblowing will lead to fair outcomes for the parties involved (Trevino & Weaver, 2001). Each of these factors, presumably, strengthens beliefs that reporting wrongdoing would benefit the organization.

For whistleblowers, a desire to remedy actions hurtful to the organization often accompanies a willingness to endure personal hardship as a result of reporting (Miceli et al., 2008). Many whistleblowers can reasonably expect to be retaliated against in some fashion (Miceli et al., 2008). In one review of studies that measured retaliation against whistleblowers, rates of retaliation ranged from 17% to 38% (Miceli, Rehg, Near, & Ryan, 1999). Some forms of retaliation can be relatively minor, such as small changes in job duties or negative performance evaluations, but many forms are severe, such as demotion, termination, lawsuits, and even physical threats (Miceli et al., 2008). With such serious personal costs, whistleblowers seem to be uniquely oriented to place group interests above personal self interest. Recently, identity fusion theorists have conceptualized the nature of just such an orientation.

Chapter 2: Identity Fusion

Recently, researchers have suggested that identity fusion occurs when people experience a visceral sense of "oneness" between their personal and group selves (Swann, Gomez, Seyle, Morales, & Huici, 2009). This sense of oneness is associated with porous boundaries between the personal self and group identity, allowing for each to synergistically energize the other in a manner not unlike the feelings of self-partner overlap in close dyadic relationships (Aron, Aron, Tudor, & Nelson, 1991). The fusion construct also shares some theoretical heritage (e.g., importance of categorical ties) with dominant interpretations of group identification inspired by self-categorization theory (e.g., Turner, Oakes, Haslam, & McGarty, 1994). However, fusion departs from these traditions in four important ways. First, fusion theory asserts that strongly fused individuals maintain a highly agentic personal self within group contexts, resisting "depersonalization" pressures highlighted by self-categorization perspectives. Instead, for strongly fused individuals, group membership is personal, as evidenced by their strong feelings of personal responsibility and agency for the group (Gomez et al., 2011). This feature of fusion theory is particularly important here, as it is clear that whistleblowers maintain a strong sense of personal agency that compels them to follow their moral principles despite considerable personal costs.

Second, fusion theory posits that at least for strongly fused individuals, personal and social identities are able to activate one another to reinforce similar behavioral outcomes. For example, when the personal selves of strongly fused persons are made salient, strongly fused persons are especially willing to endorse fighting or dying to protect the group (Swann et al., 2009). Other work has shown that strongly fused persons transform group outcomes into personal expectancies. For instance, after one's political party suffered a loss at the polls, only voters strongly fused with their party internalized the defeat, forecasting gloomier outcomes for their personal lives (e.g., personal well-being, financial status; Buhrmester et al., 2012). Similar to suffering political defeats, might witnessing acts of organizational wrongdoing activate the personal selves of strongly fused persons and thus intensify emotional responses?

Third, fusion theory posits that some group members care as much or more about their relationships with other group members as they do about more abstract, categorical ties to the group (i.e., shared qualities, outcomes, and goals). Whereas past social identity theorists have characterized social identification in terms of categorical ties and assume that group members view one another as categorically interchangeable when group identity is salient, fusion theory assumes that group members form close relational ties to one another in the form of actual or assumed relationships with individual members of the group (Markus & Kitayama, 1991; Brewer & Gardner, 1996). Strongly fused individuals' willingness to sacrifice their lives for unknown others in their group may in part trace back to their beliefs that fellow group members are not interchangeable cogs in the group machine. Similarly, in whistleblowing scenarios, strongly fused persons may be especially concerned with reducing the wrongdoing's negative effects on fellow group members.

The final principle of fusion states that strongly fused persons' allegiance to the group should tend to remain highly stable even when their allegiance is challenged by others. For example, take Cheryl Eckard, another corporate whistleblower who has publically said that she loved her job. In 2002, as a quality assurance manager at GlaxoSmithKline, she witnessed major manufacturing problems at the company's facilities in Puerto Rico. Rather than report directly to the FDA, a reasonable reporting channel given the level of wrongdoing, she reported the problems internally. After almost a full year of being ignored, ostracized, and finally fired from the company altogether, she finally made her claims public, leading to one of the largest governmental fines of its kind at the time. Like Eckard, a strongly fused person's unwavering bond with the organization may not only increase the chances of initial, spontaneous reporting; it may motivate one's persistence to see the claim through to its effective end, no matter the personal costs. This anecdotal evidence closely parallels demonstrations that strongly fused group members redouble their willingness to sacrifice for the group even when they have been ostracized by several members of the group (Gomez et al., 2011).

Chapter 3: Links between Fusion and Pro-Organizational Activity

The fusion construct was developed primarily to explain why some group members are willing to make significant, sometimes extreme personal sacrifices for the good of their groups. Such actions usually are not official group doctrine, yet some members are motivated to go beyond the call of duty. Although relatively rare compared to other types of group activity, pro-group acts involving personal risk or sacrifice can significantly impact individuals, groups, and broader society (e.g., terrorism, responses to man-made and natural disasters).

Over the past several years, a growing number of empirical investigations have borne out fusion's predictive role for various acts of personally costly pro-group action. Persons strongly (vs. weakly) fused with their country report a greater willingness to physically fight and sacrifice their lives to defend their country from threats (Swann et al., 2009; Gomez et al, 2011). In intergroup variations of moral dilemmas, strongly fused persons endorse committing suicide to save fellow country members' lives over doing nothing and watching country members die; weakly fused persons were much less willing to endorse such self-sacrifice (Swann, Gomez, Huici, Morales, & Hixon, 2010; Gomez et al., 2011). Moreover, strongly fused persons are also uniquely willing to endure severe in-group ostracism so they can retain their abilities to promote their group (Gomez et al., 2011). Rather than give up on the group when the group has given up on them, strongly fused members increase their commitment to their group (see also Swann et al., under review). Other studies have shown that strongly fused persons are especially willing to "put their money where their mouth is." For example, when presented with an opportunity to donate to fellow Spaniards in need of financial help, strongly fused persons donated more personal funds than weakly fused persons (Swann et al., 2010). In another cross-cultural investigation, Canadian, Chinese, and Indian participants played a resource-allocation computer game designed to measure helping behaviors between players (Semnani-Azad, Sycara, & Lewis, 2012). Players who were strongly fused with

their respective countries allocated more resources to fellow members in the game and made fewer selfish requests for aid than did weakly fused players.

Behaviors like actual or virtual resource donation only begin to touch on the diversity of real-world actions (e.g., whistleblowing) that may be explained under the fusion framework. Here, I seek to understand fusion's ties to two important whistleblowing behaviors in a controlled lab environment. First, when strongly versus weakly fused group members (i.e., work or university) witness organizational wrongdoing, what are their spontaneous cognitive, emotional, and behavioral responses? Second, for members who spontaneously blow the whistle, do strongly vs. weakly fused members make different mental calculations about whether to formally report the wrongdoing when doing so comes at a personal cost?

Broadly, I posit that fusion with one's group (i.e., work or university) serves as an important lens through which members interpret and respond to organizational wrongdoing. Due to the porous boundary between their personal and social identities that intensifies emotional reactions to group events, strongly fused persons should be in a more affectively-charged state after witnessing organizational wrongdoing than weakly fused persons. Specifically, I expect strongly fused students to exhibit high levels of surprise, active negative emotions (e.g., anger, upset), and relatively lower levels of positive emotions (e.g., content, upbeat) compared to weakly fused persons. In addition, based on past evidence that strongly fused persons' thoughts are oriented outward toward other group members rather than selfishly inward, I predict that strongly fused persons' thoughts will reflect deep concerns about the wrongdoing's effects on the group and other group members while concerns about oneself will be scant. Weakly fused persons should show the opposite pattern of concerns (i.e., more selfish than unselfish).

Together, strongly fused persons' concrete, immediate thoughts and feelings in response to the event should influence abstract judgments about the overall egregiousness of the wrongdoing as well as feelings of personal responsibility to report the wrongdoing versus a diffusion of responsibility. Bolstered by dual convictions that a significant moral violation has occurred and that one has a personal responsibility to report the violation, I

predict that strongly fused persons will spontaneously and formally blow the whistle despite personal costs associated with doing so. Conversely, I predict that weakly fused persons will diminish the egregiousness of the wrongdoing, diffuse reporting responsibility to others, and not spontaneously or formally report the wrongdoing to the same extent as strongly fused persons. Four preliminary studies lay the groundwork for a controlled lab study designed to provide insights into the divergent reactions of strongly and weakly fused group members.

Chapter 4: Preliminary Studies

As highlighted earlier, past research on identity fusion has emphasized the relationship of fusion to a person's willingness to sacrifice one's life for one's country. Although identity fusion theory is meant to apply to a diverse range of groups, no published studies to date have examined the specific application of fusion to work organizations or universities. Establishing construct validity for these focal groups is necessary since they are somewhat different in nature from the group "nation". For one, work organizations and universities are much smaller than nations in terms of number of members and average amount of time one has been a group member in most cases. In addition, members of work organizations and universities are also involved in their groups to promote personal economic and educational outcomes, whereas involvement in one's nation is built upon a broader social contract. In four preliminary investigations, I explore basic psychometric properties of measures of fusion with one's work organization and university, compare other/perceiver versus self-perceptions of organizational fusion (Preliminary Studies 1 & 2), and present preliminary predictive validity evidence linking fusion with one's work organization to pro-organizational intentions and whistleblowing (Preliminary Studies 3 & 4).

PRELIMINARY STUDY 1

In Study 1, ninety-four UT students (M age = 19.8; 67 females) completed the pictorial and verbal fusion scales in reference to the university (i.e., "I am one with my university") as part of a larger battery of scales in a lab session. After completing the scales, they were instructed to reflect upon their ratings and write a short statement about their thoughts and feelings toward UT (word count range: 7-85 words).

Consistent with previous validation studies (Gomez et al., 2011), the verbal measure displayed high reliability, Cronbach's α = .88. The overall mean of verbal measure scores was near the neutral mid-point of the scale, M = 4.35, SD = 1.19. Participants' scores on the verbal measure were normally distributed upon visual

inspection of a histogram with normal curve overlay. A Shapiro-Wilk test was also not significant, suggesting normality. Also consistent with past studies, the verbal measure correlated moderately strongly with the pictorial measure, $r(94) = .61$, $p<.001$. The pictorial measure mean was near the scale mid-point $M = 3.17$, $SD = 1.06$.

Overall, both the pictorial and verbal measures of fusion with one's university exhibited acceptable psychometric qualities on par with previously validated measures of fusion with one's country. Given past evidence showing that the verbal measure is more strongly related to pro-group outcomes (Gomez et al., 2011), I chose to utilize only the verbal measure for future studies.

Did perceivers agree about targets' levels of fusion based on targets' written thoughts and feelings toward UT?

High agreement signals similarity in perceivers' interpretations of 1) the fusion construct, 2) the content of written thoughts and feelings about UT, and 3) how the two are conceptually linked. Five volunteer perceivers familiar with the fusion construct were presented with targets' responses to the open-ended 'explain' question. Perceivers were not given any other information about the participants other than that they were UT students taking part in a psychology survey. Perceivers rated how fused they believed each participant to be according to the pictorial scale measure (1 = not at all fused; 5 = totally fused). To create an index of inter-rater consensus, the mean of all 10 possible pairwise correlations among the five perceivers was computed (Gosling, Ko, Mannarelli, & Morris, 2002). The mean *r* across the 10 correlations was .76. As an alternative measure of consensus, Cronbach's alpha was computed and was high, $\alpha = .94$. Together, these results suggest that people agree about how the content of responses reflects fusion levels with one's university. Based on information akin to asking a new acquaintance how he or she feels about UT generally, perceivers seem quickly able to grasp the acquaintance's level of fusion with UT.

Were perceiver impressions of targets' levels of fusion with UT accurate?

Perceivers' judgments of targets' fusion levels were averaged to represent a composite index of perceived fusion. The correlation between this composite index and target self-ratings on the pictorial scale was high, r (94) = .77. Apparently, perceivers familiar with the fusion construct were able to accurately discriminate between strongly and weakly fused participants based on short, written accounts of the targets' thoughts and feelings toward UT.

Did targets strongly versus weakly fused with UT differ in their thoughts and feelings about UT?

To learn more about the how students who are fused with UT experience their group, I examined whether the content of targets' writings was associated with their levels of fusion with UT. One set of perceivers familiar with the fusion construct read a portion of targets' written responses and agreed on 10 coding dimensions that they felt represented 5 broader constructs. A second set of perceivers then judged each response. One construct, feelings of shared strength with UT, was present in very few responses (11%) and was dropped due to insufficient power. The "individuation vs. oneness" construct captured the degree to which the student felt like an individual separate from the group (i.e. individuated, negative score) versus a sense of wholeness, oneness, or completeness with the group (i.e., oneness, positive score). The "emotionality and efficacy" construct captured the degree to which the student felt negative emotions or incompetence at UT (negative score) versus positive emotions or competence at UT (positive score). The "relational ties" construct captured the degree to which the student felt weak social bonds with other students (negative score) versus strong social bonds (positive score). Finally, the "categorical ties" construct captured the degree to which the student felt a weak sense of spirit or involvement with UT as a whole (negative score) versus a weak sense of strong spirit or involvement (positive score). Sub-dimensions, example responses, and an explanation of the coding can be found in Appendix A.

Were perceiver impressions of targets' levels of fusion with UT accurate?

Perceivers' judgments of targets' fusion levels were averaged to represent a composite index of perceived fusion. The correlation between this composite index and target self-ratings on the pictorial scale was high, r (94) = .77. Apparently, perceivers familiar with the fusion construct were able to accurately discriminate between strongly versus weakly fused participants based on short, written accounts of the targets' thoughts and feelings toward UT.

Table 1. Relationships between students' fusion levels and written response themes

	Fusion with UT (verbal scale)
Perceived Individuation vs. Oneness	r (49) = .57; p <.001
Perceived Emotionality & Efficacy	r (26) = .48; p =.01
Perceived Relational Ties	r (27) = .67; p <.001
Perceived Categorical Ties	r (51) = .49; p <.001

As seen in Table 1, fusion scores were associated with all four constructs. More strongly fused targets were more likely to report feeling a greater sense of oneness with UT, more positive emotions and feelings of efficacy, stronger relational ties to other students, and stronger categorical ties to the university as a whole. Together, these findings suggest that fusion with one's university is a construct that is detectable in participants' general thoughts and feelings about the university and is apparent to naive perceivers. Encouraged by these findings, in Preliminary Study 2 I aimed to explore similar issues in the context of fusion with one's work organization, the group context most associated with whistleblowing.

PRELIMINARY STUDY 2

In Study 2, 52 currently employed adults (*M age* = 32.0; 19 females) completed a short survey on Mechanical Turk for a small fee. Participants were presented with each item of the verbal fusion scale in reference to one's work organization (e.g., "I feel immersed in my work organization"). For each, they first rated how well they understood the meaning of the item on a Likert scale (1 = not at all; 2 = somewhat; 3 = mostly; 4 =

totally) before indicating their level of agreement with the item (1 = totally disagree…7 = totally agree). Next, participants were instructed to write down any "in-the-moment" thoughts and feelings about their organization elicited by the item. This protocol was meant to directly assess participants' interpretations of each scale item. Written responses to all seven items were merged to represent participants' written 'think-alouds' (for a conceptually similar procedure, see Swann, Stein-Seroussi, & Giesler, 1992). Finally, participants completed demographic information, including a short description of their job. Two participants indicated they were self-employed, indicating no relationship with a broader organization. These two participants were dropped from analyses, leaving N=50.

As in past studies, responses to the verbal measure displayed acceptable reliability, α = .82, had an overall mean near the neutral scale mid-point, M = 4.16, SD = 1.21, and were normally distributed. For the set of ratings of item understanding, the overall mean was well above the theoretical mid-point of the scale (M = 3.55; SD = .43; scale range = 1-4; one-sample $t(49) = 17.51$, $p < .001$). Ratings of item understanding were also acceptably reliable, α = .80. Together, these results suggest that participants had little trouble making sense of the fusion scale statements in the context of their work organization, and the scale displayed similar properties to both Preliminary Study 1 and previous studies of fusion with one's country (Gomez et al., 2011).

Did perceivers display consensus and accuracy?

Three volunteer perceivers familiar with the fusion construct were presented with the set of think-aloud responses. Perceivers were not given any other information about the participants other than their occupations and that they had taken part in an online survey. Perceivers rated how fused they believed each participant to be with their work organization (1 = not at all fused… 7 = totally fused). Consensus was again high, α = .94. During coding, one participant failed to follow the writing directions and could not be confidently coded, leaving N = 49. The correlation between the mean of perceivers' ratings and participant self-ratings on the verbal scale was positive, $r(49) = .61$, $p < .001$.

Consistent with Preliminary Study 1, perceivers were again able to agree and accurately tell apart high versus low fused participants, this time based on written accounts of their immediate thoughts and feelings toward their work organization.

Do employees who are strongly fused with their organization differ in their 'think-alouds' compared to weakly fused employees?

If participants deeply understand the meaning of scale items when completing them, then their written thoughts elicited by each item ought to be strongly related to concepts covered by the scale. To examine this, two perceivers familiar with the fusion construct rated each set of responses according to the same five constructs as in Preliminary Study 1 (i.e., individuation vs. oneness, emotionality and efficacy, relational ties, categorical ties, and shared strength). A third perceiver reviewed inconsistencies between the two sets of codes and made final judgments to be used for analyses. Shared strength was included in the analysis because a significant percentage of participants (90%) mentioned thoughts relevant to the concept.

Table 2 presents correlations between participants' fusion with their organization scores and each coded theme. As in Preliminary Study 1, all coded concepts and fusion were strongly associated. These data suggest that, as one would hope for any self-report scale that encourages conscious deliberation, strongly fused participants' self-generated thoughts elicited by the items reflect stronger themes of oneness with the organization, positive emotions, social bonds with fellow employees, belief in organizational goals and values, and reciprocal feelings of strength and helping.

Table 2: Relationships between UT students' fusion levels and written response themes

	Fusion with work org.
Individuation vs. Oneness	r (43) = .63; p <.01
Emotionality & Efficacy	r (40) = .60; p <.01
Relational ties	r (34) = .47; p <.01
Categorical ties	r (41) = .62; p <.01
Shared Strength	r (44) = .63; p <.01

Together, Preliminary Studies 1 and 2 establish that individuals broadly construe fusion with one's work and educational organizations in expected ways. I turn now to focus on issues of convergent, discriminant, and predictive validity. Preliminary Study 3 explores fusion's relationships with a range of demographic, organizational, and pro-group outcomes. Preliminary Study 4 more specifically targets fusion's association with whistleblowing phenomena.

PRELIMINARY STUDY 3

In Preliminary Study 3, 207 employed Americans (*M age* = 30.9; 110 females; 78% Caucasian) completed an online survey on Mechanical Turk in exchange for a small fee. Instead of an open-ended question to assess occupation, participants indicated their work field from a list of 23 categories generated by the Bureau of Labor Statistics (e.g., military, sales, personal care & services, etc.). Participants in the sample represented all 23 categories. "Education, Training, and Library" fields and "Sales" fields were the two most common types indicated (13.3% and 12.8% of total participants, respectively).

Participants next completed a survey battery. Multiple measures included in past studies of fusion with one's country were modified to reference the work organization context. Participants completed the fusion with one's work organization scale (α = .86), a 6-item measure of organizational identification (Mael & Ashforth, 1992; α = .86), a 5-item measure of group-directed agency (e.g., "I am responsible for my work organization's actions"; α = .91), and a 6-item measure of willingness to protect one's organization (e.g., "I'd do anything to protect my organization"; α = .83). Other measures previously shown to be associated with pro-organizational outcomes were also included: a 5-item measure of perceptions that one's organization is an ethical work environment (e.g., "Unethical behavior is punished in my organization"; α = .67; Trevino et al., 1998), a 2-item measure of job satisfaction (e.g., "In general, I am satisfied with my job"; α = .76), and a 5-item measure of willingness to take change-oriented organizational action (e.g., "I speak up and encourage others to get involved in issues that affect the organization"; α = .88; adapted from Van Dyne & LePine, 1998). Finally, participants

indicated how many years they have been employed at their organization, their organization size (small—large), and whether or not they managed other employees (0 = no; 1 = yes).

Table 3: Correlates of Organizational Fusion (Pearson *r*'s)

	Organizational Fusion
Demographics	
Gender	.05
Age	.00
Education level	.06
Organization variables	
Years employed at organization	.04
Size of organization	-.09
Manage others?	.24*
Perceptions of ethical work environment	.30*
Job satisfaction	.54*
Sense of group-directed agency	.53*
Organizational identification	.63*
Willingness to engage in Pro-org. actions	
Protect the organization	.40*
Effect extreme organizational change	.59*

Note: N=207; * = $p<.01$

Table 3 shows the correlates of fusion with each measure. Strongly fused individuals were especially likely to report a willingness to engage in a broad range of pro-organizational actions ranging from mild to extreme. They also tended to be high in organizational identification, feelings of group-directed agency, perceptions of an ethical work environment, and job satisfaction; however these relationships were not so strong as to suggest total conceptual overlap. Fusion was weakly related to management position and unrelated to organization size, tenure, and other demographic variables, suggesting that fusion does not emanate from specific demographic or structural circumstances. Overall, these results establish the foundations of convergent, discriminant, and

predictive validity for the fusion measure regarding one's work organization in ways that largely cohere with past validation studies (Gomez et al., 2011).

PRELIMINARY STUDY 4

Preliminary Study 4 aimed to directly link fusion to two existing methods of tapping whistleblowing phenomena: reporting intentions in hypothetical scenarios and reporting past actual whistleblowing action. Two hundred and ninety-five employed Americans (*M age* = 31.5; 130 females; 73% Caucasian) completed an online survey on Mechanical Turk in exchange for a small fee. Participants first completed the fusion scale (α = .88) and demographic and organization items as in Preliminary Study 3. Participants then responded to three hypothetical 'whistleblowing scenarios' developed by Sims and Keenan (1998). Each scenario asked participants to imagine they had witnessed a different kind of organizational wrongdoing (i.e., co-worker is violating company policy, discriminatory organizational practice, co-worker tampers with evidence in legal proceeding). After each scenario, participants indicated which of five actions they would take, with the most extreme response representing a willingness to blow the whistle and persevere despite negative consequences. The relationship between fusion and whistleblowing versus less extreme responses was weakly positive but not significant, r = .09, p =.16. Treating the scenario responses as continuous rather than dichotomous, however, revealed a small but significant positive correlation with fusion, r = .21, p < .01. Given that there is no past research showing that people uniformly see *only* the most extreme responses to these scenarios as constituting whistleblowing, it seems more reasonable to interpret the continuous responses rather than the dichotomized metric (but see Appendix B for details).

To measure retrospective reporting of whistleblowing (Bjorkelo et al., 2010), participants were presented with a definition of whistleblowing and asked to report whether or not they have engaged in whistleblowing at their current organization (yes/no). 13.9% of participants indicated that they had engaged in whistleblowing. Fusion was positively related to whistleblowing reporting, $r(295)$ = .22, p < .001. In a binary

logistic regression model controlling for number of years employed, gender, age, management position, education, and identification with organization, fusion remained a significant predictor of whistleblowing reporting.

Overall, evidence for links between fusion and whistleblowing was mixed but promising. Associations between fusion and whistleblowing intentions were weak but in the expected positive direction. There could be multiple interpretations of these results, but past researchers have noted that intentions involving ethical judgments are hindered by significant social desirability biases (Smith & Ellingson, 2002). Such scenarios often tap what people feel they *should* do rather than *would* do. Meta-analytic results of the whistleblowing literature further suggest issues of response bias: few dispositional and situational variables predict both intentions and actual whistleblowing behaviors (Mesmer-Magnus & Viswesvaran, 2005). Given the strong concerns with assessing and interpreting hypothetical intentions, I avoided this approach in the main investigation.

For the item asking about retrospective whistleblowing experience, as expected, employees who reported being more fused with their organization were more likely to say that they had engaged in whistleblowing action, defined broadly, than weakly fused employees. Although this relationship was fairly modest, given the ambiguities inherent in this approach (e.g., whether the whistleblowing was construed as pro-organizational or anti-organizational, whether reporting was internal or external, etc.), it was encouraging to find a statistically significant correlation.

In summary, across the four Preliminary Studies I found 1) the verbal scales of fusion with one's work organization and fusion with one's university exhibited high internal reliability, 2) perceivers and targets agreed about targets' fusion levels, 3) strongly vs. weakly fused individuals differed significantly in their written thoughts and feelings about their organizations in ways consistent with identity fusion theory, and 4) the verbal fusion measure displayed convergent, discriminant, and predictive validity consistent with past studies.

Three results from the Preliminary Studies also provided initial support for hypotheses tested in the main investigation. First, strongly fused students' writings about

their relationship with their university reflected a high degree of emotionality and efficacy. Second, their writings also reflected strong senses of perceived relational connection to other students and categorical connection to the university more generally. Together, these results support the hypothesis that in a whistleblowing scenario, strongly fused persons' reactions to wrongdoing will be both affectively charged and especially concerned with the wrongdoing's effects on group members and the group at large. Third and most importantly, Preliminary Study 4 found suggestive evidence of a link between fusion and retrospective whistleblowing actions. Based on these encouraging results, I set out to further investigate the link between fusion and whistleblowing in an experimental laboratory environment.

Chapter 5: An Experimental Investigation of Whistleblowing as it Unfolds

To examine fusion's link to actual whistleblowing behavior, I developed a novel laboratory scenario designed to stage a situation in which participants witnessed supposedly real wrongdoing and had an opportunity to blow the whistle to effect a pro-organizational outcome. In the scenario, university student participants witnessed a fellow student (confederate) blatantly cheat on a test for personal gain and to the detriment of fellow students and the university at large. Later, students in the session had opportunities to both spontaneously and formally report the wrongdoing to the experimenter.

A brief methodological note. Past researchers, mostly in fields where controlled experiments are not the norm (e.g., business ethics), have indicated reluctance to develop a deception-scenario such as mine due to concerns about believability and generalization (Miceli et al., 2008). Here, based on extensive piloting, I lay out an approach that carefully choreographs a serious instance of student cheating coupled with later opportunities to make spontaneous and formal reports.

A major goal of the study was to record and quantify the cognitive and emotional underpinnings of the whistleblowing process. Past researchers interested in understanding people's in-the-moment thoughts and feelings have relied upon simple and direct "think-aloud" instructions (Swann, Stein-Seroussi, & Giesler, 1992). The current study's purposes would have been made transparent using similar instructions, so I developed a conceptually analogous, novel 'confidant' procedure. Shortly after witnessing the wrongdoing, some participants were randomly matched with student actors (i.e., confidants) who were blind to participants' fusion scores. In a neutral, supportive manner, confidants acted as an "open ear" to participants who desired to converse about the cheating incident. Insofar as participants' reporting actions reflect their shared thoughts and feelings with the confidant, this procedure provides a clear window into the proximal underpinnings of whistleblowing without raising suspicions about the study's true purposes.

METHODS

Participants

A total of 101 students from UT Austin's introductory psychology courses volunteered in the study. Students received course credit for their participation. Nine participants were dropped from analyses for 3 separate reasons. First, three participants were certain that the confederate-cheater was an actor. Second, four participants reported during the debriefing that they did not notice the confederate cheat during the test, which was an essential aspect of the procedure. Third, two participants overheard a fellow participant report the cheating incident to the experimenter before they had an opportunity to report to the experimenter by themselves. These participants were dropped because the goal of the paradigm was to test whether participants would report the cheating incident without knowledge of other participants' reporting actions.

This left a final study sample of $N = 92$ (52 females, 40 males; mean age = 19.34 years (SD = 1.77); mean years enrolled at UT = 1.47 (SD = .80); 42.4% Caucasian, 21.7% Hispanic/Latino, 10.9% African-American, 19.6% Asian-American, 5.4% identified as 'other'). Overall, the sample broadly represented the UT Austin student population, with minorities, females, and underclassmen slightly oversampled.

I should also note that two participants in two separate sessions unexpectedly cheated on the test but were *not* dropped from the analyses. In both cases no other participants noticed the participant cheat, thus the situational dynamics were the same as in other sessions (i.e., they only saw the confederate cheat as intended). Although the actions of the spontaneous-cheaters may have contributed to their decision to refrain from reporting the confederate-cheater, dropping them from the sample was unwarranted since the essential elements of the cheating paradigm – witnessing the cheater and having an opportunity to report the cheating without influence from other participants – were intact.

Materials and Procedure

Participants were invited to participate in sessions scheduled for up to four participants (minimum 2) and up to three confederates (minimum 1). Every session

required that a confederate male play the 'cheater' role. I decided that only males would play the cheater role because when a female played the cheater role in pilot sessions, participants tended to find her cheating hard to believe. Volunteer research assistants signed up at their convenience to play the role of 'confidant' (maximum two per session). The final sample included 31 confidant-participant pairs, a few short of the intended goal due to scheduling conflicts. Last, I played the role of the experimenter in the study because I needed to be available during each session in case of an unexpected event (e.g., extreme verbal or physical confrontation of the cheater).

At the start of each session, the experimenter told participants that the study sought to understand relationships between personality, social affiliations, and objective knowledge about UT. Upon completing the consent form, participants were randomly assigned to either an individual lab room or a lab room shared with a confidant. Participants and confidants sat at computers and began to complete a survey battery. Confidants engaged participants in small talk to develop some rapport so that later possible conversation after the cheating incident would not seem out of the ordinary. The cheater was always assigned to an individual lab room.

After completing demographic items, participants responded to the fusion scale (Gomez et al., 2010) with UT as the focal group. The scale, as in the preliminary studies, was internally consistent, α = .91, and normally distributed with a mean near the objective scale midpoint (M = 4.46; SD = 1.34). Participants then completed a measure of social identification with UT (Mael & Ashforth, 1992) and fusion scales in reference to the U.S., one's religious group, one's political party, and to 'all of humanity'. Before taking a break to complete a multiple choice test of their knowledge of UT traditions and history, participants also completed a five-item measure of UT-directed agency (Gomez et al., 2010) and the Big Five Inventory 44 (BFI; John, Naumann, & Soto, 2008). Unfortunately only 64 participants' data was gathered on the BFI because I unintentionally left out the items from the battery originally.

The experimenter then instructed participants and confidants to sit in the main lab room to take the UT-knowledge test. The main lab room contained one large table with

five chairs, a second small table with two chairs that both faced the wall so that persons sitting there could not see others without turning around in their chairs. Experiment materials were kept on top of a short file drawer against a wall in the middle of the room. For each session, the experimenter instructed the cheater and participants to sit in a chair at the larger of the two tables. The confidants were instructed to sit at the small table. This seating arrangement made it so that the cheater's actions would be in plain view of participants but not in view of confidants.

Once participants sat down, the experimenter explained that they would be taking a short multiple-choice test on their knowledge about UT traditions and history. As a reward, the highest scorer on the test in the session would be given an extra hour of participation credit. Participants each semester needed to accrue 5 hours total of volunteer credit as part of their coursework, thus the extra hour was an enticing reward. This setup made the cheater's actions believable and meaningful: cheating on the test would result in an unfair personal gain that another participant deserved. In addition, cheating on the test contaminated the cheater's results, hurting the integrity of the study and more generally violating the code of academic honor of the university. Reporting the cheater under these circumstances was therefore a pro-organizational act at multiple levels not unlike other workplace circumstances (e.g., fudging billable hours to gain a yearly bonus and firm recognition).

The Cheating Incident

Most students have witnessed academic cheating at some point in time (Burton & Near, 1995). This fact may explain why it took careful modifications during piloting to choreograph a cheating incident that students would find believable. As the experimenter explained the test, he passed out the 2-page stapled test question form each with a scantron tucked inside. A second scantron was tucked inside one of the forms. This scantron represented the answer key; the word "KEY" appeared in black marker across the top with the answers bubbled in. In each session, the experimenter handed the test form with the key inside to a confidant and regular test forms to the rest of the group.

Shortly after receiving the test form with the key, the confidant said aloud, "I think you accidentally gave me the key?" and handed the key back to the experimenter. The experimenter replied, "Oops! Yeah, that's not for your eyes!" and placed the key face down on top of a pile of papers on the file drawer in the room. This interaction sequence made it clear to the group that the answer key was present in the room.

The experimenter then left participants to complete the test. Test questions were very difficult to make it believable that someone would be visibly frustrated by the questions and resort to cheating (i.e., the confederate-cheater). After several minutes, the cheater displayed visible frustration at the test (e.g., sighs, pencil tapping, looking around). The cheater then continued working on the test until the other participants finished the test. As the participants finished, the cheater anxiously looked toward the door to ensure that the experimenter was still away. The cheater then rolled in his chair a few feet toward the file drawer, grabbed the key, then rolled back with the key in hand. After furiously erasing all of his answers, he quickly began to bubble in new answers from the key. A hidden camera in the corner of the room recorded these events. Hidden microphones in the individual and paired lab rooms later captured the participant-confidant conversations and the bogus debriefings.

Soon after the cheater finished cheating and returned the key, the experimenter re-entered the room. He passed out a "Test Follow-Up," instructing participants to rate the difficulty of the test and 8 items measuring positive and negative state emotions (strong, content, upbeat, attentive, angry, annoyed, irritated, and upset). Including the rating of the test difficulty was intended to increase the plausibility of the self-reports of emotions mere moments after the confederate completed his cheating routine. These items were selected based on their possible relevance to the cheating scenario and were adapted from the PANAS-X (Watson & Clark, 1994).

Participants, confidants, and the cheater went back to their individual or paired lab rooms to finish the survey battery. The rest of the battery contained a number of personality and other individual difference scales that were not part of the primary analyses. Scales included the Brief Loquaciousness and Interpersonal Responsiveness

Test (Swann & Rentfrow, 2001), the Texas Social Behavior Inventory (Helmreich, Spence, & Stapp, 1974), items tapping one's moral foundation of UT ingroup loyalty from the Moral Foundations Questionnaire (Graham, Haidt, & Nosek, 2009), and selected items tapping activity-level, assertiveness, altruism, integrity, cautiousness, teamwork, conformity, courage, deliberateness, emotion-based decision-making, fearfulness, harm-avoidance, leadership, perfectionism, persistence, and responsibility from the International Personality Item Pool (Goldberg, 1999).

Solo Participants

Once finished with the survey, participants waited for the experimenter to initiate the bogus debriefing. The bogus debriefing provided participants with an opportunity to spontaneously blow the whistle on the cheater. Past research on employee voice has shown that employees' decisions to voice concerns to management are influenced by their expectations of how their concerns will be received (Burris, 2012; Dutton, Ashford, O'Neill, Hayes, & Wierba, 1997). Since the goal of the study was to provide participants with a fair opportunity to spontaneously blow the whistle without the undue influence of strong contextual factors, I designed the bogus debriefing to be free of concerns about possible negative judgment by the experimenter. In the bogus debriefing, the experimenter debriefed each participant in his or her individual lab room. Each room was soundproof so that their conversation could not be heard by others. Each participant was thanked for his or her participation and told the broad (false) purposes of the study. To ensure that participants felt that they had engaged in an educational experience, the experimenter asked the participant to describe each part of the study step-by-step and to feel free to bring up any questions or comments they had during their time in the lab. At that point, some participants immediately blew the whistle. Some first discussed the survey portion of the study and then brought up the cheater when discussing the test portion. Others never brought up the cheating. From the actual debriefing later, it was clear that participants felt that the bogus debriefing was the safe and proper channel to report the cheating.

If the participant did not spontaneously report the cheater, the participant signed a bogus debriefing form and was free to leave. If the participant reported the cheater, the experimenter explained that the situation was serious and that he would come back with a misconduct form from the university's research ethics office. Completion of this form represented the formal reporting opportunity. Before leaving, the experimenter made it clear that completing the form would require extra effort and time from the participant.

After several minutes, the experimenter returned with the misconduct form. The experimenter reiterated that reporting the misconduct on the form constituted a serious allegation and then explained that the participant had a choice to make in the matter. If the participant desired to continue with a formal report of the cheating incident, the participant would complete the misconduct form. The form asked for contact information and a detailed description of the misconduct by the cheater. At some time in the future, the participant would also meet with an ethics committee member to discuss the misconduct. The experimenter then gave the participant the option to renege on his or her report the cheater by suggesting that he or she could refuse to complete the misconduct form and "forget the whole thing happened." As the experimenter left to let the participant decide whether to complete the misconduct form, he told the participant that after he or she completed the form or decided not to, he or she was free to leave the lab.

This sequence empirically captured each participant's willingness to follow through on a spontaneous report with a formal report. More so than the spontaneous report, formally reporting the cheater explicitly involved spending extra time and effort in the present moment to complete the misconduct form. Formally reporting also involved spending extra time and effort in the *future* (i.e., meeting with an ethics committee member in person).

Confidant-Paired Participants

Upon reentering their lab room to complete the survey battery, the confidant began working quietly. In a number of cases, as soon as the participant entered the room, he or she initiated conversation with the confidant about the cheater. In cases in which the

participant did not immediately initiate conversation about the confidant, the confidant was trained to wait up to several minutes then begin to casually bring up the test (e.g., "Did you do well on that test? I know I didn't get many right…"). The confidant did not directly bring up the cheating because 1) if the participant began to discuss the cheating, it would have been inconsistent with the confidant's claim that he or she did not see the cheating; and 2) I did not want the confidant to prime or pressure participants to discuss the cheating if they had already made a decision to keep it to himself or herself, as doing so could have systematically influenced reporting actions.

If participants initiated conversation about the cheater, confidants acted interested and explained that they did not witness the cheating because they sat facing the wall. Confidants were trained to elicit participants' experience witnessing the cheating (e.g., "Wait, how did he cheat?") and their immediate thoughts and feelings about the incident (e.g., "Seriously? Why would he do that?"). Confidants maintained a neutral, curious, and supportive tone with participants so as not to unduly influence their reactions or reporting actions.

After the conversation died down and the participant neared completion of the survey, the experimenter entered the room to check in on their progress. The confidant then asked to be excused to the restroom. With the confidant out of the room, the participant finished the survey shortly thereafter and the experimenter began the same bogus debriefing procedure as described above with the following modification: If the participant blew the whistle, the experimenter left the room as before to fetch the misconduct form. The confidant then reentered the lab room and chatted with the participant if the participant initiated conversation. The confidant then spent a couple minutes clicking through the survey to get to the end. When the experimenter returned with the misconduct form, he asked the confidant to wait in an adjacent room to be debriefed while he provided the participant instructions regarding the misconduct form.

Actual Debriefing. As each participant left the lab, the confederate-cheater caught him or her right outside the lab door, revealed that he was an actor in the study, and conducted the actual debriefing. Much to their relief, participants learned that the study

was an elaborate setup designed to measure students' responses to cheating. The confederate explained the actors involved in the study and reassured participants that none of their actions in the lab had any personal consequences for anyone involved in the session. Participants were then probed for suspicion open-endedly, and the debriefer followed up with additional questions for clarification to judge whether the participant was so strongly suspicious as to warrant exclusion from the dataset due to hypothesis-guessing, etc. Participants were then asked about their reporting experiences during the bogus debriefing and asked to explain their courses of action (e.g., 'why did you report/not report?', etc.). These debriefings were audio recorded in all but four cases due to human error (i.e., forgot to start recorder or partial recording due to lack of hard drive space). Finally, participants read through a comprehensive debriefing form that further detailed the design and purposes of the study, benefits to participants, and sources where participants could learn more about identity fusion and whistleblowing research.

CODING & DATA REDUCTION

Immediate Reactions to Cheating during Test

A research assistant and I reviewed videotapes of the test room to ensure that the proper cheating protocol was followed and to generate possible coding categories for interactions amongst participants, the cheater, and confidants. We coded each tape for whether or not the participant verbally engaged the cheater (e.g., "Are you really cheating?"; "You shouldn't turn that in") and whether or not each participant was present in the room when another participant verbally engaged the cheater (i.e., they overheard the verbal engagement). Other possible behaviors to code for either did not occur (i.e., participant-participant interaction in the presence of the cheater) or were too infrequent and/or difficult to code accurately (i.e., gaze direction and facial expression).

Initial and Formal Reporting Behavior

Initial reporting was coded by the experimenter as soon as he finished with the bogus debriefing for each participant (i.e., reported = '1'; did not report = '0'). In all

cases it was clear from the bogus debriefing conversation whether the participant had reported or not. To code for formal reporting (coded '1') versus not reporting (coded '0'), misconduct forms were examined. In all cases it was clear whether the participant had completed the misconduct form (i.e., all form fields were completed) or not (i.e., participant left entire form blank or in one case left only a note saying "Jesus is his only judge for his virtues").

Formal Reporting Thoroughness

For the last field of the misconduct form, participants were asked to write down what they witnessed in detail. A research assistant first transcribed the responses, then she and I read them and discussed candidate coding categories. Upon skimming the responses, we decided to code for one category: level of response detail. Each transcription was rated according to the following scale: 1 = not at all detailed, 2 = a little detailed, 3 = somewhat detailed, 4 = mostly detailed, 5 = extremely detailed. Two coders then rated each transcription for level of detail. Interrater reliability was acceptable, r (25) = .71. Word count for each transcription was also calculated (M = 54.4 words; SD = 37.7; range = 11 to 193).

Self-Reported Emotions after Test

To check that the positive and negative emotion items formed two separate constructs, a Principal Components Analysis was conducted (oblique rotation). Upon visual inspection of the scree plot, it was clear that two separate factors emerged, explaining 64.7% of the cumulative variance. Pattern Matrix scores indicated clear loading of the positively valenced items on one factor (content, upbeat, strong, and attentive) and the negative items on the other (upset, irritated, annoyed, and angry). Mean scores for 'positive emotions' and 'negative emotions' based each on the respective four items were then computed. Alphas for both positive and negative emotion factors were acceptable, α's = .76 and .86, respectively. The two emotion factors were uncorrelated, $r(92) = -.07$, n.s.

Coding the Confidant-Participant Interactions

Two research assistants and I read through a sample of the confidant-participant audiotapes to develop an initial coding scheme. Twelve categories were initially developed, and after discussion and testing the category-fit on several additional tapes, nine final dimensions were agreed upon. Research assistants also transcribed the interactions. Dimensions are presented in Table X and example statements for each code are in the Appendix.

Table 4. Coded Dimensions and Frequencies

Coded Dimensions for Confidant Interactions & Debriefing Reasons for Spontaneous & Formal Reporting	Confidant Interactions Frequencies	Spontaneous Reporting Frequencies	Formal Reporting Frequencies
Immediately brought up the cheater in conversation with confidant	32%	--	--
Described cheating in detail to confidant	45%	--	--
Expressed negative emotion about cheater	36%	14%	8%*
Expressed surprise about cheating	36%	7%*	3%*
Expressed selfish concern for oneself	29%	36%	33%
Expressed unselfish concern for other students or UT more generally	32%	18%	26%
Expressed concern regarding egregiousness of cheating	48%	30%	33%
Expressed responsibility to report cheater	36%	19%	39%
Clearly expressed thoughts and feelings	55%	63%	64%

Note: Frequencies represent the number of cases in which the action dimension was coded in the affirmative (i.e., present). * indicates that the frequency % represents fewer than 10 cases and had to be excluded from further analyses due to minimum cell *n* issues.

Two coders initially attempted to score each dimension on 5 point Likert scales; however, coders expressed difficulty distinguishing between scale values. To clarify the coding, we instead employed a binary coding system similar to the one used in the preliminary studies. Coders rated whether each interaction behavior was present (coded '1') or absent (coded '0'). Coders' binary judgments were the same in 88% of cases (Cohen's kappa range: .59 to .92; M = .70). I then reviewed the cases of disagreement,

looking first for possible data entry errors. Then I reviewed the interaction transcripts and judged in favor of one coder or the other.

Coding Experimenter-Participant Bogus Debrief Interactions

A research assistant and I listened to a small random sample of tapes. It became clear that for tapes in which the participant chose not to spontaneously report, there was no useful interaction to code for. We then listened to all of the tapes in which the participant spontaneously reported the cheater to the experimenter. We tried to apply categories developed from the confidant-participant interactions, but it became clear that there was quite a bit of uniformity in how the experimenter-participant interactions unfolded. As soon as the participant brought up the cheating, the experimenter would act surprised and ask the participant to explain what he or she saw. Participants complied and described that the key had been taken. In a few cases the participant would add that he or she or another participant said something to the cheater. The experimenter would then ask if the participant was sure of what he or she saw and almost all reiterated their description. The experimenter then continued to follow his script and then left the participant's room. Given the lack of variability in the experimenter-participant interactions, it was decided that the interactions could not be reliably coded for further analyses.

Coding the Actual Debriefings

Two volunteer assistants transcribed the actual debriefings for participants' responses to the debriefer's questions about participants' reporting actions and thoughts and feelings about the course of action they chose. After transcription, assistants and I read approximately 10 transcripts and discussed whether and how to amend the coding categories developed for the confidant-participant interactions. Two categories relevant only to the confidant-participant interactions, detailing the cheating to the confidant and raising the cheating first, were not raised by participants during the debriefing and did not make sense to code. Other categories made sense to code at the levels of 'reasons for spontaneous reporting decision' and if applicable, 'reasons for formal reporting decision.'

No new categories were agreed upon to be added. One research assistant then coded the debriefings. I also rated twenty random transcripts to check for coding reliability, and codes were congruent in 94% of cases (Cohen's kappa range: .69 to .89; M = .81). Higher agreement compared to the confidant coding was unsurprising because the debriefing transcripts tended to be shorter and the interaction was more circumscribed (i.e., question and answer format). The percentage of participants who were coded in the affirmative (i.e., a '1' to indicate presence of the action) out of the total N are presented in Table X. Examples of participant statements for each code are in the Appendix.

MAIN ANALYSES

Fusion with UT and Whistleblowing Actions

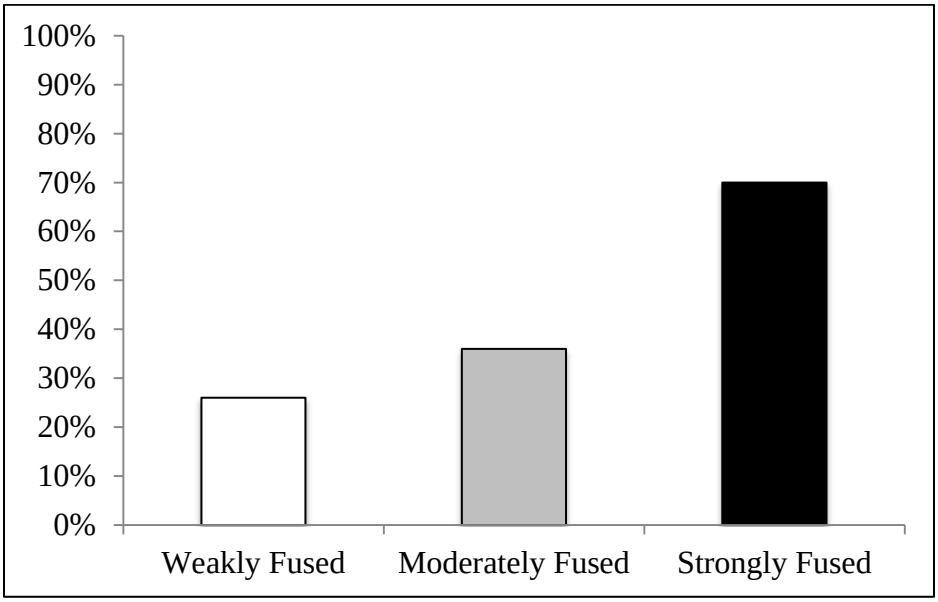
Overall, 41 out of 92 participants (45%) spontaneously blew the whistle by reporting the cheater's actions to the experimenter during the false debriefing. Twenty-five (61%) of participants who initially reported also formally reported. These distributions provided the backbone for further analyses treating these actions as dependent variables.

To examine whether fusion with UT predicted whistleblowing actions, binary logistic regression techniques were employed. Participants' spontaneous reporting actions were regressed on fusion with UT scores. As hypothesized, a significant effect of fusion emerged, B = .71, SE = .20, OR = 2.03, Wald χ^2 = 11.94, p = .001, indicating that strongly fused students were more likely to report the cheater to the experimenter than weakly fused students. As shown in Figure 1, 70% of strongly fused students (i.e., the top tertile on fusion) spontaneously reported the cheater, whereas only 36% of moderately fused students (i.e., middle tertile) and 26% of weakly fused students (i.e., bottom tertile) reported.

To test discriminant validity from social identification (Mael & Ashforth, 1992), spontaneous reporting was also regressed on fusion while controlling for scores on the social identification scale adapted to UT as the focal group. A main effect of fusion again emerged, B = .79, SE = .24, OR = 2.20, Wald χ^2 = 10.94, p = .001; however, the effect of

identification did not approach significance, B = -.22, SE = .31, OR = .80, Wald χ^2 = .50, p = .48.

Figure 1. % of students who spontaneously reported split by fusion tertile



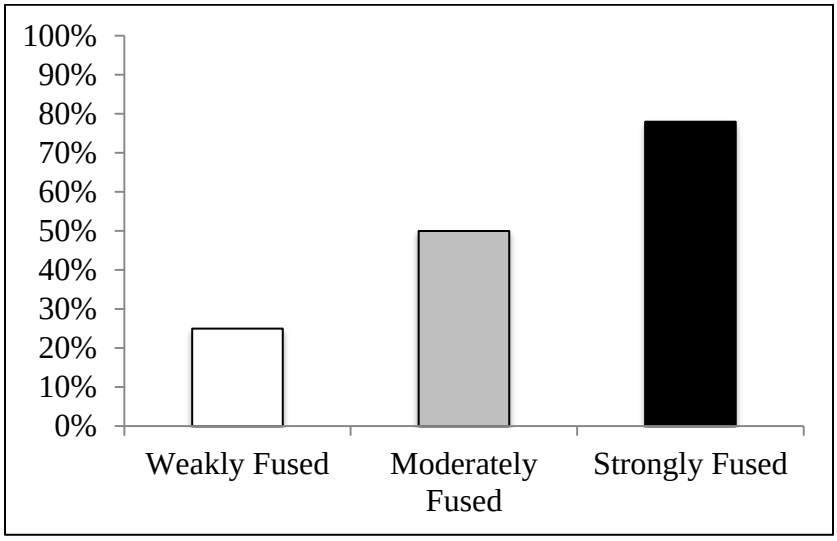
Note: Tertiles are illustrative only. Since fusion was measured with a continuous scale and the reporting outcome was binary, logistic regression was employed.

To test whether fusion with UT also predicted the formal reporting outcome, formal reporting was regressed on fusion with UT. Since the opportunity for formal reporting (versus abstention) was available only to participants who spontaneously reported the cheater, only the subset of participants who spontaneously reported (41/92 participants) were included for analysis. As hypothesized, a significant effect of fusion emerged, B = .76, SE = .33, OR = 2.13, Wald χ^2 = 5.41, p=.02, indicating that strongly fused students were more likely than weakly fused students to formally report the cheater. As shown in Figure 2, 78% of strongly fused students (i.e., the top tertile on fusion) formally reported the cheater, whereas only 50% of moderately fused students (i.e., middle tertile) and 25% of weakly fused students (i.e., bottom tertile) formally reported.

Spontaneous reporting was also regressed on fusion while controlling for scores on the social identification scale adapted to UT as the focal group. A main effect of

fusion again emerged in the expected direction, B = 1.38, SE = .47, OR = 3.99, Wald χ^2 = 8.66, p = .003. The effect of identification was significant but in the *opposite* direction of the fusion coefficient, B = -1.26, SE = .58, OR = .28, Wald χ^2 = 4.70, p = .07, suggesting that high identifiers were *less* likely to formally report than low identifiers.

Figure 2. % of students who formally reported split by fusion tertile



Mediating Role of Self-Reported Emotions

To test the hypothesis that state emotions in response to witnessing the cheater were related to both fusion and reporting outcomes, correlations amongst the different variables were computed. As seen in Table 5, negative emotions were related to fusion and both reporting outcomes, whereas much weaker, non-significant correlations were found for positive emotions.

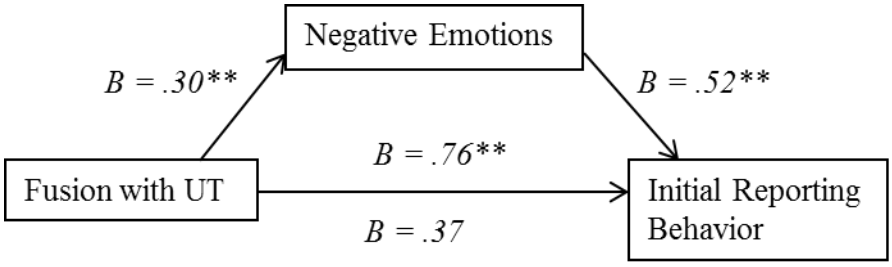
Table 5. Correlations between Self-Rated Emotions, Fusion, and Reporting

	Negative Emotions	Positive Emotions
Fusion with UT	.44**	.14
Spontaneous Reporting	.56**	.08
Formal Reporting	.48**	.19

Note: *: p < .05, **: p < .01; N = 92

Given the moderately strong associations with negative emotions, might negative emotions statistically mediate the relationship between fusion with UT and spontaneous reporting? To test this possibility, statistical mediation procedures based on bootstrapping methods developed by Hayes (2013) were utilized. As depicted in Figure 3, in a simple linear regression, fusion with UT predicted negative emotions $B = .30$, t (90) = 4.67, $p <$ .001. In a logistic regression, spontaneous reporting was regressed on negative emotions and fusion simultaneously. The indirect effect of negative emotions on spontaneous reporting was significant, $B = .52$, $p <.01$, and the direct effect of fusion was reduced to marginal significance, $B = .37$, $SE = .23$, OR = 1.45, Wald $\chi^2 = 2.58$, $p = .108$, bootstrapped 95% CI, .23 to 1.00, *N boots* = 5000. To test the directionality of effects, a second statistical mediation analysis was conducted in which negative emotions assumed independent variable status and fusion was considered the statistical mediator. The model fit poorly, resulting in a non-significant mediation effect (95% CI, -.09 to .69).

Figure 3. Self-Reported Negative Emotions Mediate Fusion Link to Spontaneous Reporting



Finally, to assess whether negative emotions mediated the fusion—formal reporting link, formal reporting was regressed on fusion and negative emotions simultaneously. The direct effect of fusion remained significant, $B = .74$, $SE = .34$, OR = 2.09, Wald $\chi^2 = 4.75$, $p = .03$, and no statistical mediation effect was observed, 95% CI, -.30 to .35. It should be noted that the low N in this analysis, 41 participants, resulted in an underpowered test able to detect only very large effects (Fritz & MacKinnon, 2007). This problem applied to all other multivariate models predicting the formal reporting outcome.

Mediating Role of UT-Directed Agency

The same analytic strategy as above was used to test whether UT-directed agency mediated the relationship between fusion with UT, spontaneous and formal reporting. First, fusion with UT predicted agency, $B = .53$, $t\ (89) = 5.68$, $p < .001$. Agency was also marginally related to spontaneous reporting, $r(92) = .19$, $p = .07$. However, when spontaneous reporting was regressed on fusion with UT and agency simultaneously, the direct effect of fusion remained significant ($p < .01$) and the indirect effect of agency was not significant (95% CI, -.21 to .22). For the formal reporting outcome, the correlation between agency and formal reporting was near zero, $r\ (41) = .02$, n.s., and no mediation was found, 95% CI, -1.44 to .15.

Whistleblowing Actions and Coded Interaction Dimensions

What thoughts and feelings were associated with students' reporting actions? A series of chi-square analyses were conducted to explore the relationships between participants' spontaneous reporting actions (i.e., report vs. report) and the 9 confidant-participant interaction dimensions (i.e., confidant interactions) as well as the 6 dimensions coded from participants' responses to the debriefing prompt 'why did you spontaneously report or not report?' Chi-square tests were also conducted to examine links between the formal reporting outcome (formally report vs. not) and the 5 dimensions coded from participants' responses to the debriefing prompt 'why did you formally report or not report?'

Did participants who spontaneously reported (vs. did not) share different thoughts and emotions with the confidant prior to reporting?

On 7 of the 9 dimensions the answer was yes. As seen in Table 6, a majority of participants who spontaneously reported were also the first to bring up the cheating in conversation with the confidant, detailed the cheating to the confidant, expressed negative emotion, surprise, unselfish concerns for others, concerns about cheating egregiousness, and feeling responsible to report. In contrast, only a minority of those who chose not to report expressed those same 7 thoughts and feelings with the confidant. Two

dimensions, voicing selfish concerns and clear thought and feeling expression, were unrelated to reporting vs. not reporting actions.

Table 6: %s of students who reported vs. did not report the cheater and also engaged in each coded action during interactions with confidants and during actual debriefing

	Confidant Interactions	Spontaneous Reporting Reasoning	Formal Reporting Reasoning
Immediately brought up the cheater in conversation with confidant	58% vs. 16% *	--	--
Described cheating in detail to confidant	67% vs. 37% ^	--	--
Expressed negative emotion about cheater	67% vs. 16%**	25% vs. 6% *	--
Expressed surprise about cheating	58% vs. 21% *	--	-
Expressed selfish concern for oneself	25% vs. 32%	20% vs. 50% **	21% vs. 53% *
Expressed unselfish concern for other students or UT more generally	67% vs. 11% **	33% vs. 6% **	42% vs. 0% **
Expressed concern regarding egregiousness of cheating	83% vs. 26% **	55% vs. 10% **	54% vs. 0% **
Expressed responsibility to report cheater	83% vs. 5% **	43% vs. 2% **	62% vs. 0% **
Clearly expressed thoughts and feelings	67% vs. 42%	88% vs. 21% **	63% vs. 67%

Note: * $p < .05$; ** $p < .01$; ^ $p < .10$ using chi-square analyses. For the Confidant Interactions and Spontaneous Reporting Columns, %'s are in reference to the spontaneous reporting outcome. For those columns, table should be read as "X% of students who spontaneously reported the cheater also engaged in Y action (from left-hand column), whereas Z% of students who did not report also engaged in Y action." For the Formal Reporting column, %'s are in reference to the *formal* reporting outcome (report vs. not). $N = 31$ for the Confidant Interactions column. $N = 88$ for the Spontaneous Reporting Reasoning column. $N = 39$ for the Formal Reporting Reasoning column. Dashes indicate that the coding dimension did not apply or too few cases were coded in the affirmative.

Did participants who spontaneously reported the cheating (vs. did not) express different thoughts and emotions during the debriefing when asked about why they did or did not spontaneously report?

On all 6 coded dimensions the answer was yes. As seen in Table 6, significantly more participants who spontaneously reported (vs. did not report) also expressed negative emotion, unselfish concerns for others, concerns about cheating egregiousness, feeling responsible to report, and expressed themselves clearly. Those who chose not to spontaneously report were more likely to voice a selfish concern than those who did spontaneously report. This last finding was the only result inconsistent with the confidant interaction results.

Did participants who formally reported the cheating vs. those who spontaneously reported but did not formally report express different thoughts and emotions during the debriefing when asked about why they did or did not formally report?

On 4 of 5 coded dimensions the answer was yes. As seen in Table 7, participants who formally reported also tended to express unselfish concerns for others, concerns about cheating egregiousness, and felt responsible to report the cheating. In contrast, those who chose not to formally report voiced selfish concerns. No differences were found on the clarity of expression dimension.

Fusion and Coded Interaction Dimensions

Given the link between fusion with UT and whistleblowing outcomes, would patterns of thoughts and emotions similar to those above be found with fusion? As above, analyses are organized into the three groups: confidant-participant interactions, spontaneous reporting reasoning, and formal reporting reasoning. Since fusion scores were on a continuous scale and coded dimensions were dichotomous, I conducted a series of point-biserial correlations (i.e., r_{pb}). All correlations are presented in Table 7.

Table 7. Correlations between fusion with UT and all coded dimensions

	Confidant Interactions	Spontaneous Reporting Reasoning	Formal Reporting Reasoning
Immediately brought up the cheater in conversation with confidant	.04	--	--
Described cheating in detail to confidant	.27	--	--
Expressed negative emotion about cheater	.46**	.14	--
Expressed surprise about cheating	.38*	--	-
Expressed selfish concern for oneself	.29	-.10	-.37*
Expressed unselfish concern for other students or UT more generally	.43*	.21*	-.01
Expressed concern regarding egregiousness of cheating	.48**	.32**	.18
Expressed responsibility to report cheater	.38*	.33**	.37*
Clearly expressed thoughts and feelings	.07	.14	.13

Note: * $p < .05$; ** $p < .01$; ^ $p < .11$; $N = 31$ for the Confidant Interactions column. $N = 88$ for the Spontaneous Reporting Reasoning column. $N = 39$ for the Formal Reporting Reasoning. column. Dashes indicate that the coding dimension did not apply or too few cases were coded in the affirmative.

Was fusion associated with thoughts and feelings expressed to the confidant?

Fusion was positively associated with 5 of the 9 dimensions. More so than weakly fused students, strongly fused students expressed negative emotion, surprise, unselfish concerns for others, concerns about cheating egregiousness, and feeling responsible to report. The remaining dimensions were positively associated with fusion but did not reach statistical significance.

Was fusion associated with thoughts and feelings in response to the debriefing prompt 'why did you spontaneously report or not?'

Fusion was positively associated with 3 of the 6 dimensions. More so than weakly fused students, strongly fused students voiced unselfish concerns for others, felt the cheating was egregious, and felt responsible to report.

Was fusion associated with thoughts and feelings in response to the debriefing prompt 'why did you formally report or not?'

Two of the five dimensions were associated with fusion. Strongly fused students were *less* likely than weakly fused students to express selfish concerns for oneself but *more* likely to voice feeling responsible to follow through with a formal report.

SUPPLEMENTAL ANALYSES

Large amounts of data were collected in this study. Some were collected to supplement the validity of the main analyses. Other data were collected to gain further insights into the fusion construct. In the following sections, I separate analyses of the supplemental data into discrete questions and answers.

Exploring Situational Influences on Reporting

The cheating paradigm contained some controlled situational variables and some that were uncontrolled but coded for afterwards. I examine each in turn.

Did confidant-pairing influence reporting frequency?

To explore whether being paired with a confidant versus being alone influenced reporting actions, I conducted two chi-square analyses. Thirty-nine percent of participants paired with a confidant spontaneously reported the cheater compared to 48% of unpaired participants. This difference did not reach significance, χ^2 (df = 1) = .65, p = .42. Fifty percent of participants paired with a confidant formally reported the cheater compared to 66% of unpaired participants. Apparently the added social dynamic of being paired with a confidant had little impact on participants' reporting behaviors. However, the fact that in both cases the reporting percentages were (non-significantly) lower when paired with a confidant hints that diffusion of responsibility could have been at work.

Did uncontrolled interactions influence reporting frequency?

In creating the cheating paradigm, I assumed that participants' experiences while witnessing the cheater would not be systematically affected by possible uncontrolled participant-cheater or participant-participant interactions during the test. A total of

sixteen participants either verbally engaged the cheater during the test about his cheating (N = 7) or were present in the room when another participant verbally engaged the cheater (N = 9). No participants mentioned the cheater to one another during the test. Given the small number of cases, statistical tests were not conducted due to unreliability issues. Nonetheless, I examined reporting actions and fusion scores for the groups of 7 and 9 cases.

Four of the nine who overheard another participant say something to the cheater spontaneously reported the cheater to the experimenter. Three of those four spontaneous reporters also formally reported. The mean fusion score for the nine cases was 4.30, very close to the overall mean.

All seven of those who said something to the cheater during the test also spontaneously reported to the experimenter, and 5 of 7 also formally reported. The mean fusion score for the 7 cases was 4.53, also very close to the overall mean. Together, these results suggest that overhearing another participant say something to the cheater was not uniformly tied to later reporting behavior; however, personally saying something to the cheater in all seven cases resulted in at least a spontaneous report. Given the presumed high level of social awkwardness involved in 'calling out' a cheater in close physical quarters, it is not surprising that such initial verbal boldness would be followed-up on in an spontaneous report to the experimenter later. Overall, these results suggest that witnessing the cheating was not systematically influenced by *other* participants and did not warrant dropping anyone further from the analyses.

Did different cheater-actors elicit different reporting frequencies?

A second assumption of the cheating paradigm was that participants' reporting behaviors would not vary systematically according to who played the role of the cheater (i.e., one of four male research assistants). To check this, I conducted a chi-square analysis that crossed participants' spontaneous reporting behaviors by the coded identity (1-4) of the assistant playing the cheater in each session. The analysis did not approach

significance, $\chi^2(df = 3) = 5.8$, $p > .1$), suggesting that reporting behaviors were not systematically influenced by who played the role of the cheater.

Addressing possible alternative explanations of findings

Since the design of the study was correlational, third variables unaccounted for in the main analyses or other statistical anomalies might be responsible for certain findings. I found no evidence for this hypothesis in the analyses reported below.

Do strongly fused students tend to express negative emotions no matter the situation?

One alternative explanation of the negative emotion statistical mediation is that strongly fused students tend to generally be higher on negative emotionality than weakly fused students, challenging the assumption that negative emotions were specifically in response to the cheating incident. Unfortunately, state emotions were not also collected prior to the cheating incident which would have provided the clearest means to test this alternative explanation. However, two trait dimensions related to negative emotionality, BFI Agreeableness and Neuroticism, were administered before the cheating took place. Contrary to the alternative explanation, fusion with UT was marginally but *positively* correlated with Agreeableness $r = .22$, $p = .08$. Fusion was not significantly related to Neuroticism, $r = -.17$, $p = .18$, but note that the sign of the r is negative, indicating that strongly fused students may tend to be slightly *less* neurotic, not more as the alternative explanation would argue. See Table 11 for all correlations between fusion, whistleblowing outcomes, and Big Five personality indices.

Can effects of fusion be explained by demographic variables?

In follow up regression analyses for the spontaneous reporting outcome, demographic covariates (i.e., age, gender, years attending UT, and ethnicity) were entered in Block 1 and fusion on Block 2. A significant effect of fusion remained, $B = .77$, $SE = .25$, OR = 2.15, Wald $\chi^2 = 9.78$, $p = .002$. Significant effects of gender, $B = -1.54$, $SE = .60$, OR = .21, Wald $\chi^2 = 6.55$, $p = .01$, and ethnicity (reference category = Caucasian), Wald $\chi^2 = 11.13$, $p = .03$, also emerged. Fifty-four percent of females overall

reported the cheater compared to only 33% of males. To understand why gender differences emerged, I conducted chi-square analyses crossing gender with all of the confidant and spontaneous reporting dimensions. *None* of the chi-square values approached significance. Spurious or real, the reason for the gender difference is unclear from the data. If real, perhaps the male cheater implicitly primed feelings of similarity or empathy amongst male participants, driving down reporting frequency. Alternatively, perhaps females felt more at ease reporting to the male experimenter than did males. See Table 12 for Pearson correlations between fusion, whistleblowing, and demographic variables.

By ethnicity, 62% percent of Caucasian students reported compared to 45% of Hispanic/Latino students, 30% of African-American students, 22% of Asian-American students, and 20% of students identified as 'other'. These differences are consistent with past work in the employment discrimination literature showing that whites are actually *more* likely to file discrimination claims than minorities, presumably due to differences in justice perceptions (e.g., Goldman, 2001). Anecdotally, during debriefing discussions with one Hispanic and one African-American male, both explained that they were socialized from a young age that 'snitches get stitches.' This colloquialism may reflect a deeper cultural dynamic that deserves more attention in future work on whistleblowing.

Can associations between fusion and emotion be explained by social identification?

Past research has shown that strongly fused persons' emotional reactions to group events are different not only from weakly fused persons, but also from highly identified group members (Buhrmester et al., 2012). To revisit this issue, I conducted a series of regressions using fusion and social identification with UT scores to predict self-reported negative emotions and coded negative emotions from the confidant interactions. For the self-reported negative emotion outcome, fusion and identification were entered as predictors on the same block. The coefficient for fusion was significant, $B = .30$, $p < .001$, but not for identification, $B = -.04$, n.s. In an analogous model with negative

emotion shared with the confidants as the outcome, the coefficient for fusion was marginal, B = .77, p = .106, and the identification coefficient was again smaller, B = .53, p = .32. In addition, in models that enter only fusion or identification as single predictors, fusion coefficients were greater than identification coefficients for both outcomes. Overall, these analyses suggest that associations between fusion and negative emotions cannot be fully explained by shared variance between fusion and identification.

Can the effects of fusion be explained by fusion with other groups?

To test whether the effect of fusion with UT on spontaneous reporting could be explained by shared variance with fusion with other groups (i.e., with the U.S.A., religious group, political party, and 'all of humanity'), the 4 non-UT fusion scores were loaded into Block 1 of a logistic regression, and fusion was added on Block 2. In this model, fusion with UT remained a significant predictor, B=.61, SE = .21, OR=1.84, Wald χ^2=8.18, p=.004, but no other predictors approached significance, B's < .28, p's <.19. Four regression models then separately tested whether the four other fusion scores predicted spontaneous reporting. Models for fusion with U.S.A., religious group, and political group did not approach significance, B's < .10. However, for the model that regressed reporting on fusion with 'all of humanity' scores, a marginal effect emerged B=.35, SE = .18, OR=1.42, Wald χ^2=3.69, p=.055, suggesting that students who felt more strongly fused with all of humanity were more likely to report than those who felt weakly fused with all of humanity. Although this relationship was unexpected, it may be the case that fusion with a diverse, egalitarian educational institution fosters fusion with all people more generally.

Analogous models were then conducted for the formal reporting outcome. Fusion scores with the four other focal groups were unrelated to formal reporting (.13 < p's < .82). In a model controlling for fusion with all four other groups, fusion with UT remained a significant predictor, B=.68, SE = .34, OR=1.98, Wald χ^2=4.03, p=.045. Overall, these results show that fusion with UT was the most robust predictor of reporting

outcomes when compared to other types of fusion, types that conceptually did not align with the cheating context to the same extent as fusion with UT.

Did measures of similar constructs across methods converge?

One trait-like self-report measure of UT-directed agency, one state self-report measure of emotions, and coded behavior dimensions (i.e., confidant-participant interactions and debriefing responses) were gathered during the study. To what extent did different measures of similar constructs converge? Two sets of measures were examined. For emotion measures, expressed negative emotions with a confidant were positively correlated with self-reported negative emotions after the test, r (31) = .70, p <.001. Expressed surprise emotion with a confidant was not correlated with self-reported positive emotions after test, r (31) = -.08, *n.s.* The null relationship between surprise and the positive emotions is consistent with the notion that surprise was not necessarily valenced in a positive or negative direction.

For responsibility measures, voiced responsibility expressed during the debriefing was positively correlated to self-reported UT-directed agency, r (88) = .33, p <.01. However, self-reported UT-directed agency and voiced responsibility with a confidant were not correlated, r (31) = .15, *n.s.* It is not surprising that relationships with UT-directed agency are inconsistent, as both coded responsibility dimensions were robustly related to reporting, but UT-directed agency was not.

Associations between fusion with UT and other variables

The preliminary studies provided crucial evidence that fusion with one's university shared similar properties to other types of fusion. Here, I conducted further tests using on the main study's data to explore other new questions not crucial to the main analyses.

Is Fusion Associated with Formal Reporting Thoroughness?

To test the possibility that strongly fused students would do an especially thorough job when formally reporting, correlations were computed between the coded

dimension 'level of reporting detail on misconduct form', word count of the transcribed responses, and fusion with UT. Coded level of reporting detail and word count were highly correlated, r (25) = .83, p <.001, but neither was correlated with fusion, r's = -.01, and .06, respectively. From reviewing their responses, differences in detail levels seemed to be due to idiosyncrasies of each student's memory of the cheating. For instance, some students noticed the cheater right away and were able to detail all elements of the cheating action. Other students noticed the cheater as he put the key back, thus their recollections were less detailed.

Did students signal fusion with UT through their physical appearance?

To explore the possibility that strongly fused students might physically embody the group by donning a UT emblem, name, or logo on their personal belongings, participants' physical appearances coded for the presence or absence of identifiable UT symbols. Twenty percent of students in the sample were identified as donning a UT symbol; however, this binary variable was not correlated with fusion, r(92) = -.05, n.s. To the extent that signaling one's fusion level is a motivated process, it may be that participating in a psychology experiment does not prime a signaling motive. Instead, a stronger test would be to measure fusion and later measure physical appearance at a group event like a football game. Understanding how and when fusion impacts self-expression in physical environments would add another layer to this important literature (Gosling, Ko, Mannarelli, & Morris, 2002).

Are fusions with different groups associated with one another?

Past work has shown that fusion with one group (e.g., country) is generally not associated with fusion with other groups (Swann et al., 2009). As shown in Table 8, fusion with UT was not significantly correlated with either fusion with the U.S.A or a political party. Fusion with UT was weakly but positively correlated with fusion with religious group and fusion with all of humanity. These findings reaffirm the conceptualization of fusion as an attitude-like construct that does not simply track a 'tendency to fuse with any group' notion.

Table 8. Correlations (*r's*) between Fusion with UT and Other Groups

	Fusion with UT
Fusion with U.S.A.	.15, *n.s.*
Fusion with Religious Group	.24, p =.02
Fusion with Political Party	.09, *n.s.*
Fusion with 'All of Humanity'	.23, p =.03

Language Analyses using LIWC

To provide further convergent evidence for the cognitive and affective underpinnings of whistleblowing actions, I explored relationships between fusion, whistleblowing, and several linguistic markers using LIWC. Based on past research and the contents of different dictionaries, I sought to test whether fusion and reporting behaviors would be positively related to overall word count, usage of 1st personal singular and plural pronouns, negative affect (especially the anger subdimension), positive affect, causal words, and use of cognitive mechanisms. First, transcriptions of the actual debriefing and confidant-participant interactions were formatted and processed using the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2007). Then, the word category data generated by LIWC was merged with the main data file.

For the confidant-participant interactions, LIWC variables of interest mostly were not significantly associated to fusion or reporting outcomes (*p's* > .20). Word count was positively but marginally associated with spontaneous reporting, r = .40, p = .09, significantly associated with formal reporting, r = .57, p = .01, but unrelated to fusion, r = -.09, n.s. Apparently with the confidants, strongly fused students were able to express their heightened emotion and concern in a similar number of words as it took weakly fused students to express less emotion and concern.

For the debriefings, consistent positive associations were found between word count and fusion, spontaneous reporting, and formal reporting, *r*'s > *.26*, *p*'s < .01. First person singular pronouns (e.g., I, me, my) were not associated with fusion, r = -.05, n.s., but were negatively associated with spontaneous reporting, r = -.33, p < .01, and formal

reporting r = -.29, p < .01. This last result is consistent with coders' ratings of selfish concerns. Finally, negative emotions and anger were positively associated with spontaneous reporting and formal reporting, r's > .18, but again not associated with fusion. Other LIWC variables of interest were unrelated to fusion and reporting outcomes. Overall, only some linguistic markers of affect and cognition were consistent with coders' ratings. I suspect two causes for the inconsistencies: First, the transcribed interactions ranged greatly in word count (3 to 506), leading to a number of unreliable estimates of word category frequency for some participants. To address this possibility, I dropped participants whose reporting reasoning consisted of less than 20 words. However, correlations amongst variables did not differ dramatically after dropping some participants. Second, the number of transcripts (less than 100) was low compared to other studies that have used LIWC to analyze thousands, even millions of large texts (Pennebaker et al., 2007). In future work, I would like to collect a much larger corpus of narrative texts produced by whistleblowers and group members who remained silent after witnessing wrongdoing and compare linguistic markers between these two groups.

Chapter 6: General Discussion

How do students respond when given an opportunity to blow the whistle? In the current study, nearly half of students (45%) spontaneously blew the whistle on the cheater. Consistent with my hypothesis, students who were strongly fused with UT were more likely than weakly fused students to spontaneously report the cheater to the experimenter. Convergent evidence from self-report and audiotaped data showed that negative emotions (e.g., anger, irritation, etc.) strongly underlie the link between fusion and spontaneous reporting of the cheater. Further behavioral evidence from participants' conversations with confidants and debriefings showed that strongly fused students voiced high levels of surprise by the cheating, expressed concerns about the cheater's negative effects on other students or the university more generally, expressed concerns about the egregiousness of the cheating, and voiced feeling personally responsible to report the cheater (i.e., a *lack* of diffused responsibility). Importantly, the cognitive and affective dimensions related to fusion were also related to students' actual spontaneous whistleblowing behaviors.

Fusion also predicted a second, more personally costly whistleblowing action: formally reporting the cheater to the university. Not surprisingly, rates of reporting the cheater to the university were a modest 61%. Consistent with expectations, strongly fused students were more likely than weakly fused students to formally report the cheater. Interestingly, in contemplating whether to complete the formal report, evidence showed that strongly fused students' emotions and concerns changed somewhat from when deciding to spontaneously report. Strongly fused students expressed *less* selfish concerns about involvement (i.e., they were undeterred by personal costs) and expressed greater feelings of personal responsibility to follow through with a formal report. Importantly, expressions of self-concern and responsibility were similarly associated with students' actual formal reporting behaviors. Overall, these findings paint a vivid picture of cognitive and affective antecedents of whistleblowing, a picture I dissect below.

SPONTANEOUS REPORTING

Without a point of reference, it is difficult to judge whether the 45% rate of spontaneous reporting is encouraging or discouraging from a justice perspective. However, this percentage is significantly higher than the 3%-11% of students in cross-sectional studies who said they have previously reported academic misconduct (Burton & Near, 1995; Sierles, Kushner, & Krause, 1988) and even higher than some estimates of academic cheating reporting *intentions* (Simon et al., 2004). The 45% is also on par with past research showing that 48% of federal employees who have witnessed organizational wrongdoing have blown the whistle (Miceli et al., 1999). These points of comparison suggest that the wrongdoing scenario created in the current study may have been considered more egregious than typical types of cheating in and outside of the college classroom.

Alternatively, the bogus debriefing scenario may have created a safer and more immediate opportunity for spontaneous reporting compared to typical college classroom experiences. Past research from the management literature suggests this may be the case. For instance, under certain conditions group members are attuned to the messages that managers send about whether and when 'organizational voice' will be positively versus negatively received (Burris, 2012). In this regard, it is worth noting that the current study was designed to lead participants to believe that reporting would be received as an act of loyalty to the group. A variation of the design might test to see how strongly vs. weakly fused individuals act when circumstances make it clear that blowing the whistle would be received as a challenge to the group. Conceivably, fusion alone may sufficiently motivate whistleblowing despite expectations of a hostile reaction. Alternatively, expectations of hostile reactions may greatly temper strongly fused members' desire to blow the whistle. Future variations on the current study's design could address these possibilities.

FUSION AND SPONTANEOUS REPORTING

The link between fusion and spontaneous reporting adds to a growing but still nascent literature on the behavioral consequences of different forms of identity fusion. Although spontaneous reporting represents a relatively mild form of going beyond the

call of duty compared to more extreme acts of self-sacrifice, studying less extreme pro-group behaviors has the distinct advantage of being much more tractable empirically. Studying milder forms of pro-group behaviors also move fusion's predictive validity forward substantially, as the bulk of past work has relied upon self-reports of extreme pro-group intentions (but see Swann et al., under review, for longitudinal evidence that fusion with one's preferred gender among transsexuals predicts sex change operations).

Going forward, it will important for research on identity fusion to continue to bridge the gap between pro-group behavior and intentions. For example, colleagues and I have recently found evidence that after the recent Boston Marathon terrorist attack, fusion with one's country was associated with pro-group self-sacrifice intentions, retrospective reports of victim support behaviors, and acts of support (e.g., money, written notes to victims) made available as part of the studies (Buhrmester, Fraser, & Swann, in prep). This work implies that fusion may trigger a plethora of pro-group actions, but presumably some boundary conditions exist. Future research that pushes such boundaries could fruitfully inform the nature of fusion.

COGNITIVE AND EMOTIONAL UNDERPINNINGS OF FUSION AND SPONTANEOUS REPORTING

Overall, several cognitive and emotion dimensions were associated with both fusion and spontaneous reporting. Based on self-reports of immediate affect after the test, negative emotions statistically mediated the relationship between fusion and spontaneous reporting. Convergent evidence from participants' interactions with confidants further supported the hypothesis that emotions serve as an antecedent to spontaneous whistleblowing action. In addition, the confidant tapes revealed that strongly fused students were more surprised by the cheating than weakly fused students, and surprise was a precursor to spontaneous reporting as well. These results buttress past evidence that strongly fused persons' experience group-related events more intensely than weakly fused persons (Buhrmester et al., 2012), and they uniquely translate the physiological arousal associated with emotions into pro-group action (Swann et al., 2010).

Looking to related literatures, findings from studies of punitive altruism are consistent with my analyses of negative emotions (Fehr & Gachter, 2002). In 'shared investment' economic games with others, participants' levels of anger at free riders in the group strongly predicted their decisions to give up personal gains to punish free riders. Anger and punishment was especially high amongst participants who tended to contribute more to the group than the average participant. The findings from the current study are conceptually analogous: those who say they give more than others to the group in general (i.e., strongly fused) were most apt to feel anger and spontaneously report the cheater. Other evolutionary theories of cooperation, such as signaling theory and kin selection are not equipped to explain findings such as these, and they broadly support a multi-dimensional account of the evolution of cooperation in societies.

These results also add another layer of understanding to the role of moral emotions in decision-making (see Hutcherson and Gross, 2011 for a review). Whereas past research has shown that certain moral violations elicit homogeneous emotional responses (e.g., Haidt, 2001), here I have shown that individual differences in fusion lead to different emotional responses, which in turn influence moral action vs. inaction. As with most other important social phenomena, these results support an interactionist approach to the study of whistleblowing that emphasizes both situational features and individual differences.

Strongly fused students' emotional responses to the wrongdoing were accompanied by cognitive concerns as well. In conversations with confidants and debriefers, strongly fused students voiced concerns about the cheater's negative effects on other students or the university more generally, expressed concerns about the egregiousness of the cheating, and voiced feeling personally responsible to report the cheater (i.e., a *lack* of diffused responsibility). Students' concerns about other students and the university at large echo evidence from the preliminary studies linking fusion and students' written thoughts about their relational and categorical ties to the university.

These results are also generally consistent with the pro-organizational behavior model of whistleblowing developed by Miceli and colleagues (2006). In it, they posit that

potential whistleblowers first estimate the severity or egregiousness of the wrongdoing. If judged to be significantly wrongful, potential whistleblowers then judge whether it is their responsibility or someone else's to report. In the current study, strongly fused students generally followed their model, finding the wrongdoing egregious and bucking pressure to diffuse responsibility to others.

COGNITIVE AND EMOTIONAL UNDERPINNINGS OF FUSION AND FORMAL REPORTING

Sixty-one percent of students who spontaneously reported the cheater went on to formally report him as well. That this percentage did not approach 100% reiterates the importance of understanding whistleblowing as a rich *process* that does not end after an initial claim is made. The steps required to formally report the wrongdoer were deliberately made to be personally costly in terms of time and effort, and participants figured these steps into their decision to formally report versus drop the initial claim. Based on participants' explanations of their actions during the debriefing, weakly fused students who spontaneously reported said they became selfishly concerned about the time and effort involved in formal reporting. Meanwhile, strongly fused students tended to reaffirm their feelings of personal responsibility and were not deterred by the personal costs involved. Interestingly, emotions driving the decision to spontaneously report and concerns about the egregiousness of the wrongdoing seemed to fall to the wayside for strongly fused students in the formal reporting stage. Instead, strongly fused students' sense of visceral oneness with the group caused them to continue to feel responsible to effect a positive group outcome. This feeling, in turn, motivated most of them to formally report the cheater.

LIMITATIONS & PRACTICAL IMPLICATIONS

The current investigation and results were qualified by several limitations. First, although past studies have demonstrated the mediating role of group-directed agency for fusion's effects on pro-group outcomes, the current study did not support such statistical mediation. While fusion was strongly associated with a self-reported measure of personal agency directed toward UT, the agency measure was only weakly related to spontaneous

reporting and unrelated to formal reporting. In neither case was agency a statistical mediator of fusion's link to reporting. Although it is unclear why this was the case, it is possible that the agency measure is better fitting for certain types of groups (e.g., country) than others. The fact that coded measures of responsibility were reliably related to fusion and reporting outcomes suggests this might be case.

Second, the study suffered somewhat from being underpowered for some statistical tests. Generally, an *N* of at least 50 and a continuous mediator variable is recommended for statistical mediation. In the current study, only 31 confidant paired participants completed the study and confidant and debriefing dimensions were coded using a binary present/absent system rather than a continuous coding scheme. Since the study involved highly choreographed deception and was time intensive from training, scheduling, and running standpoints, the total *N* for the study had to be limited (although the $N = 92$ surpassed the originally proposed $N = 80$).

Third, the only truly experimental part of the paradigm was the random assignment of participants to be with a confidant or alone. Future work could manipulate other key variables. For instance, it would be important to show that strongly fused students' whistleblowing actions are confined to situations involving the ingroup vs. an outgroup. Evidence from the current study that other types of fusion do not predict whistleblowing to the same extent as fusion with UT provides some evidence that this would likely be the case.

Fourth, and related to the previous point, since I did not experimentally manipulate either the independent variable or variables discussed as underlying mechanisms, I cannot with certainty claim causal directionality between variables. Establishing causality is a complicated process for researchers interested in highly stable constructs like fusion. However, I argue that certain features of the main investigation at least make the hypothesized causal directions more plausible than alternative hypotheses. First, unlike in a traditional cross-sectional statistical mediation involving multiple stable self-report variables, here I showed that a stable, attitude-like construct measured at the beginning of the study (i.e., fusion) later predicted *state* reactions to a discrete stimuli

(the cheater) which later predicted actual spontaneous and deliberate *behaviors* that were novel to the paradigm. The multiple methods and discrete time sequence employed strengthens the case for causality compared to self-report, cross-sectional designs. Second, I found no evidence for an alternative hypothesis arguing that emotion dispositions (i.e., agreeableness & neuroticism), not state emotions, could explain the statistical mediation effect. However, to establish causality more firmly, future work should explore novel methods that experimentally manipulate fusion and hypothesized mediators.

Finally, while the current studies add to the literature on the antecedents of whistleblowing, they do so with one major assumption in place – that whistleblowing is a pro-group action. As touched upon in the introduction, anecdotal evidence suggests that other motives exist as well and are often difficult to tease apart. For instance, some in the media have criticized the Enron whistleblower, Sherron Watkins, as blowing the whistle publicly to save her own skin during Enron's downfall. An accusation of selfishness is a common tactic to discredit whistleblowers, and in cases where there is some possible personal benefit following the act of whistleblowing, disentangling motives from one another is impossible. Future work might adapt an experimental paradigm such as mine to test the impact of adding in clear personal benefits to whistleblowing versus explicitly excluding the possibility of personal benefits.

The results from these studies could have practical implications. Much research has already uncovered the numerous positive effects of fostering high levels of organizational commitment (Moynihan & Pandey, 2007). My results suggest that organizations may wish to go a step further and seek out ways to develop high levels of fusion with one's organization. One possible avenue is to focus on a key difference between fusion and other perspectives on group alignment: the central importance of perceived relational ties to other group members. Numerous groups (e.g., sports teams, small businesses) emphasize the family-like ties between group members. A promising new avenue for researchers will be to uncover what causal antecedents foster the transition to believing that *all* group members, friends and strangers alike, share

irrevocable familial ties. Based on anthropological fieldwork, studies suggest that group rituals may be an especially important precursor (Whitehouse, 1995; 2004). A major insight from this line of work is how, in religious contexts, group members develop deep ties to one another as a result of engagement in highly arousing but low frequency ritual acts (e.g., dysphoric initiation rites). Secular organizations like fraternities and sororities and military groups seem to implicitly or explicitly realize the importance of these types of acts (e.g., hazing, Navy SEALs' "Hell Week"). It would be naïve to flatly recommend that other work and educational organizations simply follow suit, but ritual ceremonies like Texas A&M's annual "Aggie Muster," an emotional day for remembering deceased fellow Aggies, may work to build fusion in similar ways to more extreme ritual acts, albeit less intensely.

Most broadly, the current studies are representative of a larger endeavor that seeks to trace extreme, personally risky, and self-sacrificial pro-group behaviors to their earliest psychological roots. Group theorists in recent decades have largely approached group action from social cognition perspectives that minimize the importance of stable self-views that develop over years of social interaction, social comparison, and self-reflection. Identity fusion theory breaks from this tradition and as the current studies show, the theory continues to be useful as a guide to understanding the antecedents of impactful group actions.

Appendix A: Preliminary Study 1 Supplement

Two raters familiar with the fusion construct first read a portion of students' written responses. As they read responses, they looked for themes prevalent in responses and noted them. Raters then discussed their notes and developed 10 specific dimensions that they felt represented 5 broader constructs. These 5 broad dimensions and their sub-dimensions are presented in the left-hand column below. A second pair of raters then judged each response according to the scheme in the rightmost column below. Descriptions capturing the meaning of each dimension and examples were provided by the initial coding team. A third rater then compared the judgments of each of the two raters, resolved differences in judgments, and produced the final ratings for analysis.

Table 9: Coding dimensions and subdimensions, number of responses that dimension was judged to be present, and example response.

Coding Dimensions	***N***	***Example Responses***
1. Reciprocal Strength -- UT strengthens me -- I strengthen UT	10	Coded 0 = no strength/hurts: "i feel that intellectually i might be under the standards that are being asked of me." Coded 1 = high strength: "I strive to improve myself in order to make the school look better"
2. Individuation vs. Oneness -- Self and UT part vs whole	49	Coded 0 = little/no overlap: "I consider being a ut student a small part of my life."; Coded 1 = partial overlap:"UT is a huge part of my life, but I am still an individual."; Coded 2= near/total overlap: "I feel almost 100% a part of UT."
3. Emotionality & Efficacy -- Positive vs negative emotions --Sense of competence	26	Coded 0 = negative emotion: "I don't like how many students go here. I get very uncomfortable." Coded 1=positive emotion: "Whenever i go back home i always say with pride and joy i go to the University of Texas at Austin."
4. Relational ties -- Social relationships w/ students -- Shared interests w/ students	27	Coded 0 = weak ties: "I have not been out much and met a lot of new people." Coded 1 = strong ties: "I have made more friends than what I though i would have made since I came in this fall."

Table 9 continued.

5. Categorical ties -- Connection to UT the institution -- Involvement with UT & UT orgs. -- Sense of place & community	51	Coded 0 = weak ties: “I am a freshman at UT and therefore have not really established a high sense of school spirit yet.” Coded 1 = strong ties: “I feel very connected to UT Austin because it's a great school in a great city”

Appendix B: Preliminary Study 4 Supplement

For the hypothetical whistleblowing scenarios, response options for each ranged in effort from ‘do nothing’ (A; 0) to ‘whatever is necessary to correct the wrongdoing’ (E; 4). The three scenarios were as follows:

1. You have recently learned that a co-worker has been violating one of your company's policies on an ongoing basis. This violation is quite serious and would look very bad for your department if it were discovered. The coworker has more seniority than you, and he/ she is well respected in the company. You are unsure if his/her behavior is illegal, but you are sure it is not right. Response options: A Say nothing. B Confirm the violation only if I was specifically asked by management. C Mention the violation to the co-worker only, hoping that it would stop. D Report the violation to my supervisor only, letting it drop after that. E Make sure that the violation was corrected, no matter how far I would have to go.

2. In the course of your normal job duties, you have come across some organizational practices which are discriminatory towards select groups of employees. You are not a member of any of the groups being discriminated against, however. It appears that the practices have been going on for as long as the company has been in business. The information is not public, and under normal circumstances it will never be. There is no way that anyone will ever know that you have found out about these practices. Response options: A Say nothing. B Casually mention to my supervisor that I was concerned about discriminatory practices. C Quietly question the practices, stopping when resistance was given. D Openly question the practices, going as far as necessary within the company hoping to implement changes. E Openly question the practices within the company, and if necessary go public, insisting on changes.

3. This year has not been a very good one for your company. Sales are down significantly, which has led to the lay off of many excellent employees. In addition, the company is facing a major lawsuit, which has the potential of bankrupting the firm. A prospective client has indicated an interest in placing a very large order which would guarantee the reinstatement of all laid-off employees, and would keep them employed for

quite some time. However, the client is wary about the stability of your company. To improve the client's perception of the company, the sales manager has removed the reference to the lawsuit in the copy of the annual report given to the potential client. One responsibility of your position is to ensure the accuracy of company documents. Since few people ever read the entire report, it is doubtful that anyone else would notice the missing information. Response options: A Do nothing. B Talk with the sales manager, letting it drop after that. C Ask the sales manager to provide the client with complete information. D Talk with the top executives of the company, asking that the client be provided with complete information. E Talk with the top executives of the company, and make sure that the client received the complete report. E Do whatever was necessary to make sure that the client received the complete report.

Following the analytic strategy of Sims and Keenan (1998), extreme responses (4) were interpreted as intention to engage in whistleblowing and recoded as "1". The other response options (0-3) were recoded as "0" to represent no intention to blow the whistle. These recoded responses were then summed (range 0-3). Alternative methods of indexing intentions also showed weak, positive relationships to fusion that hovered above traditional significance levels (i.e., mean of raw 0-4 responses, dichotomization based on providing one or more extreme response vs. none). Upon examination of each scenario separately, it was found that fusion scores were most strongly related to whistleblowing intentions for the scenario involving a co-worker violating a company policy where the violation was left to the participant to imagine, $r = .20$, $p<.01$. This significant result in part informed the design of the lab scenario, as conceptually they are somewhat similar (i.e., co-worker violating company policy vs. a fellow student cheating on a test, a clear violation of university ethics).

Appendix C: Main Investigation Supplement

Table 10: Coded Dimensions and Examples

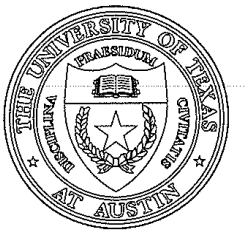
Dimension	Confidant Examples	Debriefing Examples
1. Immediately brought up the cheater in conversation with confidant	Coded 1 = brought up cheating shortly after entered room. Vs. Coded 0 = confidant chatted about test casually after several minutes.	--
2. Described cheating in detail to confidant	Coded 1 = detailed: "And then the coordinator put the key down on the table. And then, just now, the guy like took the key, put it in his lap…" vs. Coded 0 = not detailed or absent.	--
3. Expressed negative emotion about cheater	Coded 1 = high: "No, I'm just pissed off… Did you see that?" vs. Coded 0 = absence of negative emotion.	1: "I got so mad. I was like shaking. I was just so nervous."
4. Expressed surprise about cheating	Coded 1 = high: "haha I was like is this really happening?...And I was just like so confused I was waiting for the guy next to me he was just staring at him."; "Well that dude snapped! And I was like, 'Oh, that's the key!' vs. Coded 0 = absence of surprise.	0: "I thought just that you could use the credit and thought that's cool"
5. Expressed selfish concern for oneself	Coded 1 = high: "Um um I I I just need the two hours to get all five." Vs. Coded 0 = absence.	1: "I wasn't really thinking about it that much and just kind of wanted to be done with the survey."; 0 = "I was sure I wasn't going to get the credit, but I don't know, I felt you were going to screw up the whole thing."

Table 10 continued.

6. Expressed unselfish concern for other students or UT more generally	Coded 1 = high: "…but now he's [expletive] up that guys study because those are all false he doesn't know anything about UT."; "The kid just like grabbed the [expletive] key and its just like not fair to his study, it's not fair to anyone else…so he just messed up that uys study and ruined everyone else's chances" Vs. Coded 0 = absent.	1: "I was like this is going to be skewed. Everyone is going to be pissed at that guy cheating." ; "[The cheater] could have gotten extra hours and someone else kind of deserved it."
7. Expressed concern regarding egregiousness of cheating	Coded 1 = high: "it's like you're in college now sweetie. it's like be done with cheating. This is your life." Vs. Coded 0 = low: "But who cares, it's not that big of a deal"; "It's not that big of a deal! I didn't even finish it…you can get that hour on an online survey"; "Yeah it's like c'mon whatever. It's just like a quiz thing."	1: "I thought it was unfair and you were getting *extra* hours" Vs. 0: "I just didn't feel like it was a big enough deal."; "Well, like it's not fair if they do that based on the hour, but it's whatever…it's no big deal."; "Yeah, um, I don't know I just didn't care."
8. Expressed responsibility to report cheater	Coded 1 = high: "No worries I might have been the only one to see him. I'll say something."; "Yeah well it may not be our business, but it is because I saw him"; Coded 0 = low responsibility: "I mean I'm not gonna be a narc"; "I mean, I'm not even the only one there. I mean, like, he's fine…I'm good, you know what I'm saying?" (also coded as low thought clarity).	1: "I mean I saw what I saw and I could have been the only one who saw it" vs. 0: "Eh, not really any of my business."; "Yeah, but I didn't really want to get involved."
9. Clearly expressed thoughts and feelings	Coded 1 = did not contradict or sound uncertain. Vs. Coded 0 = low clarity: "yeah it's like c'mon whatever…I don't know, whatever"	0: "Umm…like…I don't know really."

Figure 4: The Experimental Misconduct Form

Below is the bogus 'misconduct form' that participants completed if they desired to formally report the cheater.



UNIVERSITY OF TEXAS AT AUSTIN
OFFICE OF RESEARCH SUPPORT
INTERNAL REVIEW BOARD
P.O. BOX 7426
MAIL CODE A3200
AUSTIN, TX 787513

EXPERIMENTAL MISCONDUCT FORM

This form is to be completed by the participant should any form of experimental misconduct occur during the course of a student's time as a participant. Completion of this form is at the discretion of the participant. By completing this form, the participant *agrees to be contacted by the IRB Office for follow-up questioning.*

The researcher conducting the study should provide the participant with this form in a private space. If you as the participant would like to obtain information about a research study, or if you have questions, concerns, complaints or wish to discuss problems or your rights as a research subject with someone unaffiliated with the study, please contact the IRB Office at -------. Anonymity, if desired, will be protected to the extent possible. As an alternative method of contact, an email may be sent to ------- or a letter sent to IRB Administrator, P.O. Box -------.

TO BE COMPLETED BY THE PARTICIPANT AT THE TIME OF THE INCIDENT:

TODAY'S TIME AND DATE:
YOUR LOCATION:
EXPERIMENT STUDY NAME:
YOUR NAME:
YOUR EMAIL:
YOUR TELEPHONE #:
PLEASE DESCRIBE ANY MISCONDUCT THAT YOU WITNESSED IN AS MUCH DETAIL AS POSSIBLE:

Table 11: Correlations between Big Five Indices, Fusion, and Whistleblowing

Big Five Indices	Fusion	Spontaneous Reporting	Formal Reporting
Openness	.36**	.24	.10
Conscientiousness	.23	.33**	.20
Extraversion	.30*	.12	.10
Agreeableness	.22	.26*	.21
Neuroticism	-.17	-.24	-.17

Note: *: $p < .05$, **: $p < .01$

Table 12: Correlations between Fusion, Whistleblowing, and Demographic Variables

Demographic Variables	Fusion	Spontaneous Reporting	Formal Reporting
Age	-.13	-.15	-.12
Gender	-.02	-.21*	-.09
# of years as UT student	-.17	-.04	-.12
# of UT family members	-.01	.16	.04

Note: *: $p < .05$, **: $p < .01$

References

Ainslie, G. (1992). Picoeconomics: The strategic interaction of successive motivational states within the person. Cambridge, England: Cambridge University Press.

Aron, A., Aron, E., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology 60*, 241-253.

Baker F., Johnson, M.W., & Bickel, W.K. (2003). Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *Journal of Abnormal Psychology, 112*(3), 382-92.

Bergman, M.E., Langhout, R.D., Cortina, L.M., & Fitzgerald, L.F. (2002). The (un)reasonableness of reporting: Antecedents and consequences of reporting sexual harassment. *Journal of Applied Psychology, 87,* 230-242.

Bjørkelo, B., Einarsen, S. and Matthiesen, S. B. (2010), Predicting proactive behaviour at work: Exploring the role of personality as an antecedent of whistleblowing behaviour. *Journal of Occupational and Organizational Psychology, 83,* 371–394.

Brewer, M. B., & Gardner, W. L. (1996). Who is this "we"? Levels of collective identity and self representations. Journal of Personality and Social Psychology, 71, 83-93.

Brief, A.P., & Motowidlo, S. (1986). Prosocial organizational behaviors. *Academy of Management Review, 4,* 710-725.

Brown, B. S., & Choong, P. (2005). An investigation of academic dishonesty among business students at public and private United States universities. *International Journal of Management, 22(2),* 201–214.

Buhrmester, M.B., Fraser, W.T., & Swann, W.B., Jr. (in prep). Who makes big sacrifices for big groups?

Buhrmester, M., Kwang, T., & Gosling, S.D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science, 6(1),* 3-5.

Buhrmester, M., Gómez, Á., Brooks, M. L., Morales, J. F., Fernández, S., & Swann, W. B. Jr. (2012). My group´s fate is my fate: "Fused" Americans and Spaniards link personal life quality to outcome of ´08 elections. *Basic and Applied Social Psychology, 34,* 527-533.

Burris, E.R. (2012). The risks and rewards of speaking up: Managerial responses to employee voice. *Academy Of Management Journal*, *55*(4), 851-875.

Burton, B. K., & Near, J. P. (1995). Estimating the incidence of wrongdoing and whistle-blowing: results of a study using randomized response technique. *Journal of Business Ethics, 14,* 18–30.

Carozza, D. (2007). Interview with Sherron Watkins: Constant warning. *Fraud Magazine.*

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

Dozier, J.B., & Miceli, M.P. (1985). Potential predictors of whistle-blowing: a prosocial behavior perspective. *Academy of Management Review, 10,* 823-836.

Dutton, J.E., Ashford, S.J., O'Neill, R.M., Hayes, E., & Wierba, E.E. (1997). Reading the wind: How middle managers assess the context for selling issues to top managers. *Strategic Management Journal, 18*¸407-423.
Eysenck S.B., Eysenck H.J., Barrett P. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences. 6,* 21–29.
Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature, 415,* 137-140.
Fehr, E. & Fischbacher, U. (2005). Human altruism – Proximate patterns and evolutionary origins. *Analyse & Kritik, 27,* 6-47.
Fritz, M.S., & McKinnon, D.P. (2003). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233-239.
Goldberg, L.R. (1999). A Broad-Bandwidth, Public-Domain, Personality Inventory Measuring the Lower-Level Facets of Several Five-Factor Models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe,* Vol. 7. (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
Gómez, Á, Brooks, M.L., Buhrmester, M., Vázquez, A., Jetten, J., & Swann, W.B. (2011). On the nature of identify fusion: Insights into the construct and a new measure. *Journal of Personality and Social Psychology, 100,* 918-933.
Gosling, S.D., Ko, S.J., Mannarelli, T., Morris, M.E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 379,* 378-398.
Graham, J.W. (1986). Principled organizational dissent: A theoretical essay. In L.L. Cummings and B.M. Staw (Eds.), *Research in organizational behavior* (Vol. 8, pp.1-52). Greenwich, CT: JAI Press.
Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96,* 1029-1046.
Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin, 130,* 769-792.
Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108,* 814-834.
Harding, T.S., Carpenter, D.D., Finelli, C.J., & Passow, H.J. (2004). Does academic dishonesty relate to unethical behavior in professional practice? An exploratory study. *Science and Engineering Ethics, 10,* 1-14.
Hayes, A. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis. New York: Guilford Press.
Healy, P.M., & Palepu, K. (2003). The Fall of Enron. *Journal of Economic Perspectives, 17,* 3.
Helmreich, R., Spence, J. T., & Stapp, J. (1974). Short form of the Texas Social Behavior Inventory. *Bulletin of the Psychonomic Society, 4,* 473-475
Honz, K., Kiewra, K.A., & Yang, Y.S. (2010). Cheating perceptions and Prevalence across academic settings. *Mid-Western Educational Reader, 23,* 10-17.

Hornsey, M. J., Blackwood, L., Winnifred, L., Fielding, K., Mavor, K., Morton, T. White, K. M. (2006). Why do people engage in collective action? Revisiting the role of perceived effectiveness. *Journal of Applied Social Psychology, 36*, 1701–1722.

Hutcherson, C.A., & Gross, J.J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology, 4,* 718-737.

Jetten, J., Spears, R., Hogg, M. A., & Manstead, A. S. R. (2000). Discrimination constrained and justified: Variable effects of group variability and in-group identification. *Journal of Experimental Social Psychology, 36*, 329–356.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.

King, G., III. (1997). The effects of interpersonal closeness and issue seriousness on blowing the whistle. *Journal of Business Communication, 34,* 419-436.

Klein, H. A., Levenburg, N. M., McKendall, M., & Mothersell, W. (2006). Cheating during the college years: how do business school students compare? *Journal of Business Ethics, 72,* 197–206.

Latane, B., & Darley, J. (1969). Bystander "Apathy". *American Scientist, 57,* 244-268.

Lee, J.Y., Heilmann, S.G., & Near, J.P. (2004). Blowing the whistle on sexual harassment: Test of a model of predictors and outcomes. *Human Relations, 57*, 297-322.

Mael, F. A., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. Journal of Organizational Behavior, 13, 103–123.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98, 224-253.

Mesmer-Magnus, J.R., & Viswesvaran, C. (2005). Whistleblowing in organizations: An examination of correlates of whistleblowing intentions, actions, and retaliation. *Journal of Business Ethics, 62,* 277-297.

Miceli, M.P. & Near, J.P. (1992). Blowing the whistle: The organizational and legal implications for companies and employees. New York: Lexington.

Miceli, M.P., & Near, J.P. (1997). Whistle-blowing as antisocial behavior. In R. Giacalone & J. Greenberg (Eds.), *Antisocial behavior in organizations* (pp. 130-149). Thousand Oaks, CA: Sage Publications.

Miceli, M.P., Near, J.P., & Dworkin, T.M. (2008). Whistle-blowing in Organizations. New York: Routledge.

Miceli, M.P., Near, J.P., & Schwenk, C.P. (1991). Who blows the whistle and why? *Industrial and Labor Relations Review, 45,* 113-130.

Miceli, M.P., Rehg, M., Near, J.P., & Ryan, K. (1999). Can laws protect whistle-blowers? Results of a naturally occurring field experiment. *Work and Occupations, 26,* 129-151.

Miethe, T.D. (1999). Whistle-blowing at work: Tough choices in exposing fraud, waste and abuse on the job (2nd ed.). Boulder, CO: Westview Press.
Milgram, S. (1974). Obedience to authority: An experimental view. Harper Collins: New York.
Near, J.P., Van Scotter, J.R., Rehg, M.T., & Miceli, M.P. (2004). Does type of wrongdoing affect the whistle-blowing process? *Business Ethics Quarterly, 14,* 219-242.
Near, J.P., & Miceli, M.P. (1985). Organizational dissidence: The case of whistle-blowing. *Journal of Business Ethics, 4,* 1-16.
Near, J.P., & Miceli, M.P. (1996). Whistle-blowing: Myth and reality. *Journal of Management, 22,* 507-526.
Ostaszewski, P. (1996). The relation between temperament and rate of temporal discounting. *European Journal of Personality*, 10, 161–172.
Pennebaker, J. W., Booth, R. E., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC2007: Operator's manual. Austin, TX: LIWC.net
Semnani-Azad, Z., Sycara, K.P., & Lewis, M. (2012). Dynamics of helping behavior and cooperation across culture. *Proceedings of Collaboration Technologies and Systems* 2012, 525-530.
Sierles, F. S., Kushner, B. D., & Krause, P. B. (1988). A controlled experiment with a medical student honor system. *Journal of Medical Education, 63*, 705–712.
Simon, C. A., Carr, J. R., McCullough, S. M., Morgan, S. J., Oleson, T., & Ressel, M. (2004). Gender, student perceptions, institutional commitments and academic dishonesty: who reports in academic dishonesty cases? *Assessment & Evaluation in Higher Education, 29*(1), 75–90.
Sims, R.L., & Keenan, J.P., (1998). Predictors of external whistleblowing: Organizational and intrapersonal variables. *Journal of Business Ethics*, 17, 411-421.
Smith, D.B., & Ellingson, J.E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology, 87,* 201-209.
Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. In J. Suls & A. G. Greenwald (Eds.), Social psychological perspectives on the self (Vol. 2, pp. 33-66). Hillsdale, NJ: Erlbaum.
Swann, W. B., Jr. (2011). Self-verification theory. In P. Van Lang, A. Kruglanski, & E.T. Higgins (Eds.) Handbook of Theories of Social Psychology, Sage: London.
Swann Jr, W. B., & Rentfrow, P. J. (2001). Blirtatiousness: cognitive, behavioral, and physiological consequences of rapid responding. *Journal of Personality and Social Psychology, 81(6),* 1160.
Swann, W. B., Jr., Stein-Seroussi, A., & Giesler, B. (1992). Why people self-verify. Journal of Personality and Social Psychology, 62, 392-401.
Swann, W. B. Jr., Gómez, A., Dovidio, J. F., Hart, S. & Jetten, J. (2010b). Dying and killing for one's group: Identity fusion moderates responses to intergroup versions of the trolley problem. Psychological Science, 21, 1176–1183.

Swann, W.B., Jr., Gómez, A., Huici, C., Morales, F., & Hixon, J. G. (2010a). Identity fusion and self-sacrifice: Arousal as catalyst of pro-group fighting, dying and helping behavior. Journal of Personality and Social Psychology, 99, 824–841.

Swann, W. B. Jr., Gómez, A., Seyle, C. D., Morales, J. F. & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. Journal of Personality and Social Psychology, 96, 995–1011.

Swann, W. B., Jr., Jetten, J.,Gómez, Á. Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review, 119*, 441-456.

Swann, W.B., Jr., Gomez, A., Vazquez, A., Gomez-Gil, E., Segovia, S., Carrillo, B., & Guillamon, A. (under review). Identity and the choices transsexuals make: Fusion with preferred sex predicts sex change surgery. Manuscript submitted for publication.

Swann, W.B., Jr., Stein-Seroussi, A. & Giesler, B. (1992). Why people self-verify. *Journal of Personality and Social Psychology, 62*, 392-401.

Swanner, J. K., Beike, D. R., & Cole, A. T. (2010). Snitching, lies and computer crashes: An experimental investigation of secondary confessions. *Law and Human Behavior; Law and Human Behavior, 34*(1), 53.

Swartz, M., & Watkins, S. (2003). *Power failure: The inside story of the collapse of Enron*. Doubleday Books.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.). *The social psychology of intergroup relations* (pp. 33–47). Monterrey, CA: Brooks-Cole.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29(1)*, 24-54.

Trevino, L.K. & Weaver, G.R. (2001). Organizational justice and ethics program "follow-through": Influence on employees' harmful and helpful behavior. *Business Ethics Quarterly, 11,* 651-671.

Trevino, L.K., Butterfield, K.D., McCabe, D.L. (1998). The ethical context in organizations: Influences on employee attitudes and behaviors. *Business Ethics Quarterly, 8,* 447-476.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). Rediscovering the social group: A self-categorization theory. Oxford, England: Basil Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. Personality and Social Psychology Bulletin, 20, 454-463.

Van Dyne, J.A., & LePine, L. (1998). Helping and voice extra role-behaviors: Evidence of construct and predictive validity. *Academy of Management Journal, 41,* 108-119.

Van Knippenberg, D. (2000). Work motivation and performance: A social identity perspective. *Applied Psychology, 49,* 357–371.

Watson, D., & Clark, L.A. (1994). The PANAS-X : Manual for the Positive and Negative Affect Schedule – Expanded Form.

Wells, G.L. & Olson, E.A. (2003). Eyewitness Testimony. *Annual Review of Psychology, 54*, 277-95.
Whitehouse, H. (1995). Inside the cult: Religious innovation and transmission in Papua New Guinea. Oxford, England: Oxford University Press.
Whitehouse, H. (2004). Modes of religiosity: A cognitive theory of religious transmission. Walnut Creek, CA: AltaMira Press.
Withey, M.J. & Cooper, W.H. (1989). Predicting exit, voice, loyalty, and neglect. *Administrative Science Quarterly, 34*, 521-539.
World Bank, 2004. The Costs of Corruption. http://go.worldbank.org/LJA29GHA80.
Zhang, J., Chiu, R.K., & Wei, L., (2009). Decision making processes of internal whistleblowing behavior in China: Empirical evidence and implications. *Journal of Business Ethics, 88,* 25-41.
Zuckerman, M. (1971). Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology, 36,* 45–52.

Vita

Michael Duane Buhrmester attended Plano Senior High School in Plano, Texas. In 2003, he began his undergraduate studies at the University of Texas at Austin and was a Dedman Scholar. In 2007, Michael received his B.A. in Plan II Honors and Psychology with Honors. He then began his graduate studies in social and personality psychology at the University of Texas at Austin.

Permanent address (or email): Buhrmester@gmail.com

This dissertation was typed by Michael Duane Buhrmester.