

Wielding a Double-Edged Sword: Public Moral Advocates are Derogated Yet Influential

by

Nadia Yasmine Bashir

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Department of Psychology
University of Toronto

© Copyright by Nadia Yasmine Bashir 2014

Wielding a Double-Edged Sword: Public Moral Advocates are Derogated Yet Influential

Nadia Yasmine Bashir

Doctor of Philosophy

Department of Psychology
University of Toronto

2014

Abstract

Individuals who publicly challenge a group's wrongdoing (e.g., whistleblowers, group critics, and individuals who receive authorization to declassify group information) are frequently reviled by group members. Past research suggests that these moral advocates are derogated because they implicitly reproach group members or make them feel immoral. Because moral advocates who take a public stand against wrongdoing expose or highlight a group's problematic practices, however, group members may instead derogate these advocates because the advocates make the group *look* immoral in the eyes of others. I examined this directly. I also examined whether threatening the group's social image may allow moral advocates to succeed in promoting change within the group, despite being derogated. Moral advocates who publicly challenged group wrongdoing were derogated by group members but were nevertheless successful in raising concern about the need for group change. Moral advocates who challenged wrongdoing privately and therefore posed less of a threat to the group's social image evaded derogation but were also unsuccessful in promoting change. Social image concerns rather than moral reproach or moral integrity concerns mediated these effects. Thus, public moral advocates wield a double-edged sword: Publicly threatening the group's social image leads moral advocates to be derogated, but it also enables them to promote group change.

Acknowledgments

I would like to thank my committee members, Penelope Lockwood, Alison Chasteen, and Wil Cunningham, for their guidance and support throughout the completion of this research. I am also grateful to Beatrice Bejan, Leah McSpadden, Man-On Tong, Sylvia Wang, and Chuqing Yang for their assistance with data collection.

Table of Contents

Acknowledgments.....	iii
Table of Contents.....	iv
List of Figures	x
List of Appendices	xi
Chapter 1 Introduction.....	1
1 Reactions to Moral Advocates	1
2 Social Image Concerns.....	3
3 Public Moral Advocacy as a Double-Edged Sword.....	5
4 The Present Research	7
Chapter 2 Study 1.....	9
1 Method.....	10
1.1 Participants.....	10
1.2 Procedure	10
2 Results	11
2.1 Manipulation checks	12
2.1.1 Target disclosure.....	12
2.1.2 Target group membership.....	12
2.2 Potential mediating variables.....	12
2.2.1 Social image concerns.....	12

2.2.2	Moral reproach.....	13
2.3	Dependent variables.....	13
2.3.1	Target evaluations.....	13
2.3.2	Group improvement concerns.....	13
2.4	Mediation Analyses	14
3	Discussion	14
Chapter 3 Study 2.....		16
1	Method.....	16
1.1	Participants.....	16
1.2	Procedure	16
2	Results	17
2.1	Manipulation checks	17
2.1.1	Wrongdoer group membership	17
2.1.2	Target group membership	17
2.2	Potential mediating variables	17
2.2.1	Social image concerns.....	17
2.2.2	Moral reproach.....	18
2.3	Dependent variables.....	18
2.3.1	Target evaluations.....	18
2.3.2	Group improvement concerns.....	18

2.4	Mediation analyses.....	18
3	Discussion	19
	Chapter 4 Study 3.....	20
1	Method.....	20
1.1	Participants.....	20
1.2	Procedure	21
2	Results	22
2.1	Manipulation check.....	22
2.2	Potential mediating variables	22
	2.2.1 Social image concerns.....	22
	2.2.2 Personal moral reproach	22
	2.2.3 Group moral reproach	22
	2.2.4 Personal moral integrity	22
	2.2.5 Group moral integrity	22
2.3	Dependent variables.....	23
	2.3.1 Target evaluations.....	23
	2.3.2 Group improvement concerns.....	23
2.4	Mediation analyses.....	23
3	Discussion	24
	Chapter 5 Study 4.....	25

1	Method.....	26
1.1	Participants.....	26
1.2	Procedure	26
2	Results	27
2.1	Manipulation checks	27
2.1.1	Target disclosure.....	27
2.1.2	Disclosure authorization	27
2.2	Potential mediating variables	27
2.2.1	Social image concerns.....	27
2.2.2	Personal moral reproach	27
2.2.3	Group moral reproach	28
2.2.4	Personal moral integrity	28
2.2.5	Group moral integrity	28
2.3	Dependent variables.....	28
2.3.1	Target evaluations.....	28
2.3.2	Group improvement concerns.....	29
2.4	Mediation analyses.....	29
3	Discussion	29
	Chapter 6 Study 5.....	32
1	Method.....	32

1.1	Participants.....	32
1.2	Procedure	32
2	Results	33
2.1	Manipulation check.....	33
2.1.1	Target disclosure.....	33
2.2	Potential mediating variables	33
2.2.1	Social image concerns.....	33
2.2.2	Personal moral reproach	33
2.2.3	Group moral reproach.....	34
2.3	Dependent variables.....	34
2.3.1	Target evaluations	34
2.3.2	Group improvement concerns.....	34
2.4	Mediation analyses.....	35
3	Discussion	35
Chapter 7 Study 6.....		37
1	Method.....	37
1.1	Participants.....	37
1.2	Procedure	38
2	Results	38
2.1	Manipulation checks	38

2.1.1	Target disclosure.....	38
2.1.2	Disclosure method	38
2.2	Potential mediating variables.....	39
2.2.1	Social image concerns.....	39
2.2.2	Personal moral reproach	39
2.3	Dependent variables.....	39
2.3.1	Target evaluations.....	39
2.3.2	Group improvement concerns.....	39
2.4	Mediation analyses.....	40
3	Discussion	40
	Chapter 8 General Discussion.....	42
1	Theoretical Implications.....	42
2	Limitations and Future Directions.....	46
3	Conclusion.....	48
	References.....	49
	Notes.....	63
	Appendices.....	81
	Figures.....	82

List of Figures

Figure 1. Mediation models tested in Study 1	82
Figure 2. Target evaluations and group improvement concerns as a function of target disclosure and disclosure method in Study 6.....	83

List of Appendices

Appendix A: Sample Target Disclosure Articles.....	63
Appendix B: Social Image Concerns Measure	75
Appendix C: Target Evaluations Measure	76
Appendix D: Group Improvement Concerns Measure	77
Appendix E: Personal Moral Reproach Measure.....	78
Appendix F: Group Moral Reproach Measure	79
Appendix G: Personal Moral Integrity Measure.....	80
Appendix H: Group Moral Integrity Measure	81

Chapter 1

Introduction

Over the course of history, many individuals have taken a stand against a group's wrongdoing by challenging it publicly. In 2004, for example, Joseph Darby revealed that American military personnel had abused Iraqi soldiers at the Abu Ghraib prison (Schorn, 2009). Similarly, in 2012, Russian resident Irina Brovkina publicly criticized the Russian government for violating human rights while constructing venues for the 2014 Winter Olympics (Buchanan & Gorbunova, 2012). Because individuals who challenge wrongdoing publicly generally act with good intentions, it seems reasonable that group members should support these individuals and embrace their calls for change within the group. In reality, however, individuals who challenge wrongdoing publicly (i.e., public moral advocates) are frequently reviled. Indeed, Darby and Brovkina were ostracized by members of their communities, became victims of smear campaigns, and received threatening messages for publicly challenging the groups' practices (Buchanan & Gorbunova, 2012; Schorn, 2009). Thus, although public moral advocates may be heroes to some people, others, particularly members of the groups that are accused of wrongdoing, may fear that these advocates have tarnished the group's social image. For example, American soldiers, and even Americans more generally, may have been concerned that the horrific events of Abu Ghraib would reflect poorly on them in the eyes of the international community, because these events were made public. Thus, public moral advocates may elicit negative reactions from group members because these advocates seem to threaten the group's social image. In the present research, I examined this directly.

1 Reactions to Moral Advocates

Individuals generally agree that questionable, harmful, and illegal behaviours are wrong and unethical (Brambilla, Sacchi, Rusconi, Cherubini, & Yzerbyt, 2012; Brown, González, Zagefka, Manzi, & Čehajić, 2008; Griskevicius, Tybur, & Van den Bergh, 2010). Thus, it may seem counterintuitive that individuals often vilify those who aim to expose and therefore eliminate these behaviours. A growing body of research demonstrates, however, that individuals who take a stand against immoral behaviours (i.e., moral advocates) can elicit scorn rather than admiration from others. Indeed, when individuals call attention to discrimination, they are viewed negatively by observers (Garcia, Reser, Amo, Redersdorff, & Branscombe, 2005; Kaiser, Dyrenforth, &

Hagiwara, 2006; Kaiser & Miller, 2001; Stone, Whitehead, Schmader, & Focella, 2011).

Similarly, individuals who advocate the need to address key social concerns (e.g., civil rights and environmental activists) receive negative evaluations from perceivers (Bashir, Lockwood, Chasteen, Nadolny, & Noyes, 2013). In addition, perceivers react negatively to individuals who criticize groups for being intolerant or racist; Hornsey, 2005; Hornsey & Imani, 2004; Hornsey, Oppes, & Svensson, 2002). Furthermore, perceivers dislike individuals who confront others for making sexist or racist remarks (Czopp & Montith, 2003; Czopp, Montith, & Mark, 2006). Thus, there is clear evidence that individuals often react negatively to those who take a stand against wrongdoing. The psychological mechanisms underlying these negative reactions, however, are less clear.

Recent research suggests that derogation of moral advocates stems from threatening social comparisons and subsequent feelings of moral reproach (Monin, Sawyer, & Marquez, 2008). Specifically, when an individual refuses to perform a morally questionable behaviour in which perceivers have engaged, perceivers anticipate that the individual will compare his/her own morally “superior” behaviour to that of perceivers. This leads perceivers to view the individual as someone who implicitly condemns them (Monin, 2007; Monin & O’Connor, 2011). Because morality is central to individuals’ self-concepts (Allison, Messick, & Goethals, 1989; Sabini & Silver, 1982), social comparisons in the moral domain tend to be highly potent. That is, the experience of being looked down upon by a morally “superior” other can be so threatening to individuals that it elicits defensive reactions, including derogation (Monin et al., 2008). Indeed, even if individuals do not believe that the “superior” other is more moral than they are, the thought of being morally reproached by this person is sufficiently threatening that it leads individuals to derogate this person (Monin, 2007). For example, because vegetarians refuse to eat meat, meat-eaters assume that vegetarians implicitly criticize meat-eaters from higher moral ground. This ultimately leads meat-eaters to derogate vegetarians (Minson & Monin, 2012).

Past research therefore provides valuable insight into the important roles of social comparison and moral reproach in eliciting derogation of moral advocates who seem to implicitly criticize the morality of others. It may also explain derogation of moral advocates who explicitly condemn the behaviour of others (e.g., an individual who directly confronts a bigot; Czopp & Monteith, 2003). In past work, however, researchers have focused on examining moral advocates who take a relatively private stand against wrongdoing (e.g., the advocates wrote a note on a

study questionnaire indicating their refusal to comply with a morally questionable study task or they confronted an individual in a one-on-one interaction; Czopp et al., 2006; Monin et al., 2008).

In many cases, in comparison, moral advocates take a stand against wrongdoing by challenging it publicly. For example, whistleblowers expose unethical practices committed by organizations (e.g., Campbell, 2013; Schorn, 2009), public leaders condemn the human rights and environmental records of other nations (e.g., Dehghan, 2014; Tejada & Burkitt, 2013), and government employees decide to declassify information that calls a group's behaviour into question (O'Keefe, 2014; Savage, 2014). In these cases, members of groups that have been accused of wrongdoing may react negatively to the moral advocate, not because they are concerned about being reproached by the advocate but rather because they are concerned about how the advocate has made the group look in the eyes of others (Sumanth, Mayer, & Kay, 2011). That is, concerns about a damaged social image may have a greater impact on perceivers than do concerns about being reproached by a seemingly sanctimonious advocate. When Joseph Darby exposed the Abu Ghraib scandal, for example, his colleagues were likely not terribly concerned about Darby's own impressions of them. Instead, they were probably far more concerned about the military's image in the eyes of the American public and international community. Indeed, following Darby's disclosure, members of the American military urged the public to avoid forming negative impressions of the military based on acts that were committed by a limited number of individuals (Leung, 2009). Such reactions appear to reflect concern about the group's social image and are consistent with the possibility that social image concerns play a key role in influencing how individuals react to moral advocates who challenge their group's behaviour publicly.

2 Social Image Concerns

When individuals recognize that they have committed a failure (e.g., performed an immoral behaviour or demonstrated incompetence), they generally become concerned about how others view them. Although many researchers have previously conceptualized these social image concerns as a component of shame (e.g., de Hooge, Breugelmans, & Zeelenberg, 2008; de Hooge, Zeelenberg, & Breugelmans, 2010; Lewis, 1971; Smith, Webster, Parrott, & Eyre, 2002), more recent work demonstrates that these concerns are distinct from both shame and guilt

(Gausel & Leach, 2011; Gausel, Leach, Vignoles, & Brown, 2012; Rees, Allpress, & Brown, 2013; Tauber & van Zomeren, 2013). Specifically, shame is an emotion that individuals experience when they believe that the failure they have committed provides evidence of a personal defect (de Hooge et al., 2008; Lewis, 1971; Tracy & Robins, 2006). Guilt is also an emotion that individuals may experience after committing a failure. Different from shame, however, guilt stems from the realization that one has performed an undesirable *behaviour*, rather than from the realization that an aspect of the *self* is undesirable (Cryder, Springer, & Morewedge, 2012; Tangney, 1996; Tracy & Robins, 2006). Whereas both shame and guilt stem from negative appraisals that individuals form of themselves after committing a failure, social image concerns are negative appraisals that individuals form when they focus on how the failure will affect how others view them. Instead of eliciting feelings of shame or guilt, these social image appraisals tend to raise concerns about being rejected or ostracized by others (Gausel & Leach, 2011; Gausel et al., 2012).

Regardless of how researchers conceptualize social image concerns, there is consensus that these appraisals have important psychological implications for individuals. Indeed, maintaining a favourable social image is important for achieving high self-esteem (Gerber & Wheeler, 2009; Leary & Kowalski, 1990), fulfilling a fundamental human need to belong (Baumeister & Leary, 1995), and securing important material benefits (Hardy & Van Vugt, 2006; MacDonald & Leary, 2005). In addition, because individuals with a negative social image may be marginalized or excluded by others (Bashir et al., 2013; Kaiser et al., 2006; Parks & Stone, 2010), individuals often experience damage to their social image to be highly threatening and painful (Maner, DeWall, Baumeister, & Schaller, 2007; Ybarra, Park, Stanik, & Lee, 2012). Individuals are, therefore, highly motivated to engage in strategies that will encourage others to view them in a positive light (Carr & Vignoles, 2011; Lun, Mesquita, & Smith, 2011). Specifically, when individuals believe that their behaviour is visible to others, they are more likely to behave morally, inhibit socially undesirable actions, and put forth effort to display competence (Leary & Kowalski, 1990).

Although researchers have primarily examined these social image concerns at the individual level, the concerns that individuals have about their group's social image also have substantial implications. Specifically, because individuals consider their membership in important social groups to be a component of their self-concept (Tajfel & Turner, 1979), they strive to maintain a

favourable image of their group in the eyes of others (Brown & Čehajić, 2008; Brown et al., 2008; Chekroun & Nugier, 2011). For example, group members are more likely to use self-presentation strategies that will portray their group favourably when outgroup members are present versus absent (Klein & Azzi, 2001). As is the case at the individual level, groups that are viewed positively by others experience both psychological and material benefits. Indeed, because social groups compete for social status and resources (Tajfel & Turner, 1979), groups with a favourable social image can obtain concrete rewards, whereas groups that have a tarnished social image can experience a variety of consequences (Golec de Zavala, 2011; Owuamalam & Zagefka, 2011). Thus, although being reproached by a sanctimonious other may often be unpleasant experiences for group members, believing that one's group has a tarnished social image may typically be more consequential and may, therefore, have a greater impact on individuals.

Because maintaining a favourable image of one's group is important, individuals often derogate ingroup members who jeopardize the group's social image by performing undesirable behaviors publicly (Chekroun & Nugier, 2011; Hornsey et al., 2005). In contrast to many types of group deviants (e.g., group members who smoke in non-smoking zones, behave aggressively, and display incompetence; Chekroun & Nugier, 2011; Marques & Yzerbyt, 1988; Marques, Yzerbyt, & Leyens, 1988), moral advocates who challenge wrongdoing publicly often act with good intentions, display behaviours that can be considered heroic, and strive to help the group improve (Buchanan & Gorbunova, 2012; Hornsey, 2006; Hornsey et al., 2005; O'Keefe, 2014; Schorn, 2009). Thus, it is possible that group members would react positively to a moral advocate who publicly challenges their group's wrongdoing and would embrace the change promoted by the advocate. On the other hand, because these moral advocates pose a threat to the group's social image, members of the group may actively derogate them. In the present research, therefore, I examined whether social image concerns, rather than concerns about moral reproach, elicit negative reactions to public moral advocates.

3 Public Moral Advocacy as a Double-Edged Sword

I also examined whether threatening a group's social image is a double-edged sword for public moral advocates. That is, I assessed whether the same psychological process that elicits derogation of public moral advocates may also enable them to successfully promote moral

change within the group. For example, Edward Snowden has faced considerable backlash from U.S. government officials for publicly exposing controversial surveillance practices implemented by the National Security Agency (NSA). Nonetheless, these same individuals strongly support new legislation to limit the NSA's access to telephone data. They also admit that Snowden's actions triggered this concern for reform (Macdonald, 2014). Public moral advocates may, therefore, successfully boost concern about addressing group wrongdoing despite facing derogation.

Researchers have previously argued that if individuals derogate moral advocates, they will likely fail to adopt the moral behaviours emulated or promoted by these advocates (Monin et al., 2008). Indeed, individuals generally exhibit only other negative responses to moral advocates whom they dislike (e.g., disagreeing with the opinions of moral advocates; Hornsey et al., 2002). Furthermore, when individuals believe that their social image has been threatened, they are more likely to exhibit withdrawal or avoidance behaviours than to make active attempts to compensate for failures or misdeeds (Iyer, Schmader, & Lickel, 2007; Rees et al., 2013; Tangney, Miller, Flicker, & Barlow, 1996). Specifically, if individuals feel that there is little that they can do to salvage their social image, they may withdraw from the situation in which their image was damaged to minimize feelings of rejection (Gausel & Leach, 2011; Gausel et al., 2012). Thus, the majority of evidence suggests that if moral advocates are derogated because they appear to threaten a group's image, they will also be ineffective in stimulating group change.

Recent research suggests, however, that heightened social image concerns may also promote behaviours aimed at repairing a damaged social image. Indeed, when reflecting on historical transgressions that have been committed by their group (e.g., the Holocaust, the enslavement of African Americans, etc.), individuals become concerned about the implications of these transgressions for their group's image and therefore endorse behaviour change aimed at making reparations for their group's misdeeds (Brown & Čehajić, 2008; Brown et al., 2008; Schmader & Lickel, 2006). Individuals may find accusations about current or relatively recent group misdeeds to be more threatening than reminders about well-known historical transgressions, and may therefore be particularly likely to disregard information disclosed by many types of public moral advocates. Nevertheless, individuals may recognize that endorsing the change implicitly promoted by the advocate can help to restore their group's social image. Just as publicly expressing regret for wrongdoing may appease observers and improve a group's apparent

integrity in the eyes of others (van Leeuwen, van den Bosch, Castano, & Hopman, 2010; Wohl, Hornsey, & Bennett, 2012), individuals may believe that they can re-establish their group's moral standing by expressing concern about the need for group change. Thus, if moral advocates seem to pose a threat to the social image of a group, they may in fact succeed in raising support for group change, despite eliciting negative evaluations.

Although some evidence suggests that individuals who confront prejudice in private settings can reduce stereotyped responses among bigots despite being derogated (Czopp et al., 2006), these derogation and stereotyping responses seem to be driven by distinct psychological processes. In this case, therefore, it may be possible to improve individuals' evaluations of confronters without reducing the positive influence of confronters on prejudiced responses. In the case of public moral advocates, however, the same psychological process may influence both individuals' derogation of the advocates and their receptiveness to the advocates' implicit calls for change. That is, the threat that public moral advocates pose to a group's social image may play a critical role not only in eliciting negative evaluations of the advocates but also in allowing these advocates to raise concern about group change. Thus, public moral advocates may face a paradox: Any attempts that they make to reduce the threat that they pose to the group's image and therefore the derogation that they elicit may, ironically, also undermine their influence on the group. Thus, if public moral advocates wish to boost concern about group change, they may need to shoulder the burden of backlash.

4 The Present Research

In six studies, I examined the extent to which social image concerns, rather than concerns about moral reproach, influence individuals' reactions to public moral advocates. In Studies 1a and 1b, I examined individuals' reactions to a target who publicly exposed incriminating information about their group's academic integrity (Study 1a) or environmental practices (Study 1b). I also examined whether social image concerns rather than moral reproach concerns accounted for individuals' reactions. In Study 2, I assessed whether individuals' reactions were more pronounced when the advocate posed a greater threat to the group's image (i.e., when the advocate publicly challenged wrongdoing committed by participants' ingroup versus outgroup). In Studies 3 and 4, I ruled out an additional alternative mechanism for the reactions observed in Studies 1 and 2. That is, I verified that participants' reactions were not simply due to concerns

that the advocate had made them feel immoral. In Study 4, I also examined whether participants exhibited these reactions even in response to public moral advocates who were authorized to challenge the group's wrongdoing (e.g., targets who released group information as part of their job responsibilities). In Study 5, I obtained further evidence for the role of social image concerns by examining whether group members reacted similarly to advocates who publicly challenged nonmoral group failures. Lastly, in Study 6, I examined an ironic consequence of the role of social image concerns in influencing individuals' reactions to public moral advocates. Specifically, I assessed whether moral advocates who challenge group wrongdoing privately rather than publicly evade derogation but are also ineffective in raising concern about group change. Across studies, I predicted that targets who exposed group wrongdoing would be derogated but would nevertheless be successful in raising concern about group change, because they posed a threat to the group's social image.

The present research makes several important theoretical contributions. Specifically, these studies are the first to demonstrate that moral advocates who challenge a group's wrongdoing publicly are derogated by group members, not because they seem to reproach group members (Minson & Monin, 2012; Monin et al., 2008), but because they seem to threaten the group's social image. In addition, the present work provides evidence that the same psychological process that elicits derogation of public moral advocates may nevertheless allow these advocates to raise concern about change within the perpetrating group. That is, by threatening the group's social image, public moral advocates may raise concern about group change despite facing backlash. Thus, these findings indicate that social image concerns can simultaneously elicit positive and negative reactions from group members, and are therefore more nuanced than past research suggests (Brown et al., 2008; Chekroun & Nugier, 2011; Hornsey et al., 2005).

Chapter 2

Study 1

In Study 1, I examined how group members react when a moral advocate publicly challenges wrongdoing committed by their group. Specifically, I examined whether social image concerns rather than moral reproach concerns lead group members to derogate the advocate but also make them receptive to the advocate's implicit calls for group change. I assessed this possibility by examining individuals' reactions to a target who exposed either incriminating or neutral information about their group's academic integrity (Study 1a) or environmental behaviour (Study 1b). I chose these domains to demonstrate that group members react similarly regardless of whether the wrongdoing is relatively high (i.e., academic) versus low (i.e., environmental) in relevance to the group's identity (i.e., university students). These domains also allowed me to provide evidence that group members respond similarly regardless of whether the wrongdoing is a clear violation of an institutional policy (e.g., cheating on an exam) versus a practice that is socially undesirable but not formally prohibited (e.g., wasting energy). By examining reactions to the release of incriminating and nonincriminating group information in these domains, I aimed to show that individuals' reactions to public moral advocates stem from concerns about the group's social image rather than simply from a disliking for individuals who violate social norms by releasing confidential information. I predicted that the target who exposed incriminating rather than neutral group information would be more likely to make participants concerned about their group's social image rather than make them concerned about being reproached by him. These social image concerns, I expected, would ultimately lead participants to derogate the target but to also be receptive to the group change advocated by the target.

I also examined how participants' reactions would differ depending on the group membership of the target. Research based on social identity theory indicates that ingroup targets who perform undesirable behaviours are more threatening to group members than are outgroup members who do the same (Chekroun & Nugier, 2011; Marques et al., 1988). Individuals may, therefore, be especially likely to derogate a moral advocate and express concern about group improvement when the advocate is an ingroup versus outgroup member. Research on group criticism demonstrates, however, that individuals are more receptive to ingroup versus outgroup critics of their group (Hornsey & Imani, 2004; Sutton, Elder, & Douglas, 2006). Thus, alternatively,

individuals may react more favourably to ingroup versus outgroup public moral advocates. In Studies 1a and 1b, I tested these possibilities by varying the target's group membership.

1 Method

1.1 Participants

Participants in Study 1a were 83 undergraduate students who received course credit. One participant whose score on the group improvement measure was more than three standard deviations beyond the mean was excluded. Accordingly, 26 male and 56 female participants ($M_{age} = 18.67$ years, $SD = 1.43$) were included in analyses.¹

Participants in Study 1b were 184 undergraduate students who received course credit. Five participants scored more than three standard deviations beyond the mean on key measures (i.e., the dependent measures or manipulation checks) and were therefore excluded from the analyses. Accordingly, 76 male participants, 102 female participants, and one participant who did not report his/her biological sex ($M_{age} = 19.01$ years, $SD = 3.19$) were included in analyses.

1.2 Procedure

Participants in Study 1a first read an article ostensibly published in a national newspaper. The article indicated that a student named Andrew had publicly leaked a report about the academic integrity of students at participants' university. Participants were randomly assigned to read a version of the article indicating that the leaked report contained either incriminating information (e.g., "A confidential report recently leaked to CBC reveals that University of Toronto students are particularly likely to engage in academic misconduct [i.e., plagiarism and cheating]") or neutral, nonincriminating information (e.g., "A confidential report recently leaked to CBC reveals that University of Toronto students are no more likely to engage in academic misconduct (i.e., plagiarism and cheating) than are students at other Ontario universities"). Participants were also randomly assigned to read a version of the article indicating that Andrew was a student at participants' own or another university. That is, participants were assigned to one of four conditions in a 2(target disclosure: incriminating, neutral) \times 2(target group membership: ingroup, outgroup) design.

Participants then rated that the extent to which the target's disclosure would threaten their group's social image (e.g., "I feel that the report about UofT students will have a negative impact on the reputation of UofT students;" modified from Brown et al., 2008; six items; Cronbach's $\alpha = .97$) along a 7-point scale anchored at 1(*Strongly disagree*) and 7(*Strongly agree*). In addition, participants rated the degree to which they believed that the target would reproach them by viewing them as immoral (i.e., "If Andrew Baker knew me, he would think that I am..." [reverse-scored]; modified from Minson & Monin, 2012), along a 7-point scale anchored at 1(*Not at all moral*) and 7(*Very moral*). Next, participants rated their evaluations of the target (e.g., "obnoxious," "annoying," [higher values reflect more negative perceptions]; modified from Hornsey & Imani, 2004; nine items; $\alpha = .84$) along a 7-point scale anchored at 1(*Not at all characteristic of Andrew*) and 7(*Very characteristic of Andrew*). They also rated their concerns about the need for group members to improve their behaviour (e.g., "UofT students should make more of an effort to be ethical in the classroom;" modified from Hornsey & Imani, 2004; five items; $\alpha = .89$) along 7-point scales anchored at 1(*Strongly disagree*) and 7(*Strongly agree*). To ensure that the target disclosure manipulation was successful, participants rated the type of information disclosed by the target ("The academic integrity report indicates that UofT students are particularly likely to cheat and plagiarize;" two items; $r = .67$) along a 7-point scale anchored at 1(*Strongly disagree*) and 7(*Strongly agree*). To verify the success of the target group membership manipulation, participants were asked to recall the university that Andrew attended.

Participants in Study 1b completed the same procedure used in Study 1a but were randomly assigned to learn that Andrew had leaked either incriminating or neutral information about their group's environmental practices. They then completed the same social image concerns ($\alpha = .94$), moral reproach, target evaluations ($\alpha = .81$), group improvement concerns ($\alpha = .87$; items were modified to fit the environmental domain), target disclosure ($r = .59$), and target group membership measures used in Study 1a.^{2,3}

2 Results

Because Study 1b was essentially a replication of Study 1a in a different domain, I chose to report analyses based on a combined set of data from the two studies. I find the same pattern of results when analyzing data for each study separately.

2.1 Manipulation checks

2.1.1 Target disclosure

One participant did not respond to these items. Thus, data from 260 participants are included in the analysis. A 2(target disclosure: incriminating, neutral) \times 2(target group membership: ingroup, outgroup) \times 2(moral domain: academic, environmental) analysis of variance (ANOVA) revealed that participants viewed the disclosed information as more incriminating when the target exposed incriminating ($M = 5.48$, $SD = .99$) versus neutral ($M = 3.46$, $SD = 1.15$) group information, $F(1,252) = 206.72$, $p < .001$, $\eta_p^2 = .45$. Although there was also a marginal disclosure \times group membership \times domain interaction, $F(1,252) = 3.00$, $p = .09$, $\eta_p^2 = .01$, simple effects showed that the disclosure effect was significant within all four membership \times domain cells, $ts > 5.66$, $ps < .001$, $ds > .71$. All other effects were nonsignificant, $F_s < 1.09$, $ps > .29$, η_p^2 s $< .005$.

2.1.2 Target group membership

To verify participants' awareness of the target's group membership, I conducted separate chi-square tests for participants who had read about an ingroup versus outgroup target.⁴ Because eight participants did not provide a response to this item, data from 253 participants are included in the analysis. Participants who had read about an ingroup target were more likely to recall that the target was an ingroup (96.80%) versus outgroup (3.20%) member, $\chi^2(1, N = 125) = 109.51$, $p < .001$. Similarly, participants who had read about an outgroup target were more likely to recall that the target was an outgroup (97.66%) versus ingroup (2.34%) member, $\chi^2(1, N = 128) = 116.28$, $p < .001$.⁵

2.2 Potential mediating variables

2.2.1 Social image concerns

A 2 \times 2 \times 2 ANOVA revealed the predicted main effect of target disclosure: Participants were more concerned about their group's social image when the target exposed incriminating ($M = 5.10$, $SD = 1.30$) versus neutral ($M = 3.44$, $SD = 1.45$) group information, $F(1,253) = 98.15$, $p < .001$, $\eta_p^2 = .28$. There was also a significant main effect of domain. Participants were more concerned about their group's image when the target disclosed academic ($M = 4.80$, $SD = 1.72$) versus environmental ($M = 4.00$, $SD = 1.50$) group information, $F(1,253) = 18.99$, $p < .001$, $\eta_p^2 = .07$. There was also a marginal disclosure \times domain interaction, $F(1,253) = 2.76$, $p = .10$, $\eta_p^2 =$

.01. Simple effects showed that, in both domains, participants were more concerned about their group's social image when the target exposed incriminating ($M_{\text{Academic}} = 5.83$, $SD = 1.09$; $M_{\text{Environmental}} = 4.76$, $SD = 1.25$) versus neutral ($M_{\text{Academic}} = 3.77$, $SD = 1.61$; $M_{\text{Environmental}} = 3.29$, $SD = 1.36$) group information, $ts > 6.98$, $ps < .001$, $ds > .87$. All other effects were nonsignificant, $F_s < 1.50$, $ps > .22$, $\eta_p^2s < .007$.

2.2.2 Moral reproach

A $2 \times 2 \times 2$ ANOVA revealed a significant main effect of domain. Participants anticipated greater reproach from the target when he disclosed environmental ($M = 3.99$, $SD = 1.38$) versus academic ($M = 3.02$, $SD = 1.33$) group information, $F(1,253) = 20.53$, $p < .001$, $\eta_p^2 = .08$. All other effects were nonsignificant, $F_s < 2.15$, $ps > .14$, $\eta_p^2s < .009$.

2.3 Dependent variables

2.3.1 Target evaluations

A $2 \times 2 \times 2$ ANOVA revealed the predicted main effect of target disclosure: Participants rated the target more negatively when he exposed incriminating ($M = 4.15$, $SD = .89$) versus neutral ($M = 3.78$, $SD = .78$) group information, $F(1,253) = 12.86$, $p < .001$, $\eta_p^2 = .05$. There was also a main effect of domain. Participants rated the target more negatively when he disclosed academic ($M = 4.21$, $SD = .94$) versus environmental ($M = 3.85$, $SD = .79$) group information, $F(1,253) = 11.08$, $p = .001$, $\eta_p^2 = .04$. All other effects were nonsignificant, $F_s < 1.81$, $ps > .17$, $\eta_p^2s < .008$.

2.3.2 Group improvement concerns

A $2 \times 2 \times 2$ ANOVA revealed the predicted main effect of target disclosure: Participants expressed greater concern about the need for group change when the target exposed incriminating ($M = 5.49$, $SD = 1.00$) versus neutral ($M = 4.96$, $SD = .97$) information, $F(1,253) = 16.65$, $p < .001$, $\eta_p^2 = .06$. There was also a significant main effect of domain. Participants expressed more concern about group improvement when the target disclosed academic ($M = 5.42$, $SD = 1.04$) versus environmental ($M = 5.13$, $SD = 1.00$) group information, $F(1,253) = 4.42$, $p = .04$, $\eta_p^2 = .02$. All other effects were nonsignificant, $F_s < 1.13$, $ps > .29$, $\eta_p^2s < .005$.

2.4 Mediation Analyses

I argue that social image concerns rather than moral reproach concerns influence individuals' reactions to moral advocates who challenge wrongdoing publicly. To examine this possibility, I tested two multiple mediator models based on 5000 bootstrapped resamples (Preacher & Hayes, 2008) to assess whether social image concerns and moral reproach mediated the effects of target disclosure on both target evaluations and group improvement concerns. In these models, either target evaluations or group improvement concerns was the dependent variable, target disclosure was the independent variable, and both social image concerns and moral reproach were parallel mediators. Target group membership, moral domain, and all of the interactions were covariates (see Figure 1).

The 95% bias-corrected confidence interval of the indirect effect of target disclosure on evaluations through social image concerns did not contain zero ($ab = .16$, $SE = .04$, $CI = [.09, .24]$). The same was true of the indirect effect of target disclosure on group improvement concerns through social image concerns ($ab = .16$, $SE = .04$, 95% $CI = [.08, .25]$). The indirect effects of disclosure on evaluations ($ab = .003$, $SE = .006$, 95% $CI = [-.007, .02]$) and group improvement ($ab = -.01$, $SE = .01$, 95% $CI = [-.04, .005]$) through moral reproach were not significant. The direct effects of disclosure on evaluations ($c' = .03$, $SE = .06$, 95% $CI = [-.09, .15]$) and group improvement ($c' = .12$, $SE = .07$, 95% $CI = [-.03, .26]$) were nonsignificant when the mediating variables were included in the models. Thus, participants who learned that the target had publicly exposed incriminating versus neutral group information were more likely to both derogate the target and express greater concern about group improvement because they believed that the target's disclosure threatened their group's social image, rather than because they believed that the target would reproach them.

3 Discussion

Studies 1a and 1b provide evidence that social image concerns rather than moral reproach concerns influence group members' reactions to moral advocates who challenge group wrongdoing publicly. Indeed, participants who learned that the target had publicly exposed their group's wrongdoing believed that the disclosure would have a more negative impact on their group's image. This ultimately elicited disapproval of the target. Concerns about moral reproach by the target, in comparison, did not account for the observed effect.

By threatening a group's social image, however, public moral advocates may also succeed in stimulating change within the group. Indeed, past research suggests that moral advocates who are disliked may be ineffective in promoting change among others (Hornsey et al., 2002; Monin et al., 2008). The present study demonstrates, in comparison, that when moral advocates challenge wrongdoing publicly and therefore threaten a group's social image, these actions may in fact allow to them successfully raise concern about group change.

Comparing group members' reactions to targets who exposed academic versus environmental information was not a central focus of this research. Nonetheless, participants in the present studies were more likely to express concern about their group's image, derogate the target, and indicate concern about group change when the target exposed academic versus environmental information. Because adhering to standards of academic integrity is especially important for students and central to society's expectations of students, participants in the present studies may have been especially sensitive to disclosures about their group's academic integrity record. Participants may have viewed the academic information as more threatening to the group's social image and, therefore, have reacted more strongly to the targets who exposed academic versus environmental information. Alternatively, participants may have been more sensitive to disclosures about the group's academic integrity because academic misconduct is arguably more unambiguously unethical than are environmentally harmful practices. In any case, the absence of target disclosure \times moral domain interactions provides evidence that participants reacted similarly to the moral advocate in both domains.

Although participants reacted differently to the target who exposed wrongdoing rather than neutral group information, the target's group membership did not moderate these effects. These findings are consistent with research indicating that group members react similarly to ingroup and outgroup targets who provide negative feedback about the group publicly (Hornsey et al., 2005). In the present study, social image concerns mediated participants' reactions to the targets. Thus, aspects of the target's disclosure that were expected to affect the public's image of the group (e.g., *what* was disclosed) may have had a greater influence on participants' reactions than did aspects of the disclosure that were expected to have only a limited impact on the group's social image (e.g., *who* disclosed it).

Chapter 3

Study 2

Studies 1a and 1b provide evidence from two domains that social image concerns underlie group members' reactions to moral advocates who challenge the group's wrongdoing publicly. To the extent that this is true, group members should not react equally strongly to public moral advocates who target any group. Instead, they should be particularly likely to derogate a public moral advocate and express concern about group improvement when the advocate targets their ingroup. Indeed, individuals generally perceive an event or incident that reflects poorly on their ingroup rather than on an outgroup to pose a greater threat to their group's social image (Chekroun & Nugier, 2011). Thus, public moral advocates who target individuals' ingroup versus outgroup should pose a greater threat to the ingroup's social image and, ultimately, elicit stronger reactions from individuals. In Study 2, I examined this directly.

1 Method

1.1 Participants

Participants were 83 undergraduate students who received course credit or \$10. One participant whose score on the group improvement concern measure was more than three standard deviations below the mean was excluded. Accordingly, 36 male and 46 female participants ($M_{age} = 19.65$ years, $SD = 1.86$) were included in analyses.

1.2 Procedure

Participants completed a procedure similar to that used in Study 1a with the exception that the article indicated that the target had publicly exposed academic misconduct committed by participants' ingroup or an outgroup. In addition, the article identified the target as either an ingroup or outgroup member. In all conditions, the target disclosed incriminating group information. In sum, participants were randomly assigned to one of four conditions in a 2(wrongdoer group membership: ingroup, outgroup) \times 2(target group membership: ingroup, outgroup) design. Participants then completed the same social image concerns ($\alpha = .96$), moral reproach, target evaluations ($\alpha = .84$), group improvement concern ($\alpha = .84$), and target group membership manipulation check measures used in Studies 1a and 1b. To verify the success of the wrongdoer group membership manipulation, participants also rated the extent to which the

article indicated that their ingroup was involved in academic misconduct along a 7-point scale anchored at 1(*Strongly disagree*) and 7(*Strongly agree*).

2 Results

2.1 Manipulation checks

2.1.1 Wrongdoer group membership

A 2(wrongdoer group membership) \times 2(target group membership) ANOVA revealed a main effect of wrongdoer group membership: Participants were more likely to indicate that the target had exposed wrongdoing committed by their ingroup when the targets exposed ingroup ($M = 5.61$, $SD = 1.05$) versus outgroup ($M = 2.50$, $SD = 1.55$) misconduct, $F(1,78) = 106.86$, $p < .001$, $\eta_p^2 = .58$. The main effect of target group membership and the interaction were nonsignificant, $F_s < .25$, $p_s > .62$, η_p^2 s $< .004$.

2.1.2 Target group membership

As in Study 1, I conducted separate chi-square tests for participants who read about an ingroup versus outgroup target. Because eleven participants did not provide a response to this item, data from 71 participants are included in the analysis. Participants who had read about an ingroup target were more likely to recall that the target was an ingroup (89.19%) versus outgroup (10.81%) member, $\chi^2(1, N = 37) = 22.73$, $p < .001$. Participants who had read about an outgroup target were more likely to recall that the target was an outgroup (91.18%) versus ingroup (8.82%) member, $\chi^2(1, N = 34) = 23.06$, $p < .001$.⁶

2.2 Potential mediating variables

2.2.1 Social image concerns

A 2 \times 2 ANOVA revealed the predicted main effect of wrongdoer group membership: Participants were more concerned about their group's social image when the target exposed incriminating information about their own ($M = 5.51$, $SD = 1.04$) versus another ($M = 3.34$, $SD = 1.39$) group, $F(1,78) = 60.97$, $p < .001$, $\eta_p^2 = .44$. The main effect of target group membership and the interaction were nonsignificant, $F_s < .60$, $p_s > .44$, η_p^2 s $< .009$.

2.2.2 Moral reproach

A 2×2 ANOVA revealed neither a main effect of wrongdoer group membership nor a main effect of target group membership, $F_s < 1.36$, $p_s > .24$, η_p^2 s $< .02$. Although there was a significant wrongdoer membership \times target membership interaction, $F(1,78) = 3.99$, $p = .05$, $\eta_p^2 = .05$, simple effects tests showed that the wrongdoer group membership effect was nonsignificant within both target group membership conditions, $t_s < 1.61$, $p_s > .22$, $d_s < .36$.

2.3 Dependent variables

2.3.1 Target evaluations

A 2×2 ANOVA revealed the predicted main effect of wrongdoer group membership: Participants rated the target more negatively when he disclosed wrongdoing committed by participants' ingroup ($M = 4.45$, $SD = .78$) versus outgroup ($M = 4.04$, $SD = .88$), $F(1,78) = 5.30$, $p = .02$, $\eta_p^2 = .06$. The main effect of target group membership and the interaction were nonsignificant, $F_s < 1.89$, $p_s > .17$, η_p^2 s $< .03$.

2.3.2 Group improvement concerns

A 2×2 ANOVA revealed a marginal main effect of wrongdoer group membership: Participants expressed greater concern about group improvement when the target exposed wrongdoing committed by participants' ingroup ($M = 5.24$, $SD = 1.04$) versus outgroup ($M = 4.87$, $SD = .76$), $F(1,78) = 3.14$, $p = .08$, $\eta_p^2 = .04$. The main effect of target group membership and the interaction were nonsignificant, $F_s < 2.42$, $p_s > .12$, η_p^2 s $< .04$.

2.4 Mediation analyses

Consistent with Study 1, the 95% bias-corrected confidence interval of the indirect effect of wrongdoer group membership on target evaluations through social image concerns did not contain zero ($ab = .18$, $SE = .09$, $CI = [.009, .37]$). The same was true for the indirect effect of wrongdoer group membership on group improvement concerns through social image concerns ($ab = .16$, $SE = .09$, 95% $CI = [.003, .36]$). The indirect effects of wrongdoer group membership on evaluations ($ab = .002$, $SE = .02$, 95% $CI = [-.03, .06]$) and group improvement concerns ($ab = .003$, $SE = .01$, 95% $CI = [-.02, .06]$) through moral reproach were nonsignificant. The direct effects of wrongdoer group membership on evaluations ($c' = .03$, $SE = .12$, 95% $CI = [-.21, .26]$)

and group improvement ($c' = .01$, $SE = .13$, 95% CI = [-.25, .28]) were nonsignificant when the mediating variables were included in the models. Thus, participants who learned that a target had publicly exposed wrongdoing committed by their own versus another group rated the target more negatively and expressed greater concerns about group improvement because they were more concerned about their group's social image.

3 Discussion

These findings provide further support for the argument that concerns about maintaining a positive social image rather than concerns about moral reproach influence group members' reactions to moral advocates who challenge group wrongdoing publicly. Indeed, participants exhibited stronger reactions to the advocate when he challenged wrongdoing committed by their ingroup versus outgroup, an action that should seem to pose a greater threat to the group's social image. In addition, consistent with Study 1, social image concerns rather than moral reproach concerns mediated these effects.⁷

Chapter 4

Study 3

Studies 1 and 2 provide evidence that individuals' reactions to public moral advocates are driven by social image concerns rather than moral reproach concerns. It remains possible, however, that participants' reactions were actually due to concerns about their moral integrity; the advocate may have simply made participants feel bad about their morality. Indeed, when individuals are confronted for making a prejudiced remark, they subsequently feel negatively about themselves (Czopp et al., 2006). Furthermore, when individuals realize that their behaviour does not meet a particular moral standard, they may become concerned about their morality (Doosje, Branscombe, Spears, & Manstead, 1998; Monin, 2007; Peetz, Gunn, & Wilson, 2010). These moral integrity concerns may be sufficiently threatening that they elicit backlash against a public moral advocate who makes individuals aware that they may not be meeting acceptable standards for moral behaviour. In addition, to improve their moral standing, individuals may simultaneously become concerned about the need for their group to improve morally.

I argue, however, that a damaged social image may typically be more costly for individuals than is a damaged image of oneself or one's group. Indeed, individuals' social standing and the social and material benefits derived from their social standing (e.g., prestige) are determined more by the characteristics that individuals appear to possess in the eyes of others (e.g., others' impressions of their morality) than by the characteristics that they personally believe they possess (e.g., their own impressions of their morality; Berger, Rosenholtz, & Zelditch, 1980; Hardy & Van Vugt, 2006). Thus, social image concerns may play a greater role in influencing individuals' reactions to public moral advocates than do moral integrity concerns. In Study 3, I tested this possibility by assessing whether social image concerns, moral integrity concerns, and moral reproach explain individuals' reactions to moral advocates who challenge wrongdoing publicly.

1 Method

1.1 Participants

Participants were 55 undergraduate students who received course credit. One participant who reported that the newspaper article was fake and one participant who reported difficulties

understanding the task instructions were excluded. Accordingly, 25 male participants, 26 female participants, and two participants who did not report their biological sex ($M_{age} = 18.78$ years, $SD = 1.43$) were included in analyses.

1.2 Procedure

Participants completed the same procedure as in Study 1a, except that they were randomly assigned to read about an ingroup target who disclosed either incriminating or neutral group information. Because the target's group membership did not moderate the effects observed in Studies 1 and 2, I excluded this manipulation from Study 3 to simplify the experimental design. In this study and all subsequent studies reported, the target was always an ingroup member.

After completing the same social image concerns measure used in Studies 1 and 2 ($\alpha = .97$), participants completed a measure assessing their concern about their personal moral integrity (e.g., "Reading the academic integrity report makes me realize that I am an unethical person;" three items; $\alpha = .82$). To ensure that moral integrity concerns did not fail to emerge as a significant mediator simply because I assessed this construct at an individual rather than group level, participants also completed a measure assessing their concern about their group's moral integrity (e.g., "As a UofT student, I feel bad that UofT students are not meeting standards for academic integrity;" five items, $\alpha = .92$). Both measures were adapted from previous research (Brown et al., 2008). Participants made their ratings on a 7-point scale with endpoints labeled 1(*Strongly disagree*) and 7(*Strongly agree*). Similarly, in addition to the personal moral reproach measure included in Studies 1 and 2, participants rated the extent to which they anticipated that their group would be reproached by the target ("Andrew Baker thinks that UofT students are..." [reverse-scored]; modified from Minson & Monin, 2012) along 7-point scales anchored at 1(*Not at all moral*) and 7(*Very moral*). Lastly, they completed the same target evaluations ($\alpha = .75$), group improvement concern ($\alpha = .87$), and target disclosure manipulation check ($r = .64$) measures used in the previous studies.

2 Results

2.1 Manipulation check

Two participants did not respond to the manipulation check items. Thus, data from 51 participants are included in this analysis. A one-way ANOVA revealed that participants viewed the incriminating ($M = 5.44$, $SD = .78$) versus neutral ($M = 3.31$, $SD = 1.46$) group information as more incriminating, $F(1,49) = 41.94$, $p < .001$, $\eta_p^2 = .46$.

2.2 Potential mediating variables

2.2.1 Social image concerns

A one-way ANOVA revealed that participants were more concerned about their group's social image when the target exposed incriminating ($M = 5.71$, $SD = 1.22$) versus neutral ($M = 3.83$, $SD = 1.35$) group information, $F(1,51) = 28.01$, $p < .001$, $\eta_p^2 = .36$.

2.2.2 Personal moral reproach

There was no significant effect of disclosure on personal moral reproach, $F(1,51) = 1.13$, $p = .29$, $\eta_p^2 = .02$.

2.2.3 Group moral reproach

Participants believed that the target was more likely to morally reproach their group when he released incriminating ($M = 5.54$, $SD = .91$) versus neutral ($M = 3.78$, $SD = 1.34$) information, $F(1,51) = 31.20$, $p < .001$, $\eta_p^2 = .38$.

2.2.4 Personal moral integrity

There was no significant effect of disclosure on personal moral integrity concerns, $F(1,51) = .08$, $p = .78$, $\eta_p^2 = .002$.

2.2.5 Group moral integrity

There was no significant effect of disclosure on concerns about the group's moral integrity, $F(1,51) = 1.84$, $p = .18$, $\eta_p^2 = .04$.

2.3 Dependent variables

2.3.1 Target evaluations

A one-way ANOVA revealed that participants rated the target more negatively when he exposed incriminating ($M = 4.24$, $SD = .71$) versus neutral ($M = 3.73$, $SD = .69$) group information, $F(1,51) = 6.84$, $p = .01$, $\eta_p^2 = .12$.

2.3.2 Group improvement concerns

Participants also expressed marginally greater concern about group improvement when the target exposed incriminating ($M = 5.61$, $SD = .96$) versus neutral ($M = 5.15$, $SD = 1.06$) information, $F(1,51) = 2.74$, $p = .10$, $\eta_p^2 = .05$.

2.4 Mediation analyses

Using the same procedure as in Studies 1 and 2, I assessed the extent to which social image concerns, moral integrity concerns, and concerns about moral reproach mediated the effects of target disclosure on target evaluations and group improvement concerns. In these multiple mediator models, all five potential mechanisms (i.e., social image concerns, personal moral integrity concerns, group moral integrity concerns, personal moral reproach, and group moral reproach) were entered as parallel mediators.

The 95% bias-corrected confidence interval of the indirect effect of target disclosure on target evaluations through social image concerns did not contain zero ($ab = .14$, $SE = .08$, $CI = [.001, .33]$). The indirect effects of target disclosure on evaluations through personal moral integrity concerns ($ab = .01$, $SE = .03$, 95% $CI = [-.04, .10]$), group moral integrity concerns ($ab = -.02$, $SE = .03$, 95% $CI = [-.13, .02]$), personal moral reproach ($ab = -.02$, $SE = .03$, 95% $CI = [-.13, .01]$) and group moral reproach ($ab = .09$, $SE = .08$, 95% $CI = [-.05, .28]$) were nonsignificant. When the mediating variables were included in the model, the direct effect of disclosure on evaluations was nonsignificant, $c' = .07$, $SE = .13$, 95% $CI = [-.20, .33]$.

Similarly, the indirect effect of target disclosure on group improvement concerns through social image concerns was significant ($ab = .22$, $SE = .12$, 95% $CI = [.06, .53]$) whereas the indirect effects of disclosure on group improvement through personal moral integrity concerns ($ab = .01$, $SE = .05$, 95% $CI = [-.10, .08]$), group moral integrity concerns ($ab = .03$, $SE = .04$, 95% $CI = [-$

.03, .14]), personal moral reproach ($ab = -.03$, $SE = .04$, 95% CI = [-.16, .02]) and group moral reproach ($ab = -.02$, $SE = .11$, 95% CI = [-.25, .19]) were nonsignificant. The direct effect of disclosure on group improvement was nonsignificant when the mediating variables were included in the model, $c' = .02$, $SE = .17$, 95% CI = [-.33, .36]. Thus, social image concerns rather than concerns about moral integrity or moral reproach influenced participants' reactions to the target.⁸

3 Discussion

By ruling out an additional alternative mechanism for the effects observed in Studies 1 and 2, these findings provide further evidence for the role of social image concerns in influencing group members' reactions to public moral advocates. Indeed, the present findings demonstrate that group members' reactions to public moral advocates do not stem from threatening personal or group appraisals in general. Concerns about social image, moral integrity, and moral reproach all represent threatening appraisals that individuals may form when their group fails in some way (Gausel et al., 2012; Lewis, 1971; Minson & Monin, 2012). In this study, however, participants who learned that a target had leaked incriminating information about their group derogated the target and expressed concern about group improvement because they were concerned about their social image. Participants' concerns about either their own or their group's moral integrity as well as their expectations that either they or their group would be reproached by the advocate did not account for these reactions.

Chapter 5

Study 4

In Studies 1-3, the moral advocate who exposed incriminating information not only released unflattering information about the group, but he also did so in an unauthorized manner. It remains unknown, therefore, whether individuals exhibit similar reactions to moral advocates who are authorized to release information that challenges or exposes a group's wrongdoing (e.g., a consultant or company spokesperson who is given the task of publicly releasing information that calls the company's practices into question). That is, group members may have derogated the moral advocate who challenged group wrongdoing in Studies 1-3 simply because he was unauthorized to do so. When individuals are instead authorized to release information, even incriminating information, group members may not derogate them. Alternatively, because receiving permission to disclose incriminating group information does not alter the content of information released publicly, individuals may infer that the public will view their group negatively regardless of whether or not the advocate was permitted to deliver this information. Group members may, therefore react similarly to public moral advocates regardless of whether or not they are authorized to challenge the group's practices.

Studies 1 and 3 suggest that group members may indeed react similarly to public moral advocates who are authorized versus unauthorized to challenge group wrongdoing publicly. Specifically, these studies provide evidence that group members do not derogate moral advocates who leak information about the group's wrongdoing simply because these advocates violate rules. Indeed, participants reacted differently to a target who leaked incriminating versus neutral group information, despite the fact that both targets engaged in rule-breaking behaviour. Furthermore, these studies demonstrate that individuals' reactions to public moral advocates stem from concerns about the impact of the advocates' actions on the group's social image. To the extent that permission to expose group wrongdoing does not alter the threat that public moral advocates pose to group's social image, group members may react similarly to authorized and unauthorized public moral advocates. That is, public moral advocates may elicit derogation but also succeed in raising concern about group improvement, regardless of whether or not they are authorized to challenge the group's behaviour. I examined this in Study 4.

1 Method

1.1 Participants

Participants were 193 undergraduate students who received course credit. Two participants who did not consent to the use of their data in analyses were excluded. Accordingly, 79 male participants, 111 female participants, and one participant who did not report his/her biological sex ($M_{age} = 18.78$ years, $SD = 1.44$) were included in analyses.

1.2 Procedure

Participants were first randomly assigned to read either the incriminating or neutral academic integrity article used in Study 1a. In addition, participants within each information condition were randomly assigned to learn that the target was either authorized to disclose the information that he released (e.g., “As an OFUA [Ontario Federation of University Affairs] student member, Baker was assigned the task of releasing the report from the OFUA’s secure electronic database to the media”) or was not authorized to do so (e.g., “As an OFUA student member, Baker was not permitted to share the content of any OFUA report with other individuals”). That is, participants were randomly assigned to one of four conditions in a 2(target disclosure: incriminating, neutral) \times 2(disclosure authorization: authorized, unauthorized) design.

Participants then completed the same social image concerns ($\alpha = .96$), personal moral reproach, group moral reproach, personal moral integrity ($\alpha = .83$), group moral integrity ($\alpha = .88$), target evaluations ($\alpha = .87$), and group improvement concerns ($\alpha = .91$) measures used in the previous studies. They also completed the target disclosure manipulation check measure used in Study 1a ($r = .51$). In addition, to verify that participants were aware of the target’s authorization status, participants rated the extent to which the target was permitted to disclose the information (e.g., “Andrew was permitted to release the academic integrity report;” two items; $r = .75$) along 7-point scale anchored at 1(*Strongly disagree*) and 7(*Strongly agree*).

2 Results

2.1 Manipulation checks

2.1.1 Target disclosure

Because one participant did not complete the manipulation check measures, data from 190 participants are included in the manipulation check analyses. A 2(target disclosure) \times 2(disclosure authorization) ANOVA revealed that participants viewed the incriminating ($M = 5.43$, $SD = 1.09$) versus neutral ($M = 3.27$, $SD = 1.07$) group information as more incriminating, $F(1,186) = 188.51$, $p < .001$, $\eta_p^2 = .50$. The main effect of disclosure authorization and the target disclosure \times disclosure authorization interaction were nonsignificant, $F_s < .34$, $p_s > .56$, η_p^2 s $< .003$.

2.1.2 Disclosure authorization

A 2×2 ANOVA revealed that participants viewed the target as more authorized to disclose the information when he was ($M = 5.05$, $SD = 1.27$) versus was not ($M = 2.33$, $SD = 1.40$) authorized to do so, $F(1,186) = 194.76$, $p < .001$, $\eta_p^2 = .51$. The main effect of target disclosure and the interaction were nonsignificant, $F_s < .42$, $p_s > .51$, η_p^2 s $< .003$.

2.2 Potential mediating variables

2.2.1 Social image concerns

A 2×2 ANOVA revealed the predicted main effect of target disclosure: Participants were more concerned about their group's social image when the target exposed incriminating ($M = 5.54$, $SD = 1.38$) versus neutral ($M = 3.30$, $SD = 1.27$) group information, $F(1,187) = 134.71$, $p < .001$, $\eta_p^2 = .42$. The main effect of authorization and the interaction were nonsignificant, $F_s < .99$, $p_s > .32$, η_p^2 s $< .006$.

2.2.2 Personal moral reproach

A 2×2 ANOVA revealed a marginal main effect of disclosure: Participants were more likely to believe that the target would personally reproach them when he disclosed incriminating ($M = 3.56$, $SD = 1.40$) versus neutral ($M = 3.22$, $SD = 1.42$) information, $F(1,187) = 2.77$, $p = .10$, η_p^2

= .02. The main effect of authorization and the interaction were nonsignificant, $F_s < .09$, $p_s > .77$, η_p^2 s < .001.

2.2.3 Group moral reproach

A 2×2 ANOVA revealed a significant main effect of disclosure: Participants believed that the target was more likely to morally reproach their group when he released incriminating ($M = 5.67$, $SD = 1.08$) versus neutral ($M = 3.49$, $SD = 1.29$) information, $F(1,187) = 162.30$, $p < .001$, $\eta_p^2 = .47$. There was also a significant target disclosure \times authorization interaction, $F(1,187) = 5.80$, $p = .02$, $\eta_p^2 = .03$. Simple effects tests showed that, within both authorization conditions, participants anticipated greater group-directed moral reproach from the target who released incriminating ($M_{Authorized} = 5.47$, $SD = 1.16$; $M_{Unauthorized} = 5.86$, $SD = .98$) versus neutral ($M_{Authorized} = 3.71$, $SD = 1.24$; $M_{Unauthorized} = 3.28$, $SD = 1.31$) information, $t_s > 7.27$, $p_s < .001$, $d_s > 1.06$. The main effect of authorization was nonsignificant, $F(1,187) = .02$, $p = .90$, $\eta_p^2 < .001$.

2.2.4 Personal moral integrity

There were no significant effects on personal moral integrity concerns, $F_s < .97$, $p_s > .32$, η_p^2 s < .006.

2.2.5 Group moral integrity

There was a main effect of target disclosure on concerns about the group's moral integrity: Participants were more concerned about the group's moral integrity when the target exposed incriminating ($M = 4.83$, $SD = 1.57$) versus neutral ($M = 4.20$, $SD = 1.72$) information, $F(1,187) = 6.96$, $p = .009$, $\eta_p^2 = .04$. There was also a marginal main effect of authorization. Participants were more concerned about their group's moral integrity when the target was unauthorized ($M = 4.74$, $SD = 1.72$) versus authorized ($M = 4.30$, $SD = 1.61$) to expose the information, $F(1,187) = 3.29$, $p = .07$, $\eta_p^2 = .02$. The interaction was nonsignificant, $F(1,187) = .99$, $p = .32$, $\eta_p^2 = .005$.

2.3 Dependent variables

2.3.1 Target evaluations

A 2×2 ANOVA revealed the predicted main effect of target disclosure: Participants rated the target more negatively when he exposed incriminating ($M = 4.11$, $SD = .85$) versus neutral ($M = 3.49$, $SD = .90$) group information, $F(1,187) = 23.93$, $p < .001$, $\eta_p^2 = .11$. There was also a main

effect of target authorization: Participants rated the target more negatively when he was unauthorized ($M = 3.96$, $SD = .89$) versus authorized ($M = 3.64$, $SD = .94$) to disclose the information, $F(1,187) = 6.57$, $p = .01$, $\eta_p^2 = .03$. The disclosure \times authorization interaction was nonsignificant, $F(1,187) = .61$, $p = .44$, $\eta_p^2 = .003$.

2.3.2 Group improvement concerns

A 2×2 ANOVA on participants' group improvement concerns revealed the predicted main effect of target disclosure: Participants expressed greater concern about group improvement when the target exposed incriminating ($M = 5.54$, $SD = 1.05$) versus neutral ($M = 5.05$, $SD = 1.19$) group information, $F(1,187) = 9.01$, $p = .003$, $\eta_p^2 = .05$. The main effect of authorization and the interaction were nonsignificant, $F_s < .55$, $p_s > .46$, η_p^2 s $< .004$.

2.4 Mediation analyses

As in Study 3, social image concerns mediated the effect of target disclosure on evaluations ($ab = .28$, $SE = .06$, 95% CI = [.17, .42]). In this case, group moral reproach concerns also mediated the target disclosure effect on evaluations ($ab = .10$, $SE = .05$, 95% CI = [.003, .21]), whereas personal moral integrity ($ab = .003$, $SE = .008$, 95% CI = [-.005, .03]), group moral integrity ($ab = -.01$, $SE = .01$, 95% CI = [-.05, .01]), and personal moral reproach ($ab = .009$, $SE = .01$, 95% CI = [-.005, .04]) did not. Similarly, social image concerns mediated the effect of target disclosure on improvement concerns ($ab = .16$, $SE = .08$, 95% CI = [.02, .32]). Group moral integrity concerns also mediated the target disclosure effect on improvement concerns ($ab = .05$, $SE = .03$, 95% CI = [.01, .12]), whereas personal moral integrity ($ab = .01$, $SE = .02$, 95% CI = [-.02, .06]), personal moral reproach ($ab = -.02$, $SE = .02$, 95% CI = [-.08, .001]), and group moral reproach ($ab = .12$, $SE = .08$, 95% CI = [-.04, .27]) did not. The direct effects of disclosure on evaluations ($c' = -.08$, $SE = .09$, 95% CI = [-.25, .09]) and group improvement ($c' = -.09$, $SE = .11$, 95% CI = [-.31, .14]) were nonsignificant when the mediating variables were included in the models.

3 Discussion

This study demonstrates that moral advocates who publicly challenge group wrongdoing are derogated, even when they are authorized to challenge the group's practices. Indeed, participants expressed greater concern about their group's moral integrity and evaluated the target more

negatively when he was unauthorized versus authorized to release group information. Participants may have believed that their group's wrongdoing was more serious if the target was willing to violate rules to expose it. Nonetheless, the target's authorization to challenge group wrongdoing did not moderate the effect of target disclosure on individuals' reactions. These findings provide further evidence that individuals do not derogate only those public moral advocates who choose to break rules in order to take a stand against a group's wrongdoing. Given that maintaining a favourable social image is critical to fulfilling basic human needs (Baumeister & Leary, 1995; Gerber & Wheeler, 2009; Leary & Kowalski, 1990; MacDonald & Leary, 2005), individuals may be most sensitive to characteristics and actions of public moral advocates that will affect how their group is viewed by outsiders. Aspects of public moral advocates that are unlikely to substantially affect the group's social image (e.g., group membership and authorization status) may have only a limited impact on individuals' reactions to these advocates. Thus, although it may seem that individuals should derogate only those moral advocates who boldly defy authority and use radical methods to expose wrongdoing publicly, the present findings suggest that individuals react negatively even to those who become public moral advocates simply by virtue of their responsibility to expose group wrongdoing. In essence, group members seem to adopt a "shoot the messenger" approach, even when the target has full authorization to deliver the message.

The present findings also provide converging evidence that social image concerns, rather than simply concerns about moral reproach or moral integrity, influence individuals' reactions to public moral advocates. As in Study 3, even when I considered the possibility that moral reproach or moral integrity may account for individuals' reactions, social image concerns explained why the target who disclosed incriminating group information was both derogated and successful in raising concern about group improvement. In fact, social image concerns was the only variable that emerged as a significant mediator in both mediation analyses in Study 4. Although group moral reproach significantly mediated the disclosure effect on evaluations, it did not mediate the disclosure effect on group improvement concerns. Similarly, group moral integrity mediated the disclosure effect on group improvement, but did not mediate the disclosure effect on evaluations. Furthermore, whereas social image concerns also mediated the disclosure effects on evaluations and improvement concerns in Study 3, neither group moral reproach nor group moral integrity emerged as significant mediators in Study 3. Thus, although

concerns about moral reproach and moral integrity may play a key role in influencing individuals' reactions in some moral contexts (Czopp et al., 2006; Monin et al., 2008), social image concerns seem to be more important to individuals when a moral advocate challenges their group's wrongdoing publicly.

Chapter 6

Study 5

Studies 1-4 demonstrate that group members' reactions to moral advocates who publicly expose group wrongdoing are driven by social image concerns rather than concerns about moral reproach. To the extent that this is true, group members should exhibit similar reactions to targets who expose negative group information that is unrelated to morality (e.g., an individual who publicly exposes a group's incompetence). Indeed, individuals exhibit concern about their group's social image in response to both moral and nonmoral group failures (Gausel & Leach, 2011; Smith et al., 2002). Thus, regardless of whether advocates expose information about a group's moral versus nonmoral shortcomings, they may be derogated yet succeed in raising concern about the need for group improvement. On the other hand, if group members' reactions to public moral advocates are instead influenced by moral reproach concerns, individuals should react differently to targets who expose moral versus nonmoral group failures (Minson & Monin, 2012; Monin et al., 2008). In Study 5, I examined these possibilities directly.

1 Method

1.1 Participants

Participants were 200 undergraduate students who received course credit. Four participants whose session was interrupted by a fire drill, one participant whose computer malfunctioned during the session, and one participant who had already taken part in a study on the same topic were excluded. Accordingly, 101 male and 93 female participants ($M_{age} = 19.19$ years, $SD = 1.73$) were included in analyses.

1.2 Procedure

Participants were first randomly assigned to learn that a target had disclosed either incriminating or neutral information about their group's morality or competence. Participants in the morality conditions read either the incriminating or neutral misconduct article used in Study 1a.

Participants in the competence conditions read either an incriminating or neutral article about the group's competence (e.g., "A confidential report recently leaked to CBC reveals that University of Toronto students are [not particularly socially skilled/as socially skilled as are students at other Ontario universities]"). That is, participants were randomly assigned to one of four

conditions in a 2(target disclosure: incriminating, neutral) \times 2(group failure domain: morality, competence) design. Participants then completed the same social image concerns ($\alpha = .94$), personal moral reproach, group moral reproach, target evaluations ($\alpha = .83$), and group improvement concerns ($\alpha = .84$; items were modified in the competence domain) measures used in Study 4. They also completed the target disclosure manipulation check measure used in the previous studies. The items in this measure were the same as those used in the previous study (i.e., in the morality conditions) or modified to fit the competence domain (i.e., in the competence conditions; $r = .48$).

2 Results

2.1 Manipulation check

2.1.1 Target disclosure

A 2(target disclosure) \times 2(group failure domain) ANOVA revealed that participants viewed the incriminating ($M = 5.25$, $SD = 1.25$) versus neutral ($M = 3.23$, $SD = 1.15$) group information as more incriminating, $F(1,190) = 136.83$, $p < .001$, $\eta_p^2 = .42$. The main effect of failure domain and the disclosure \times failure domain interaction were nonsignificant, $F_s < 1.15$, $ps > .28$, η_p^2 s $< .007$.⁹

2.2 Potential mediating variables

2.2.1 Social image concerns

A 2 \times 2 ANOVA revealed the predicted main effect of target disclosure: Participants were more concerned about their group's social image when the target exposed incriminating ($M = 5.17$, $SD = 1.22$) versus neutral ($M = 3.58$, $SD = 1.52$) group information, $F(1,190) = 67.05$, $p < .001$, $\eta_p^2 = .26$. There was also a main effect of failure domain: Participants were more concerned about their group's social image when the target exposed morality ($M = 4.76$, $SD = 1.45$) versus competence ($M = 4.00$, $SD = 1.64$) information, $F(1,190) = 13.70$, $p < .001$, $\eta_p^2 = .07$. The interaction was nonsignificant, $F(1,190) = .48$, $p = .49$, $\eta_p^2 = .003$.

2.2.2 Personal moral reproach

A 2 \times 2 ANOVA revealed no significant effects, $F_s < .62$, $ps > .43$, η_p^2 s $< .004$.

2.2.3 Group moral reproach

A 2×2 ANOVA revealed significant main effects of disclosure and failure domain. Participants believed that the target was more likely to morally reproach their group when he released incriminating ($M = 4.81$, $SD = 1.45$) versus neutral ($M = 3.58$, $SD = 1.19$) information, $F(1,190) = 50.62$, $p < .001$, $\eta_p^2 = .21$. In addition, participants anticipated greater group-directed moral reproach when the target exposed morality ($M = 4.75$, $SD = 1.57$) versus competence ($M = 3.64$, $SD = 1.09$) information, $F(1,190) = 40.46$, $p < .001$, $\eta_p^2 = .18$. There was also a significant disclosure \times failure domain interaction, $F(1,190) = 12.01$, $p = .001$, $\eta_p^2 = .06$. Simple effects tests showed that within both failure domains, participants anticipated greater group-directed moral reproach from the target who disclosed incriminating ($M_{\text{Morality}} = 5.62$, $SD = 1.24$; $M_{\text{Competence}} = 3.96$, $SD = 1.12$) versus neutral ($M_{\text{Morality}} = 3.83$, $SD = 1.34$; $M_{\text{Competence}} = 3.34$, $SD = .98$) group information, $ts > 2.58$, $ps < .01$, $ds > .37$.

2.3 Dependent variables

2.3.1 Target evaluations

A 2×2 ANOVA revealed the predicted main effect of target disclosure: Participants rated the target more negatively when he exposed incriminating ($M = 4.05$, $SD = .75$) versus neutral ($M = 3.78$, $SD = .82$) group information, $F(1,187) = 5.99$, $p = .02$, $\eta_p^2 = .03$. The main effect of failure domain and the interaction were nonsignificant, $F_s < .91$, $ps > .34$, η_p^2 s $< .006$.

2.3.2 Group improvement concerns

A 2×2 ANOVA revealed the predicted main effect of target disclosure: Participants expressed greater concern about group improvement when the target exposed incriminating ($M = 4.86$, $SD = 1.04$) versus neutral ($M = 4.49$, $SD = 1.27$) group information, $F(1,190) = 4.85$, $p = .03$, $\eta_p^2 = .03$. There was also a main effect of failure domain: Participants were more concerned about group improvement when the target exposed morality ($M = 5.10$, $SD = 1.00$) versus competence ($M = 4.25$, $SD = 1.18$) information, $F(1,190) = 29.32$, $p < .001$, $\eta_p^2 = .13$. The disclosure \times failure domain interaction was nonsignificant, $F(1,190) = 2.48$, $p = .12$, $\eta_p^2 = .01$.

2.4 Mediation analyses

As in Studies 1-4, social image concerns mediated the effect of target disclosure on evaluations ($ab = .11$, $SE = .04$, 95% CI = [.04, .20]) and group improvement concerns ($ab = .12$, $SE = .05$, 95% CI = [.02, .22]). Group moral reproach independently mediated the disclosure effect on improvement concerns ($ab = -.10$, $SE = .05$, 95% CI = [-.19, -.007]) but not on evaluations ($ab = .02$, $SE = .03$, 95% CI = [-.05, .08]). Personal moral reproach did not emerge as a significant mediator in either analysis ($ab_{\text{Evaluation}} = -.0005$, $SE = .005$, 95% CI = [-.01, .01]; $ab_{\text{Improvement}} = .003$, $SE = .02$, 95% CI = [-.03, .04]). The direct effects of disclosure on evaluations ($c' = .01$, $SE = .07$, 95% CI = [-.12, .15]) and group improvement ($c' = .15$, $SE = .09$, 95% CI = [-.04, .33]) were nonsignificant when the mediating variables were included in the models.

3 Discussion

The results of Study 5 demonstrate that group members react similarly to public moral advocates and targets who expose nonmoral group shortcomings. Participants were more likely to express concern about their group's social image, anticipate group-directed reproach from the target, and indicate concern about group change when the target exposed morality versus competence information. Because individuals are generally more concerned about their group appearing moral than competent (Leach, Ellemers, & Barreto, 2007), participants may have been more sensitive to disclosures about their group's morality versus competence. Nevertheless, failure domain did not moderate the target disclosure effects on participants' reactions. Indeed, regardless of whether the target publicly challenged the group's morality or competence, participants became concerned about their group's image and therefore derogated the target but also expressed greater concern about the need for group change.

The present findings also provide additional evidence that social image concerns underlie group members' reactions to public moral advocates. Specifically, social image concerns consistently mediated the effects of disclosure on both evaluations and group improvement in Studies 4 and 5. Group moral reproach, in comparison, mediated only the effect on evaluations (but not on group improvement) in Study 4 and only the effect on group improvement (but not on evaluations) in Study 5. These findings provide greater evidence that social image concerns rather than only moral reproach concerns influence individuals' reactions to public moral advocates. Furthermore, they demonstrate that the processes underlying group members'

reactions to public moral advocates are representative of a broader set of processes underlying group members' reactions to targets who publicly expose the group's shortcomings, whether moral or nonmoral. Thus, the present findings may explain reactions not only to those individuals who publicly challenge a group's immoral behaviour (e.g., whistleblowers like Joseph Darby; Schorn, 2009) but also to those individuals who publicly challenge a group's incompetence or other nonmoral failures (e.g., critics of the UK government's management of National Health Service hospitals; Goodman & Reed, 2013; Murphy, 2013).

Chapter 7

Study 6

Studies 1-5 demonstrate that group members derogate moral advocates who challenge group wrongdoing publicly because these advocates threaten the group's social image. It may therefore seem intuitive that public moral advocates can improve the evaluations that they receive if they challenge group wrongdoing privately (i.e., disclose or discuss incriminating group information only among members of the perpetrating group), a practice that does not threaten the group's social image. Furthermore, given the argument that derogation of moral advocates may undermine the ability of these advocates to promote moral behaviour change (Hornsey et al., 2002; Monin et al., 2008), moral advocates who improve the evaluations that they elicit by challenging wrongdoing privately may also be more effective in stimulating group change.

Studies 1-5 suggest, however, that reducing social image concerns to improve evaluations of public moral advocates may have the ironic consequence of undermining the influence of these advocates. Indeed, given that heightened social image concerns not only elicit derogation of public moral advocates but also raise concerns about group improvement, efforts to challenge group wrongdoing in a way that limits social image threat may also fail to raise concern among group members about the need for group change. Thus, although group members may disparage a moral advocate who challenges their group's practices publicly and therefore threatens the group's social image, this social image threat may be just what group members need to recognize the importance of group change. I examined this possibility directly in Study 6. I predicted that moral advocates would elicit derogation and raise concern about group change only when they challenged wrongdoing publicly.

1 Method

1.1 Participants

Participants were 91 undergraduate students who received course credit. Two participants with target evaluation scores more than three standard deviations beyond the mean were excluded. Accordingly, 39 male and 50 female participants ($M_{age} = 18.84$ years, $SD = 1.55$) were included in analyses.

1.2 Procedure

Participants completed the same procedure as in Study 1a with the exception that, in place of the target group membership manipulation, participants were randomly assigned to read a version of the article indicating that the target had exposed the group's wrongdoing either publicly or privately (i.e., only to members of participants' ingroup). That is, participants were randomly assigned to one of four conditions in a 2(target disclosure: incriminating, neutral) \times 2(disclosure method: public, private) design. Participants then completed the same social image concerns ($\alpha = .94$), personal moral reproach, target evaluations ($\alpha = .82$), group improvement concern ($\alpha = .91$), and target disclosure manipulation check ($r = .68$) measures used in the previous studies. In addition, to verify the success of the disclosure method manipulation, participants rated the public's awareness of the disclosure (e.g., "The public is aware of the leaked report;" two items; $r = .46$) along a 7-point scale anchored at 1(*Strongly disagree*) and 7(*Strongly agree*).

2 Results

2.1 Manipulation checks

2.1.1 Target disclosure

A 2(target disclosure) \times 2(disclosure method) ANOVA revealed a main effect of target disclosure: Participants viewed the incriminating ($M = 5.33$, $SD = 1.17$) versus neutral ($M = 3.01$, $SD = 1.02$) information as more incriminating, $F(1,85) = 97.86$, $p < .001$, $\eta_p^2 = .54$. The main effect of disclosure method and the interaction were nonsignificant, $F_s < .79$, $ps > .37$, η_p^2 s $< .01$.

2.1.2 Disclosure method

A 2 \times 2 ANOVA revealed a main effect of disclosure method: Participants indicated that the public was more aware of the disclosure when the target exposed information publicly ($M = 5.33$, $SD = 1.15$) versus privately ($M = 3.47$, $SD = 1.24$), $F(1,85) = 52.67$, $p < .001$, $\eta_p^2 = .38$. The main effect of target disclosure and the interaction were nonsignificant, $F_s < 1.00$, $ps > .32$, η_p^2 s $< .02$.

2.2 Potential mediating variables

2.2.1 Social image concerns

A 2×2 ANOVA revealed a main effect of target disclosure, $F(1,85) = 50.57, p < .001, \eta_p^2 = .37$, and a target disclosure \times disclosure method interaction, $F(1,85) = 4.41, p = .04, \eta_p^2 = .05$. Simple effects tests showed that, in both disclosure method conditions, participants were more concerned about their group's social image when the target disclosed incriminating ($M_{Private} = 4.94, SD = 1.46; M_{Public} = 5.61, SD = .89$) versus neutral ($M_{Private} = 3.64, SD = 1.33; M_{Public} = 3.22, SD = 1.10$) group information, $t_s > 3.60, p_s < .002, d_s > .78$. This effect was, however, larger in the public ($d = 1.39$) versus private ($d = .78$) disclosure condition. The main effect of disclosure method was nonsignificant, $F(1,85) = .26, p = .62, \eta_p^2 = .003$.

2.2.2 Personal moral reproach

A 2×2 ANOVA revealed no significant effects on personal moral reproach, $F_s < .84, p_s > .36, \eta_p^2_s < .02$.

2.3 Dependent variables

2.3.1 Target evaluations

A 2×2 ANOVA revealed the predicted target disclosure \times disclosure method interaction, $F(1,85) = 3.82, p = .05, \eta_p^2 = .04$ (see Figure 2). As expected, simple effects tests revealed a significant effect of target disclosure within the public conditions, $t(85) = 2.14, p = .04, d = .46$: When the target disclosed group information publicly, participants rated the target who exposed incriminating versus neutral information more negatively. The simple effect of target disclosure within the private conditions was not significant, $t(85) = .60, p = .55, d = .13$. The main effects of target disclosure and disclosure method were nonsignificant, $F_s < 1.25, p_s > .26, \eta_p^2_s < .02$.

2.3.2 Group improvement concerns

A 2×2 ANOVA revealed a main effect of target disclosure, $F(1,85) = 6.58, p = .01, \eta_p^2 = .07$, that was qualified by the predicted target disclosure \times disclosure method interaction, $F(1,85) = 8.15, p = .005, \eta_p^2 = .09$. Simple effects tests revealed a significant target disclosure effect within the public conditions, $t(85) = 3.77, p < .001, d = .82$: When the target disclosed group information publicly, participants expressed more concern about the need for group improvement

when he disclosed incriminating versus neutral information. The simple effect of target disclosure within the private disclosure conditions was not significant, $t(85) = .21, p = .84, d = .05$.¹⁰ The main effect of disclosure method was also nonsignificant, $F(1,85) = 1.18, p = .28, \eta_p^2 = .01$ (see Figure 2).

2.4 Mediation analyses

To assess the extent to which social image concerns and moral reproach mediated the interaction of target disclosure and disclosure method on both target evaluations and group improvement concerns, I conducted mediated moderation analyses with 5000 bootstrapped resamples for each of the two dependent variables (Hayes, 2013). In these models, either target evaluations or group improvement was the dependent variable, target disclosure was the independent variable, disclosure method was the moderator, and both social image concerns and moral reproach were parallel mediators. Social image concerns mediated the target disclosure \times disclosure method interaction on both evaluations ($ab = .09, SE = .06, 95\% CI = [.008, .24]$) and group improvement concerns ($ab = .11, SE = .08, 95\% CI = [.007, .32]$). The indirect effects of the interaction on evaluations ($ab = -.002, SE = .02, 95\% CI = [-.07, .03]$) and improvement concerns ($ab = .08, SE = .08, 95\% CI = [-.07, .25]$) through moral reproach, in comparison, were nonsignificant. When the mediating variables were included in the model, the target disclosure \times disclosure method interaction no longer significantly predicted evaluations, $c' = .10, SE = .07, 95\% CI = [-.04, .25]$, but continued to significantly predict group improvement concerns, $c' = .22, SE = .10, 95\% CI = [.02, .42]$. Thus, participants reacted differently to the target who challenged group wrongdoing publicly versus privately because they were more likely to exhibit heightened social image concerns when the disclosure was public. Moral reproach concerns, in comparison, did not account for the observed moderation.

3 Discussion

This study demonstrates that public moral advocates bear the burden of wielding a double-edged sword. Although the threat that they pose to a group's social image may lead them to be derogated, it may also allow them to be influential. Indeed, targets who challenged group wrongdoing publicly elicited significantly more backlash than did the other three targets, but they were also the only targets who significantly raised concerns about the need for group improvement. Targets who challenged group wrongdoing privately were not derogated any more

than were targets who released neutral group information, but they also failed to raise concern about group change. Mediation analyses revealed that public but not private moral advocates both elicited backlash and raised concern about group improvement because they posed a greater threat to the group's social image. Thus, although correlational studies suggest that individuals are more receptive to moral advocates who challenge wrongdoing privately versus publicly (Near & Miceli, 1986, 2008), the present research indicates that moral advocates may need to challenge wrongdoing publicly to effectively raise concern about change.

These findings also further demonstrate that concerns about maintaining a positive social image influence reactions to public moral advocates. If individuals' reactions to public moral advocates stem not from social image concerns but from concerns about being reproached by the advocate or about feeling immoral, individuals' reactions should not vary across public and private contexts. The results of this study, however, do not support this possibility. Instead, participants derogated the advocate and expressed concern about group change only when the advocate challenged wrongdoing in a way that threatened the group's image (i.e., publicly). Thus, the present findings are more consistent with the argument that social image concerns influence individuals' reactions to public moral advocates.

Chapter 8

General Discussion

The present studies reveal that group members derogate moral advocates who challenge their group's wrongdoing publicly, not because they are concerned about being morally reproached by the advocates but rather because they are concerned about how the advocates have made their group look in the eyes of others. Furthermore, these studies demonstrate that the same psychological process that elicits derogation of public moral advocates (i.e., social image threat) is also instrumental in allowing the advocates to promote group change. In Study 1, I provided converging evidence for these effects across two domains (e.g., academic and environmental integrity). In Study 2, I demonstrated that individuals exhibited more pronounced reactions when the advocate publicly targeted their ingroup versus outgroup, and therefore posed a greater threat to the group's image. In Studies 3 and 4, I provided evidence that individuals' reactions to public moral advocates do not simply stem from moral integrity concerns. In Study 4, I also demonstrated that because social image concerns influence individuals' reactions to public moral advocates, individuals are more sensitive to *what* moral advocates publicly expose or discuss about the group rather than *how* they do so. Specifically, participants reacted similarly to public moral advocates who were authorized versus unauthorized to challenge the group's practices. In Study 5, I obtained further evidence for the role of social image concerns by demonstrating that individuals exhibit similar reactions to public moral advocates versus targets who publicly challenge a group's nonmoral shortcomings. Lastly, in Study 6, I demonstrated that moral advocates may need to challenge a group's practices publicly and therefore threaten the group's social image to stimulate change. Although moral advocates who challenged wrongdoing privately were able to evade derogation, they were also unsuccessful in raising concern about group change.

1 Theoretical Implications

The present findings make several theoretical contributions to research on moral advocates, moral behaviour change, social image concerns, whistleblowers, and group critics. First, the present studies provide greater insight into the psychological mechanisms underlying individuals' reactions to moral advocates. Researchers have previously suggested that individuals derogate moral advocates and resist the efforts of these advocates to promote moral

behaviour change because individuals feel morally reproached by them (Minson & Monin, 2012; Monin et al., 2008). In these past studies, however, participants encountered a moral advocate who condemned their behaviour privately. The present studies demonstrate, in comparison, that when a moral advocate challenges wrongdoing publicly, social image concerns play a more central role in influencing individuals' reactions. Indeed, participants' social image concerns consistently mediated the effects observed in Studies 1-6, whereas their concerns about being morally reproached did not. Although group moral reproach significantly mediated the effect of disclosure on evaluations in Study 4 and the effect of disclosure on group improvement concerns in Study 5, neither of these effects were replicated in any of the other studies. In addition, participants' reactions to public moral advocates were similar to their reactions to targets who publicly exposed nonmoral failures, a finding that is more consistent with the possibility that social image concerns rather than moral reproach concerns underlie individuals' reactions (Gausel & Leach, 2011; Gausel et al., 2012). Furthermore, rather than responding similarly to moral advocates who challenged wrongdoing publicly versus privately, participants derogated the advocates and expressed concern about group improvement only when the advocates challenged group wrongdoing publicly. Whereas the former pattern of results would be expected if moral reproach influenced individuals' reactions, the latter would be expected if social image concerns influenced individuals' reactions. Because individuals have a strong need to belong and can reap many social and material benefits from maintaining a favourable social image (Baumeister & Leary, 1995; Hardy & Van Vugt, 2006), it may be more threatening for individuals to be perceived by others as immoral than it is to view the self as immoral or to be reproached by a seemingly morally superior other (Haidt, 2012). Thus, although moral reproach may play an important role in contexts in which advocates privately condemn the behaviour of others, it may not be the primary factor that drives individuals' reactions to moral advocates who challenge wrongdoing publicly.

Second, the present findings contribute to research on moral advocates and moral behaviour change by demonstrating that public moral advocates may wield a double-edged sword. Past research suggests that the psychological processes that elicit derogation of moral advocates who challenge wrongdoing privately may either undermine or have no influence on the ability of these advocates to promote moral behaviour change. Indeed, the majority of evidence indicates that moral advocates who challenge wrongdoing privately are not only derogated but tend to

elicit globally negative reactions from perceivers (Hornsey et al., 2002; Monin et al., 2008). This research suggests that efforts to reduce derogation of moral advocates may also improve receptiveness to advocates' implicit calls for change. There is also limited evidence that some private moral advocates (i.e., prejudice confronters in one-on-one interactions) can elicit derogation yet stimulate change. In this case, distinct psychological processes appear to influence the derogation and change responses (Czopp et al., 2006), suggesting that efforts to reduce derogation of moral advocates will not undermine the advocates' abilities to promote change. The present research demonstrates, however, that when moral advocates challenge wrongdoing publicly, the same psychological process elicits derogation of these advocates and also allows these advocates to successfully promote change. Indeed, moral advocates who challenged a group's wrongdoing publicly threatened the group's image and this ultimately led them to be derogated but to also succeed in raising concern about group change. When moral advocates challenged the group's wrongdoing privately and therefore posed less of a threat to the group's image, they were not derogated, but they were also unsuccessful in raising concern about change. Thus, if public moral advocates take steps to reduce the threat that they pose to a group's social image and therefore the backlash that they experience, they may also, ironically, undermine their influence on the group.

Third, the present findings contribute to research on social image concerns by revealing that such concerns can simultaneously elicit positive and negative reactions from group members. Previously, researchers have found that targets who engage in behaviours that threaten their ingroup's social image elicit unfavourable evaluations from other ingroup members (Chekroun & Nugier, 2011; Hornsey et al., 2005). This work suggests that threatening a group's social image concerns elicits negative reactions from group members. In other work, however, researchers have found that group members exhibit concern about making amends for historical ingroup misdeeds when they believe that these misdeeds threaten their ingroup's social image (Brown & Čehajić, 2008; Brown et al., 2008). This work demonstrates that social image threats can elicit favourable reactions from group members. Past research may therefore seem to suggest that different evaluative contexts determine whether individuals react positively versus negatively when their group's social image has been threatened. The present findings demonstrate, however, that heightened social image concerns can simultaneously elicit backlash against targets who publicly challenge group wrongdoing, but also boost apparent concern about

the need to address the wrongdoing. Thus, the present research reveals that social image concerns and the moral advocates who activate them elicit responses that are more complex and multi-faceted than those revealed by previous research.

Fourth, the present findings contribute to research on two specific types of public moral advocates: whistleblowers and intergroup critics. Indeed, despite the increasingly common presence of high-profile whistleblowers on the world stage (e.g., Edward Snowden; Larson, 2012) and the pronounced backlash that they experience (Near & Miceli, 1986; Schorn, 2009), researchers have not systematically examined the psychological processes underlying individuals' reactions to whistleblowers and the evaluative and motivational implications of these processes. Indeed, past research on whistleblowing has primarily examined predictors of whistleblowing behaviour (e.g., Near & Miceli, 2011; Waytz, Dungan, & Young, 2013) or used correlational methods to examine the consequences of whistleblowing (e.g., Miceli & Near, 1994; Near & Miceli, 1986). Although this work provides valuable insight into the reasons why whistleblowers expose wrongdoing and the variety of factors that may predict backlash toward them, they do not identify factors that have a causal impact on individuals' reactions to whistleblowers, nor do they explain why individuals often react negatively to whistleblowers. The present research is therefore the first to provide insight into the psychological processes underlying reactions to whistleblowers based on a systematic examination of individuals' perceptions. It also demonstrates the need to consider the role of social image concerns in theoretical models of whistleblowing.

Although I focused on examining reactions to moral advocates who publicly exposed group wrongdoing (e.g., whistleblowers and individuals who receive authorization to declassify information describing wrongdoing), I argue that the present findings also provide greater insight into reactions to intergroup critics, individuals who call attention to known problematic characteristics or practices of groups (e.g., racism; Hornsey et al., 2005; Hornsey, Trembath, & Gunthorpe, 2004). Previously, researchers examining responses to intergroup critics have focused primarily on comparing reactions to ingroup versus outgroup targets who criticize a group in private contexts (Hornsey, 2005, 2006; Hornsey & Imani, 2004; Hornsey et al., 2002). It remains possible, however, that individuals who criticize groups publicly elicit reactions similar to those examined in the present research. Just as social image concerns influence reactions to targets who expose wrongdoing publicly, they may also underlie reactions to

individuals who call attention to problematic practices that are already known to the public. In future research, it will be important to verify this directly.

2 Limitations and Future Directions

In Studies 3 and 4, I measured participants' concerns about their own morality and their group's morality. Whereas the personal moral integrity items may primarily assess shame-related concerns (i.e., concerns about immoral character or attributes; e.g., Shepherd, Spears, & Manstead, 2013; Tracy & Robins, 2006), the group moral integrity items may primarily assess guilt-related concerns (i.e., concerns about immoral behaviour; e.g., Branscombe, Slugoski, & Kappen, 2004; Doosje et al., 1998). Thus, although the present research suggests that neither personal shame nor collective guilt explains group members' reactions to the public moral advocates, it is possible that either personal guilt or collective shame does. That is, group members may derogate public moral advocates and express concern about group change because they personally feel guilty about the wrongdoing or because they feel ashamed of their group's collective immoral character. Given that both shame and guilt reflect concerns about how individuals view themselves or their group, rather than how they are viewed by others (Gausel & Leach, 2011; Rees et al., 2013; Tracy & Robins, 2006), individuals should experience similar levels of shame and guilt when moral advocates expose wrongdoing privately versus publicly. In Study 5, however, participants derogated the advocate and expressed concern about group change only when the advocate exposed wrongdoing publicly. Thus, the present findings are more consistent with the argument that social image concerns, rather than personal or collective feelings of guilt or shame, influence individuals' reactions to public moral advocates. In future research, it will be important to verify this empirically.

In the present research, I assessed participants' concern about the need for group change, rather than actual behaviour aimed at achieving group change. Furthermore, participants' concern about group improvement stemmed from concerns about the group's social image and was elevated only when group wrongdoing was exposed publicly. Individuals' motives for expressing concern about group change in response to allegations of group wrongdoing may, therefore, be superficial and represent a defensive attempt to "save face" rather than a genuine desire to address the group's shortcomings. Thus, it remains unknown whether public moral advocates are effective in stimulating enduring group change. Because individuals are motivated to appear

consistent (Leary & Cox, 2008), however, group members who publicly express concern about the need for change within the group may feel compelled to actually perform behaviours that will stimulate change. If group members sustain these behaviours over a period of time, new standards and norms about acceptable behaviours may develop and ultimately maintain group change over time (Czopp et al., 2006; Rabinovich & Morton, 2010). In any case, although threatening a group's social image may lead group members to support change for group-serving reasons, the present findings suggest that it may lead to a greater degree of change than does challenging group wrongdoing in a way that does not threaten the group's image (i.e., privately). Thus, although it may seem intuitive to encourage individuals to behave more morally by convincing them that they are immoral or by having a highly moral individual reprimand them, the present findings suggest that it may be more effective in some situations to make individuals believe that others will view them *as though* they are immoral.

Previous research suggests that strongly identified group members react more defensively to group threats than do weakly identified group members (Doosje et al., 1998; Miron, Branscombe, & Biernat, 2010; Tauber & van Zomeren, 2013). Results of additional analyses reveal, however, that group identification did not moderate any of the effects reported above. It may be that the limited number of participants low in group identification (i.e., across studies, only 11-17% of participants scored below the group identification scale midpoint) made it difficult to detect moderation by group identification. Replicating the present studies with a sample in which there is a greater proportion of participants low in group identification may, therefore, yield significant moderating effects. I also note, however, that differences between strongly and weakly identified group members tend to emerge only when a group's alleged wrongdoing or failure is ambiguous (Doosje et al., 1998). In such cases, highly identified group members can construct justifications that make it possible to downplay the severity of their group's behaviour (Wohl, Branscombe, & Klar, 2006). When a group's behaviour is clearly wrong or inadequate, however, these justifications are more difficult to construct and group members tend to respond similarly regardless of their group identification (Iyer, Jetten, & Haslam, 2012). Thus, when public moral advocates challenge behaviour that is clearly wrong or inadequate, strongly and weakly identified group members may exhibit similar reactions. In future research, it will be important to examine this directly.

In the present research, group members derogated a moral advocate who publicly challenged wrongdoing committed by their group. Past research suggests, however, that perceivers who are implicated in morally questionable behaviour derogate targets who implicitly condemn versus comply with the behaviour, whereas perceivers who are not implicated in the behaviour react more favourably to targets who condemn versus comply with the behaviour (Monin et al., 2008; Monin & O'Connor, 2011). It may be the case, therefore, that uninvolved, third-party observers respond more favourably to public moral advocates than to nonadvocates. Indeed, participants in Study 2 were less likely to derogate a target who exposed incriminating information about participants' outgroup versus ingroup.

On the other hand, perceivers who are not implicated in the wrongdoing challenged by public moral advocates may also exhibit backlash toward them. Indeed, public moral advocates may threaten the ability of third-party observers to view key organizations and social systems as legitimate and fair (Kaiser et al., 2006; Sumanth et al., 2011). That is, when moral advocates more directly challenge the legitimacy of existing social systems by highlighting unethical practices within organizations or discrimination against minority group members (Hornsey & Imani, 2004; Hornsey et al., 2002; Jost, Banaji, & Nosek, 2004), perceivers may feel threatened by these advocates and therefore derogate them, regardless of their own involvement in the wrongdoing. It will be important to examine these possibilities directly in future research.

3 Conclusion

Although public moral advocates often strive to promote positive change within groups, group members may perceive them to be harmful rather than helpful to the group. Indeed, the present findings indicate that public moral advocates wield a double-edged sword. Publicly challenging group wrongdoing and therefore threatening the group's image may allow public moral advocates to successfully encourage change within the group. This may, however, come at a social evaluative cost to the advocates themselves.

References

- Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, 7, 275-295. doi: 10.1521/soco.1989.7.3.275
- Bashir, N. Y., Lockwood, P., Chasteen, A. L., Nadolny, D., & Noyes, I. (2013). The ironic impact of activists: Negative stereotypes reduce social change influence. *European Journal of Social Psychology*, 43, 614–626. doi: 10.1002/ejsp.1983
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497-529. doi: 10.1037/0033-2909.117.3.497
- Berger, J., Rosenholtz, S. J., & Zelditch, M., Jr. (1980). Status organizing processes. *Annual Review of Sociology*, 6, 479-508. doi: 10.1146/annurev.so.06.080180.002403
- Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, 51, 149-166. doi: 10.1111/j.2044-8309.2010.02011.x
- Branscombe, N. R., Slugoski, B., & Kappen, D. M. (2004). The measurement of collective guilt: What it is and what it is not. In N. R. Branscombe & B. Doosje (Eds.), *Collective guilt: International perspectives* (pp. 16–34). New York: Cambridge University Press.
- Brown, R., & Čehajić, S. (2008). Dealing with the past and facing the future: Mediators of the effects of collective guilt and shame in Bosnia and Herzegovina. *European Journal of Social Psychology*, 38, 669–684. doi: 10.1002/ejsp.466

- Brown, R., González, R., Zagefka, H., Manzi, J., & Čehajić, S. (2008). Nuestra culpa: Collective guilt and shame as predictors of reparation for historical wrongdoing. *Journal of Personality and Social Psychology*, 94, 75–90. doi: 10.1037/0022-3514.94.1.75
- Buchanan, J., & Gorbunova, Y. (2012, August 10). Sochi Olympics have an ugly side. *The Washington Post*. Retrieved from http://www.washingtonpost.com/opinions/sochi-olympics-have-an-ugly-side/2012/08/10/e86770be-e24c-11e1-98e7-89d659f9c106_story.html
- Campbell, C. (2013, June 10). Edward Snowden comes forward as NSA whistle-blower, surfaces in Hong Kong. *Time*. Retrieved from <http://world.time.com/2013/06/10/edward-snowden-comes-forward-as-nsa-whistleblower-surfaces-in-hong-kong/>
- Carr, H. L., & Vignoles, V. L. (2011). Keeping up with the Joneses: Status projection as symbolic self-completion. *European Journal of Social Psychology*, 41, 518-527. doi: 10.1002/ejsp.812
- Chekroun, P., & Nugier, A. (2011). “I’m ashamed because of you, so please, don’t do that!”: Reactions to deviance as a protection against a threat to social image. *European Journal of Social Psychology*, 41, 479-488. doi: 10.1002/ejsp.809
- Cryder, C. E., Springer, S., & Morewedge, C. K. (2012). Guilty feelings, targeted actions. *Personality and Social Psychology Bulletin*, 38, 607-618. doi: 10.1177/0146167211435796
- Czopp, A. M., & Monteith, M. J. (2003). Confronting prejudice (literally): Reactions to confrontations of racial and gender bias. *Personality and Social Psychology Bulletin*, 29, 532-544. doi: 10.1177/0146167202250923

- Czopp, A. M., Monteith, M. J., & Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology*, *90*, 784-803. doi: 10.1037/0022-3514.90.5.784
- de Hooge, I. E., Breugelmans, S. M., & Zeelenberg, M. (2008). Not so ugly after all: When shame acts as a commitment device. *Journal of Personality and Social Psychology*, *95*, 933–943. doi: 10.1037/a0011991
- de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2010). Restore and protect motivations following shame. *Cognition and Emotion*, *24*, 111–127. doi: 10.1080/02699930802584466
- Dehghan, S. K. (2014, March 12). Ban Ki-moon rebukes Iranian president for human rights failings. *The Guardian*. Retrieved from <http://www.theguardian.com/world/2014/mar/12/ban-ki-moon-iran-rouhani-failing-human-rights>
- Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. R. (1998). Guilty by association: When one's group has a negative history. *Journal of Personality and Social Psychology*, *75*, 872-886. doi: 10.1037//0022-3514.75.4.872
- Garcia, D. M., Reser, A. H., Amo, R. B., Redersdorff, S., & Branscombe, N. R. (2005). Perceivers' responses to in-group and out-group members who blame a negative outcome on discrimination. *Personality and Social Psychology Bulletin*, *31*, 769-780. doi: 10.1177/0146167204271584
- Gausel, N., & Leach, C. W. (2011). Concern for self-image and social image in the management of moral failure: Rethinking shame. *European Journal of Social Psychology*, *41*, 468-478. doi: 10.1002/ejsp.803
- Gausel, N., Leach, C. W., Vignoles, V. L., & Brown, R. (2012). Defend or repair? Explaining responses to in-group moral failure by disentangling feelings of shame, rejection, and

- inferiority. *Journal of Personality and Social Psychology*, 102, 941-960. doi: 10.1037/a0027233
- Gerber, J., & Wheeler, L. (2009). On being rejected: A meta-analysis of experimental research on rejection. *Perspectives on Psychological Science*, 4, 468-488. doi: 10.1111/j.1745-6924.2009.01158.x
- Golec de Zavala, A., & Kossowska, M. (2011). Correspondence between images of terrorists and preferred approaches to counterterrorism: The moderating role of ideological orientations. *European Journal of Social Psychology*, 41, 538-549. doi: 10.1002/ejsp.810
- Goodman, P., & Reed, J. (2013, July 17). Is the Tory attack on the NHS justified? *The Guardian*. Retrieved from <http://www.theguardian.com/commentisfree/2013/jul/17/is-tory-attack-nhs-justified>.
- Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, 98, 392-404. doi: 10.1037/a0017346
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Pantheon.
- Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, 32, 1402-1413. doi: 10.1177/0146167206291006
- Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Hornsey, M. J. (2005). Why being right is not enough: Predicting defensiveness in the face of group criticism. *European Review of Social Psychology*, 16, 301-334. doi: 10.1080/10463280500436040

- Hornsey, M. J. (2006). Ingroup critics and their influence on groups. In T. Postmes & J. Jetten (Eds.), *Individuality and the group: Advances in social identity* (pp. 74-92). Thousand Oaks, CA: Sage.
- Hornsey, M. J., De Bruijn, P., Creed, J., Allen, J., Ariyanto, A., & Svensson, A. (2005). Keeping it in-house: How audience affects responses to group criticism. *European Journal of Social Psychology*, 35, 291-312. doi: 10.1002/ejsp.246
- Hornsey, M. J., & Imani, A. (2004). Criticizing groups from the inside and the outside: An identity perspective on the intergroup sensitivity effect. *Personality and Social Psychology Bulletin*, 30, 365–383. doi: 10.1177/0146167203261295
- Hornsey, M. J., Oppes, T., & Svensson, A. (2002). ‘It’s OK if we say it but you can’t’: Responses to intergroup and intragroup criticism. *European Journal of Social Psychology*, 32, 293–307. doi: 10.1002/ejsp.90
- Hornsey, M. J., Trembath, M., & Gunthorpe, S. (2004). ‘You criticize because you care’: Identity attachment, constructiveness, and the intergroup sensitivity effect. *European Journal of Social Psychology*, 34, 499-518. doi: 10.1002/ejsp.212
- Iyer, A., Jetten, J., & Haslam, S. A. (2012). Sugaring o’er the devil: Moral superiority and group identification help individuals downplay the implications of ingroup rule-breaking. *European Journal of Social Psychology*, 42, 141-149. doi: 10.1002/ejsp.864
- Iyer, A., Schmader, T., & Lickel, B. (2007). Why individuals protest the perceived transgressions of their country: The role of anger, shame, and guilt. *Personality and Social Psychology Bulletin*, 33, 572–587. doi: 10.1177/0146167206297402
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25, 881-919. doi: 10.1111/j.1467-9221.2004.00402.x

- Kaiser, C. R., & Miller, C.T. (2001). Stop complaining! The social costs of making attributions to discrimination. *Personality and Social Psychology Bulletin*, 27, 254–263. doi: 10.1177/0146167201272010
- Kaiser, C. R., Dyrenforth, P. S., & Hagiwara, N. (2006). Why are attributions to discrimination interpersonally costly? A test of system- and group-justifying motivations. *Personality and Social Psychology Bulletin*, 32, 1523–1536. doi: 10.1177/0146167206291475
- Klein, O., & Azzi, A. E. (2001). The strategic confirmation of meta-stereotypes: How group members attempt to tailor an out-group's representation of themselves. *British Journal of Social Psychology*, 40, 279–293. doi: 10.1348/014466601164759
- Larson, E. (2012, October 30). U.K. whistle-blowing cases rise 276% in four years, Kroll says. *Bloomberg News*. Retrieved from <http://www.bloomberg.com/news/2012-10-30/u-k-whistle-blowing-cases-rise-276-in-four-years-kroll-says.html>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93, 234–249. doi:10.1037/0022-3514.93.2.234
- Leary, M. R., & Cox, C. B. (2008). Belongingness motivation: A mainspring of social action. In J. Y. Shah & W. L. Gardner. (Ed.), *Handbook of motivation science* (pp. 27–40). New York, NY: The Guilford Press.
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, 107, 34–47. doi: 10.1037//0033-2909.107.1.34
- Leung, R. (2009, February 9). Abuse at Abu Ghraib. *CBS News*. Retrieved from http://www.cbsnews.com/2100-500164_162-615781.html

- Lewis, H. B. (1971). *Shame and guilt in neurosis*. New York, NY: International Universities Press.
- Lun, J., Mesquita, B., & Smith, B. (2011). Self- and other-presentational styles in the Southern and Northern United States: An analysis of personal ads. *European Journal of Social Psychology, 41*, 435-445. doi: 10.1002/ejsp.804
- Macdonald, N. (2014, March 26). Obama's 'Snowden reform' of NSA spying won't help Edward Snowden. *Canadian Broadcasting Corporation*. Retrieved from <http://www.cbc.ca/news/world/obama-s-snowden-reform-of-nsa-spying-won-t-help-edward-snowden-1.2586328>
- MacDonald, G., & Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin, 131*, 202–223. doi: 10.1037/0033-2909.131.2.202
- Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the “porcupine problem.” *Journal of Personality and Social Psychology, 92*, 42-55. doi: 10.1037/0022-3514.92.1.42
- Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology, 18*, 287-292. doi: 10.1002/ejsp.2420180308
- Marques, J. M., Yzerbyt, V. Y., & Leyens, J.- P. (1988). The ‘black sheep effect’: Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology, 18*, 1-16. doi: 10.1002/ejsp.2420180102

- Miceli, M. P., & Near, J. P. (1994). Relationships among value congruence, perceived victimization, and retaliation against whistle-blowers. *Journal of Management*, 20, 773-794. doi: 10.1177/014920639402000405
- Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science*, 3, 200-207. doi: 10.1177/1948550611415695
- Miron, A. M., Branscombe, N. R., & Biernat, M. (2010). Motivated shifting of justice standards. *Personality and Social Psychology Bulletin*, 36, 768–779. doi: 10.1177/0146167210370031
- Monin, B. (2007). Holier than me? Threatening social comparisons in the moral domain. *International Review of Social Psychology*, 20, 53-68.
- Monin, B., & O'Connor, K. (2011). Reactions to defiant deviants: Deliverance or defensiveness? In J. Jetten & M. J. Hornsey (Eds.), *Rebels in groups: Dissent, deviance, difference, and defiance* (pp. 261-280). Chichester, UK: Wiley-Blackwell.
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, 95, 76-93. doi: 10.1037/0022-3514.95.1.76
- Murphy, B. (2013, July 17). Tories facing online backlash over virulent criticism of NHS as IloveourNHS campaign highlights thousands of positive treatment stories. *TNT Magazine*. Retrieved from <http://www.tntmagazine.com/news/uk/tories-facing-online-backlash-over-virulent-criticism-of-nhs-as-iloveournhs-campaign-highlights-thousands-of-positive-treatment-stories>
- Near, J. P., & Miceli, M. P. (1986). Retaliation against whistle blowers: Predictors and effects. *Journal of Applied Psychology*, 71, 137-145. doi: 10.1037//0021-9010.71.1.137

- Near, J. P., & Miceli, M. P. (2008). Wrongdoing, whistle-blowing, and retaliation in the U.S. government. What have researchers learned from the Merit Systems Protection Board (MSPB) Survey results? *Review of Public Personnel Administration*, 28, 263-281. doi: 10.1177/0734371X08319153
- Near, J. P., & Miceli, M. P. (2011). Integrating models of whistle-blowing and wrongdoing: A proposal for a new research agenda. In J. Jetten & M. J. Hornsey (Eds.), *Rebels in groups: Dissent, deviance, difference, and defiance* (pp. 302-323). Chichester, UK: Wiley-Blackwell.
- O'Keefe, E. (2014, April 2). Two key senators say they will vote to declassify parts of CIA report. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/blogs/post-politics/wp/2014/04/02/two-key-senators-say-they-will-vote-to-declassify-cia-report>
- Owuamalam, C. K., & Zagefka, H. (2011). Downplaying a compromised social image: The effect of metastereotype valence on social identification. *European Journal of Social Psychology*, 41, 528-537. doi: 10.1002/ejsp.805
- Parks, C. D., & Stone, A. B. (2010). The desire to expel unselfish members from the group. *Journal of Personality and Social Psychology*, 99, 303-310. doi: 10.1037/a0018403
- Peetz, J., Gunn, G. R., & Wilson, A. E. (2010). Crimes of the past: Defensive temporal distancing in the face of past in-group wrongdoing. *Personality and Social Psychology Bulletin*, 36, 598-611. doi: 10.1177/0146167210364850
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879-891. doi: 10.3758/BRM.40.3.879

- Rabinovich, A., & Morton, T. A. (2010). Who says we are bad people? The impact of criticism source and attributional content on responses to group-based criticism. *Personality and Social Psychology Bulletin*, 36, 524-536. doi: 10.1177/0146167210362980
- Rees, J. H., Allpress, J. A., & Brown, R. (2013). Nie wieder: Group-based emotions for ingroup wrongdoing affect attitudes toward unrelated minorities. *Political Psychology*, 34, 387-407. doi: 10.1111/pops.12003
- Sabini, J., & Silver, M. (1982). Moral reproach. In J. Sabini & M. Silver (Eds.), *Moralities of everyday life* (pp. 35–53). Oxford, England: Oxford University Press
- Savage, C. (2014, April 1). Letter tells of U.S. searches for emails and calls. *The New York Times*. Retrieved from <http://www.nytimes.com/2014/04/02/us/politics/letter-puts-focus-on-us-searches-for-americans-emails-and-calls.html>
- Schmader, T., & Lickel, B. (2006). The approach and avoidance function of guilt and shame emotions: Comparing reactions to self-caused and other caused wrongdoing. *Motivation and Emotion*, 30, 43–56. doi: 10.1007/s11031-006-9006-0
- Schorn, D. (2009, February 11). Exposing the truth of Abu Ghraib. *CBS News*. Retrieved from http://www.cbsnews.com/8301-18560_162-2238188.html?pageNum=4
- Shepherd, L., Spears, R., & Manstead, A. S. R. (2013). “This will bring shame on our nation”: The role of anticipated group-based emotions on collective action. *Journal of Experimental Social Psychology*, 49, 42-57. doi: 10.1016/j.jesp.2012.07.011
- Smith, R. H., Webster, J. M., Parrott, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, 83, 138–159. doi: 10.1037/0022-3514.83.1.138

- Stone, J., Whitehead, J., Schmader, T., & Focella, E. (2011). Thanks for asking: Self-affirming questions reduce backlash when stigmatized targets confront prejudice. *Journal of Experimental Social Psychology*, 47, 589-598. doi: 10.1016/j.jesp.2010.12.016
- Sumanth, J. J., Mayer, D. M., & Kay, V. S. (2011). Why good guys finish last: The role of justification motives, cognition, and emotion in predicting retaliation against whistleblowers. *Organizational Psychology Review*, 1, 165-184. doi: 10.1177/2041386611398283
- Sutton, R. M., Elder, T. J., & Douglas, K. M. (2006). Reactions to internal and external criticism of outgroups: Social convention in the intergroup sensitivity effect. *Personality and Social Psychology Bulletin*, 32, 563-575. doi: 10.1177/0146167205282992
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Tangney, J. P. (1996). Conceptual and methodological issues in the assessment of shame and guilt. *Behavior Research and Therapy*, 34, 741-754. doi: 10.1016/0005-7967(96)00034-4
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, 70, 1256-1269. doi: 10.1037/0022-3514.70.6.1256
- Tauber, S., & van Zomeren, M. (2013). Outrage towards whom? Threats to moral group status impede striving to improve via outgroup-directed outrage. *European Journal of Social Psychology*, 43, 149-159. doi: 10.1002/ejsp.1930
- Tejada, C., & Burkitt, L. (2013, December 5). Biden criticizes China on media, human rights. *The Wall Street Journal*. Retrieved from <http://online.wsj.com/news/articles/SB10001424052702303997604579239751549316272>

- Tracy, J. L., & Robins, R. W. (2006). Appraisal antecedents of shame and guilt: Support for a theoretical model. *Personality and Social Psychology Bulletin*, 32, 1339-1351.
- van Leeuwen, E., Van Den Bosch, M., Castano, E., & Hopman, P. (2010). Dealing with deviants: The effectiveness of rejection, denial, and apologies on protesting the public image of a group. *European Journal of Social Psychology*, 40, 282-299. doi: 10.1002/ejsp.622
- Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology*, 49, 1027-1033. doi: 10.1016/j.jesp.2013.07.002
- Wohl, M. J. A., Branscombe, N. R., & Klar, Y. (2006). Collective guilt: Emotional reactions when one's group has done wrong or been wronged. *European Review of Social Psychology*, 17, 1-37. doi: 10.1080/10463280600574815
- Wohl, M. J. A., Hornsey, M. J., & Bennett, S. H. (2012). Why group apologies succeed and fail: Intergroup forgiveness and the role of primary and secondary emotions. *Journal of Personality and Social Psychology*, 102, 306-322. doi:10.1037/a0024838
- Ybarra, O., Park, H., Stanik, C., & Lee, D. S. (2012). Self-judgment and reputation monitoring as a function of the fundamental dimensions, temporal perspective, and culture. *European Journal of Social Psychology*, 42, 200-209. doi: 10.1002/ejsp.854

Notes

¹ In the studies reported, I observed a similar pattern of results when I included all participants in analyses.

² In this study, I also manipulated the target's identity as either an environmentalist or a nonenvironmentalist. When analyzing the data for Study 1b separately, this variable did not interact with any of the other independent variables on any measures. I obtained the same pattern of results on all measures regardless of whether this variable was included in analyses.

³ At the beginning of each study, participants completed a measure assessing the extent to which they identify with their university group. Given that my *a priori* hypotheses did not involve group identification, I did not include group identification as a variable in my initial set of analyses. I have, however, re-analyzed my data to assess whether group identification moderates the effects that I report. I summarize the results of these additional analyses in the general discussion.

⁴ Because the variance within some target disclosure \times target group membership \times moral domain cells was zero (i.e., all participants correctly identified the target's group membership), it was not possible to conduct a logistic regression predicting target group membership manipulation check responses from disclosure, target group membership, and domain. When analyzing the data separately within each disclosure \times membership \times domain cell, participants were significantly more likely to correctly versus incorrectly identify the target's group membership, $\chi^2_s > 11.84$, $ps < .002$.

⁵ In all studies, I observed a similar pattern of results when I excluded participants who had incorrectly identified the target's group membership.

⁶ As in Study 1, participants within each wrongdoer membership \times target membership cell were significantly more likely to correctly versus incorrectly identify the target's group membership, $\chi^2_s > 8.90$, $ps < .004$.

⁷ One could argue that the target's group membership did not moderate the results of Studies 1a and 1b because the university ingroup-outgroup distinction used in these studies was not

meaningful to participants. In Study 2, however, this ingroup-outgroup distinction did have an impact on participants when it was used to manipulate the group targeted by the advocate's disclosure. These findings suggest, therefore, that university group membership was psychologically meaningful to participants in Studies 1a and 1b.

⁸ The inclusion of several potential mediators in Studies 3 and 4 may raise concerns about multicollinearity in these analyses. I note, however, that the variance inflation factors for all five potential mediators were small in both studies (VIFs < 2.17).

⁹ Separate analyses within each domain showed that morality condition participants viewed the incriminating versus neutral information to be more incriminating with respect to the group's morality, $t(190) = 9.03, p < .001, d = 1.31$. Competence condition participants viewed the incriminating versus neutral information to be more incriminating with respect to the group's competence, $t(190) = 7.51, p < .001, d = 1.09$.

¹⁰ Framed differently, the simple effects of disclosure method on target evaluations and group improvement concerns were significant within the incriminating condition, $ts > 2.00, ps < .05, ds > .43$, but not within the neutral condition, $ts < .78, ps > .44, ds < .17$.

Appendix A: Sample Target Disclosure Articles

Study 1a: Neutral Disclosure Condition

Confidential Report Leaked to Public Reveals How Ethical University of Toronto Students Are

A confidential report recently leaked to CBC reveals that University of Toronto students are no more likely to engage in academic misconduct (i.e., plagiarism and cheating) than are students at other Ontario universities.

The report was commissioned by the Ontario Federation for University Affairs (OFUA), a group of student and faculty representatives from universities across Ontario who have been elected to oversee a wide variety of issues at schools across the country, such as program development, student life, and academic performance. The goal of the OFUA is to allow universities to work together to strengthen post-secondary education institutions in Ontario.

Recently, in response to data revealing rising rates of plagiarism and cheating among students at universities in Ontario, the OFUA has implemented a new program designed to encourage academic integrity among university students. This program requires universities to implement initiatives on their campuses to reduce rates of academic misconduct and monitor the academic integrity of their students. University officials decided to have the OFUA manage the assessments at all universities in Ontario because it would allow all schools to be evaluated using the same standards. It would also make it easier for administrative officials at each university to compare their school's performance to the average performance of all schools in Ontario. The reports, which are commissioned by the OFUA, are conducted by the independent research firm CESI. Each final report summarizing a school's performance is strictly confidential and disclosed only to administrative officials at that university and members of the OFUA. The reports are intended to be used for internal evaluation purposes only and are not permitted to be released to other universities or to the public.

Earlier this week, CBC obtained a copy of the University of Toronto's report from Andrew Baker, a fourth-year student at the University of Toronto [Queen's University]. As an OFUA student member, Baker was able to gain access to the report through the OFUA's secure electronic database. Although OFUA members are not permitted to share the content of any

OFUA report with other individuals, in a statement to CBC, Baker emphasized the critical need to leak the University of Toronto's report, noting that "UofT is a public institution, so the public should be given some reassurance that UofT students are as likely as students at other schools to earn their grades fairly."

An inspection of the report reveals that UofT students don't commit any more academic offences than do students at most universities in Ontario:

In comparison to students at other schools, a similar percentage of UofT students purchase essays and term papers from such services as essayexperts.ca and customessays.com with the intention of submitting them to course instructors as their own work, a practice that is considered an academic offence at all post-secondary institutions in Ontario. In addition, a large-scale analysis of student papers using turnitin.com, a software program commonly used by course instructors to detect plagiarism, revealed that an average number of students at UofT plagiarize essays and papers by copying and pasting portions of websites into their assignments without providing references for their sources.

The assessment also examined rates of academic misconduct during tests and exams and found that UofT students don't cheat on tests and exams any more often than do students at other Ontario universities. Indeed, UofT students aren't any more likely to use a variety of unauthorized aids during tests and exams (e.g., copying answers from other students' tests, hiding cheat sheets in pockets, writing notes on water bottle labels). Furthermore, in comparison to other schools, a similar proportion of UofT students engage in unethical practices while preparing for tests and exams, such as purchasing unauthorized copies of restricted exams, exams to which students are prohibited to have access both before and after the exam in order to prevent students from sharing the content of the exams with those who may write them in the future.

In addition, the report indicated that the number of extreme incidents of academic misconduct among UofT students isn't any higher than are rates of extreme academic misconduct at other schools: Students at UofT aren't any more likely to engage in such highly unethical practices as breaking into professors' email accounts to obtain unauthorized copies of exams, hiding cell phones in sleeves and jacket pockets during exams to share answers with other students, and

writing notes on parts of the body that are hidden under clothing but can be read during a trip to the washroom.

Lastly, online polls completed by a large number of undergraduate students show that an average percentage of UofT students admit to cheating or plagiarizing at least once in the past. In particular, UofT students were no more likely than were students at other schools to indicate that they have asked classmates for help on assignments, either in person, via email, or on social networking sites, without clarifying with the instructor beforehand that they were permitted to do so.

The report also indicates that, in comparison to students at other schools, UofT students are not more likely to engage in misconduct by submitting the same essay or paper to more than one course. Overall, the report reveals that the total score for University of Toronto students on this assessment of academic integrity was a 73%, which is similar to the provincial average of 72% (Note: Higher scores indicate less misconduct).

When contacted by CBC to ask for his reaction to the contents of the report, Minister of Training, Colleges and Universities John Milloy revealed that government officials are considering calling for changes to government funding for undergraduate academic programs at the University of Toronto, a move aimed at adjusting undergraduate tuition fees for students at the school, according to the school's performance. "We had no idea of how UofT students are performing when it comes to academic conduct and we probably never would have known if UofT's academic integrity report hadn't been leaked. It's important that we reward schools that make an effort to uphold standards of proper academic conduct" Milloy noted. "We want to send the message to students at universities across Ontario that it's very important to behave ethically in the classroom."

Furthermore, in response to UofT's performance on the academic integrity report, such employers as IBM and the Government of Canada, which hire a large number of UofT students following graduation, have discussed the possibility of increasing the number UofT students that they intend to recruit for jobs in the future. Marilyn Coxwell, a member of the IBM Board of Directors, noted that "Academic integrity among students influences the value of a University of Toronto degree. The academic integrity of the students we aim to hire affects how confident we can be that they are really learning the skills that their degrees would typically guarantee."

Aside from such responses from the government and potential employers, some individuals have speculated that the report may also influence the university's standing in world university rankings and its reputation among prominent donors. Indeed, at a time when donors are becoming increasingly conscious of the ethical integrity of the institutions they support financially, this report of UofT students may ensure that donors will be willing to support fundraising initiatives aimed at improving undergraduate programs and resources at the University of Toronto.

Study 1a: Negative Disclosure Condition

Confidential Report Leaked to Public Reveals that University of Toronto Students Not As Ethical As We May Think

A confidential report recently leaked to CBC reveals that University of Toronto students are particularly likely to engage in academic misconduct (i.e., plagiarism and cheating).

The report was commissioned by the Ontario Federation for University Affairs (OFUA), a group of student and faculty representatives from universities across Ontario who have been elected to oversee a wide variety of issues at schools across the country, such as program development, student life, and academic performance. The goal of the OFUA is to allow universities to work together to strengthen post-secondary education institutions in Ontario. Recently, in response to data revealing rising rates of plagiarism and cheating among students at universities in Ontario, the OFUA has implemented a new program designed to encourage academic integrity among university students. This program requires universities to implement initiatives on their campuses to reduce rates of academic misconduct and monitor the academic integrity of their students. University officials decided to have the OFUA manage the assessments at all universities in Ontario because it would allow all schools to be evaluated using the same standards. It would also make it easier for administrative officials at each university to compare their school's performance to the average performance of all schools in Ontario. The reports, which are commissioned by the OFUA, are conducted by the independent research firm CESI. Each final report summarizing a school's performance is strictly confidential and disclosed only to administrative officials at that university and members of the OFUA. The reports are intended to be used for internal evaluation purposes only and are not permitted to be released to other universities or to the public.

Earlier this week, CBC obtained a copy of the University of Toronto's report from Andrew Baker, a fourth-year student at the University of Toronto [Queen's University]. As an OFUA student member, Baker was able to gain access to the report through the OFUA's secure electronic database. Although OFUA members are not permitted to share the content of any OFUA report with other individuals, in a statement to CBC, Baker emphasized the critical need to leak the University of Toronto's report, noting that "UofT is a public institution, so if UofT students are behaving unethically by cheating the system, the public should know."

An inspection of the report reveals the "unethical behaviour" noted by Baker. Indeed, it indicates that UofT students commit more academic offences than do students at most universities in Ontario:

In comparison to students at other schools, a greater percentage of UofT students purchase essays and term papers from such services as essayexperts.ca and customessays.com with the intention of submitting them to course instructors as their own work, a practice that is considered an academic offence at all post-secondary institutions in Ontario. In addition, a large-scale analysis of student papers using turnitin.com, a software program commonly used by course instructors to detect plagiarism, revealed that many students at UofT plagiarize essays and papers by copying and pasting portions of websites into their assignments without providing references for their sources.

The assessment also examined rates of academic misconduct during tests and exams and found that UofT students cheat on tests and exams more often than do students at other Ontario universities. Indeed, UofT students are more likely to use a variety of unauthorized aids during tests and exams (e.g., copying answers from other students' tests, hiding cheat sheets in pockets, writing notes on water bottle labels). Furthermore, a greater proportion of UofT students engage in unethical practices while preparing for tests and exams, such as purchasing unauthorized copies of restricted exams, exams to which students are prohibited to have access both before and after the exam in order to prevent students from sharing the content of the exams with those who may write them in the future.

In addition, the report indicated that the number of extreme incidents of academic misconduct is particularly high among UofT students: Students at UofT are more likely to engage in such highly unethical practices as breaking into professors' email accounts to obtain unauthorized

copies of exams, hiding cell phones in sleeves and jacket pockets during exams to share answers with other students, and writing notes on parts of the body that are hidden under clothing but can be read during a trip to the washroom.

Lastly, online polls completed by a large number of undergraduate students show that a very large percentage of UofT students admit to cheating or plagiarizing at least once in the past. In particular, many UofT students indicated that they have asked classmates for help on assignments, either in person, via email, or on social networking sites, without clarifying with the instructor beforehand that they were permitted to do so.

The report did indicate that UofT students are unlikely to engage in misconduct by submitting the same essay or paper to more than one course. Overall, however, the report reveals that the total score for University of Toronto students on this assessment of academic integrity was a disappointing 43%, which is well below the provincial average of 72% (Note: Higher scores indicate less misconduct).

When contacted by CBC to ask for his reaction to the contents of the report, Minister of Training, Colleges and Universities John Milloy revealed that government officials are considering calling for cuts to government funding for undergraduate academic programs at the University of Toronto, a penalty aimed at raising undergraduate tuition fees for students at the school. “We had no idea of how UofT students are performing when it comes to academic conduct and we probably never would have known if UofT’s academic integrity report hadn’t been leaked. “We simply can’t reward this type of academic dishonesty among UofT students,” Milloy noted. We can’t send the message to students at other schools that they can be as unethical in the classroom as are students at UofT.

Furthermore, in response to UofT’s poor performance on the academic integrity report, such employers as IBM and the Government of Canada, which hire a large number of UofT students following graduation, have discussed the possibility of reducing the number UofT students that they intend to recruit for jobs in the future. Marilyn Coxwell, a member of the IBM Board of Directors, noted that “High rates of academic misconduct among students decrease the value of a University of Toronto degree. If students are cheating and plagiarizing, it means that we can’t be confident that they are really learning the skills that their degrees would typically guarantee.”

Aside from such penalties imposed by the government and potential employers, some individuals have speculated that the unfavourable report may also influence the university's standing in world university rankings and its reputation among prominent donors. Indeed, at a time when donors are becoming increasingly conscious of the ethical integrity of the institutions they support financially, this unflattering report of UofT students may reduce the willingness of donors to support fundraising initiatives aimed at improving undergraduate programs and resources at the University of Toronto.

Study 1b: Neutral Disclosure Condition

Environmental Report Leaked to Public Reveals How “Green” University of Toronto Students Are

A confidential report recently leaked to CBC reveals that University of Toronto students are not any more environmentally unfriendly than are students at other Ontario universities.

The report was commissioned by the Ontario Federation for University Affairs (OFUA), a group of student and faculty representatives from universities across Ontario who have been elected to oversee a wide variety of issues at schools across the country, such as program development, student life, and academic performance. The goal of the OFUA is to allow universities to work together to strengthen post-secondary education institutions in Ontario. Recently, in response to mounting scientific evidence of climate change and concerns about the potential social and economic consequences that people living in Ontario will face in the coming decades, the OFUA has implemented a new program designed to encourage environmentally friendly behaviour among university students. This program requires universities to implement pro-environmental initiatives on their campuses and monitor the environmental progress of their students.

University officials decided to have the OFUA manage the assessments at all universities in Ontario because it would allow all schools to be evaluated using the same standards. It would also make it easier for administrative officials at each university to compare their school's performance to the average performance of all schools in Ontario. The reports, which are commissioned by the OFUA, are conducted by the independent research firm Ecofys. Each final report summarizing a school's performance is strictly confidential and disclosed only to administrative officials at that university and members of the OFUA. The reports are used solely for internal evaluation purposes and are not released to other universities or to the public.

Earlier this week, CBC obtained a copy of the University of Toronto's report from Andrew Baker, a fourth-year student at the University of Toronto [Queen's University]. As an OFUA student member, Baker was able to gain access to the report through the OFUA's secure electronic database. Although OFUA members are not permitted to share the content of any OFUA report with other individuals, in a statement to CBC, Baker emphasized the critical need to leak the University of Toronto's report, noting that "UofT is a public institution, so the public should be given some reassurance that UofT students are as likely as students at other schools to behave sustainably."

An inspection of the report reveals that UofT students don't perform any more environmentally irresponsible behaviours than do students at most universities in Ontario:

In comparison to students at other universities, UofT students use a similar percentage of disposable coffee cups and purchase as many bottled beverages at food outlets on campus.

Energy audits indicate that students living in residence at UofT use an average amount of non-essential energy, primarily due to the extensive use of small appliances, such as mini fridges, space heaters, and fans. But it's not just students who live on campus who are performing at this level: Polls of students who live off campus show that even those who don't live in residence consume comparable amounts of energy to that amount used by most university students in Ontario.

The assessment also examined student cafeterias and other dining facilities on campus and found that UofT students don't create any more food and packaging waste than do students at other Ontario universities and they are also not any more likely to recycle improperly. Furthermore, in comparison to students at other schools, UofT students eat a similar proportion of high-carbon footprint foods (i.e., foods that require extensive environmental resources to produce), such as red meat, dairy, and imported vegetables.

In addition, the report indicates that the rate at which UofT students participate in environmental clubs and organizations on campus is comparable to that of other Ontario universities. The same is true of the amount of environmentally friendly school supplies (e.g., recycled paper and notebooks) purchased by students at campus bookstores.

Lastly, online assessments completed by a large number of undergraduate students indicate that UofT students engage in the same degree of environmentally friendly behaviours when they are spending time off campus in comparison to when they are on campus. The report also indicates that UofT students use public transit as much as do students at other Ontario universities. Overall, the report reveals that the total score for University of Toronto students on this assessment was 73%, which is equal to the provincial average (Note: Higher scores indicate lower rates of environmentally harmful behaviours).

When contacted by CBC to ask for his reaction to the contents of the report, Minister of Training, Colleges and Universities John Milloy revealed that government officials are considering calling for changes to government funding for undergraduate academic programs at the University of Toronto, a move aimed at adjusting undergraduate tuition fees for students at the school, according to the school's environmental performance. "We had no idea how UofT students are performing when it comes to protecting the environment and we probably never would have known if UofT's environmental assessment report hadn't been leaked. It's important that we recognize universities that make an effort to protect the environment," Milloy noted. "We want to send the message to students at universities across Ontario that it's very important to protect the environment."

Aside from such responses by government officials, some individuals have speculated that the report may also influence the university's reputation among prominent donors. Indeed, at a time when donors are becoming increasingly conscious of the environmental sustainability records of the institutions they support financially, this report of UofT students may ensure that donors will be willing to support fundraising initiatives aimed at improving undergraduate programs and resources at the University of Toronto.

Study 1b: Negative Disclosure Condition

Environmental Report Leaked to Public Reveals That University of Toronto Students Are Not As "Green" As They May Think

A confidential report recently leaked to CBC reveals that University of Toronto students are not very environmentally friendly.

The report was commissioned by the Ontario Federation for University Affairs (OFUA), a group of student and faculty representatives from universities across Ontario who have been elected to oversee a wide variety of issues at schools across the country, such as program development, student life, and academic performance. The goal of the OFUA is to allow universities to work together to strengthen post-secondary education institutions in Ontario. Recently, in response to mounting scientific evidence of climate change and concerns about the potential social and economic consequences that people living in Ontario will face in the coming decades, the OFUA has implemented a new program designed to encourage environmentally friendly behaviour among university students. This program requires universities to implement pro-environmental initiatives on their campuses and monitor the environmental progress of their students.

University officials decided to have the OFUA manage the assessments at all universities in Ontario because it would allow all schools to be evaluated using the same standards. It would also make it easier for administrative officials at each university to compare their school's performance to the average performance of all schools in Ontario. The reports, which are commissioned by the OFUA, are conducted by the independent research firm Ecofys. Each final report summarizing a school's performance is strictly confidential and disclosed only to administrative officials at that university and members of the OFUA. The reports are used solely for internal evaluation purposes and are not released to other universities or to the public.

Earlier this week, CBC obtained a copy of the University of Toronto's report from Andrew Baker, a fourth-year student at the University of Toronto [Queen's University]. As an OFUA student member, Baker was able to gain access to the report through the OFUA's secure electronic database. Although OFUA members are not permitted to share the content of any OFUA report with other individuals, in a statement to CBC, Baker emphasized the critical need to leak the University of Toronto's report, noting that "UofT is a public institution, so if UofT students are performing poorly, the public should know."

An inspection of the report reveals the "poor performance" noted by Baker. Indeed, it indicates that UofT students are less environmentally friendly than are students at most universities in Ontario:

In comparison to students at other universities, UofT students use a higher percentage of disposable coffee cups and purchase more bottled beverages at food outlets on campus.

Energy audits indicate that students living in residence at UofT use a large amount of non-essential energy, primarily due to the extensive use of small appliances, such as mini fridges, space heaters, and fans. But it's not just students who live on campus who are at fault: Polls of students who live off campus show that even those who don't live in residence consume more energy than do most university students in Ontario.

The assessment also examined student cafeterias and other dining facilities on campus and found that UofT students create more food and packaging waste than do students at other Ontario universities and they are also more likely to recycle improperly. Furthermore, UofT students eat a greater proportion of high-carbon footprint foods (i.e., foods that require extensive environmental resources to produce), such as red meat, dairy, and imported vegetables.

In addition, the report indicates that participation in environmental clubs and organizations on campus is particularly low among UofT students, as is the amount of environmentally friendly school supplies (e.g., recycled paper and notebooks) purchased by students at campus bookstores.

Lastly, online assessments completed by a large number of undergraduate students indicate that UofT students are just as environmentally unfriendly when they are spending time off campus in comparison to when they are on campus. The report did indicate that UofT students perform well on the use of public transit. Overall, however, the report reveals that the total score for University of Toronto students on this assessment was a disappointing 43%.

When contacted by CBC to ask for his reaction to the contents of the report, Minister of Training, Colleges and Universities John Milloy revealed that government officials are considering calling for cuts to government funding for undergraduate academic programs at the University of Toronto, a penalty aimed at raising undergraduate tuition fees for students at the school. "We had no idea that UofT students are performing this poorly when it comes to protecting the environment and we probably never would have known about this if UofT's environmental assessment report hadn't been leaked. We simply can't reward this type of environmentally harmful behaviour among UofT students," Milloy noted. "We can't send the message to students at other schools that they can be as environmentally unfriendly as are students at UofT."

Aside from such penalties imposed by the government, some individuals have speculated that the unfavourable report may also influence the university's reputation among prominent donors.

Indeed, at a time when donors are becoming increasingly conscious of the environmental sustainability records of the institutions they support financially, this unflattering report of UofT students may reduce the willingness of donors to support fundraising initiatives aimed at improving undergraduate programs and resources at the University of Toronto.

Appendix B: Social Image Concerns Measure

You will now view a series of statements about how others may react to the report about UofT students. Using the scale provided, please indicate the extent to which you agree with each of the statements.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

1. I feel that the report about UofT students will lead other people to have a negative impression of UofT students
2. I feel that the report about UofT students will lead members of the community to view UofT students negatively
3. I feel that the report about UofT students will have a negative impact on the reputation of UofT students
4. I am concerned that the report about UofT students will reflect poorly on me because it will lead other people to have a negative image of UofT students
5. I am worried that I will be at a disadvantage because the report about UofT students will make people evaluate UofT students negatively
6. I am concerned that the report about UofT students will make others think less highly of me because it will make them have an unfavourable image of UofT students.

Appendix C: Target Evaluations Measure

Please rate Andrew on the following traits, using the scale provided.

1	2	3	4	5	6	7
Not at all characteristic of Andrew			Neither characteristic nor uncharacteristic			Very characteristic of Andrew

1. Annoying
2. Offensive
3. Arrogant
4. Selfish
5. Self-centered
6. Obnoxious
7. Traitorous
8. Greedy
9. Insulting

Appendix D: Group Improvement Concerns Measure

Please indicate the extent to which you agree or disagree with the following statements, using the scale provided.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

1. UofT students should make more of an effort to be ethical in the classroom
2. UofT students should aim to show more concern for academic integrity
3. UofT should try to improve their performance on assessments of academic integrity
4. UofT students should strive to be as honest in the classroom as students at other schools in Ontario
5. UofT students need to correct the problems raised by the report

Appendix E: Personal Moral Reproach Measure

Please respond to the following question from the perspective of Andrew Baker:

If Andrew Baker knew me, he would think that I am...

1	2	3	4	5	6	7
Not at all						Very moral
moral						

Appendix F: Group Moral Reproach Measure

Please respond to the following question from the perspective of Andrew Baker:

Andrew Baker thinks that UofT students are...

1	2	3	4	5	6	7
Not at all						Very moral
moral						

Appendix G: Personal Moral Integrity Measure

On the following screens, you will see questions about how you view yourself. Please indicate the extent to which you agree with each question, using the scale provided.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

1. Due to the report about UofT students' academic behaviour, I think that I am predisposed to cheat and plagiarize at school
2. When I think of the academic behaviour of UofT students, I think that I am academically dishonest
3. Reading the academic integrity report makes me realize that I am an unethical person

Appendix H: Group Moral Integrity Measure

On the following screens, you will see questions about how you view your group. Please indicate the extent to which you agree with each question, using the scale provided.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

1. As a UofT student, I feel guilty that UofT students are not behaving ethically at school
2. As a UofT student, I feel regret for the fact that UofT students are cheating and plagiarizing
3. As a UofT student, I feel bad that UofT students are not meeting standards for academic integrity.
4. As a UofT student, I can easily feel guilty about the poor academic integrity of UofT students
5. I feel guilt when I receive information indicating that UofT students are cheating and plagiarizing

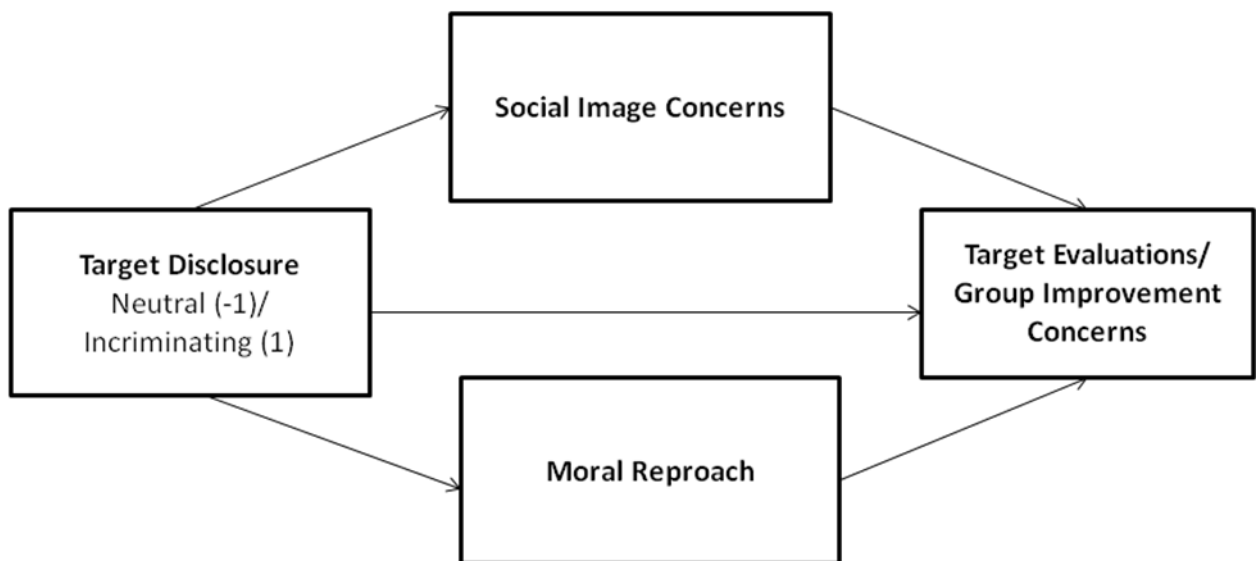


Figure 1. Mediation models tested in Study 1 to assess the indirect effects of target disclosure on both target evaluations and group improvement concerns via (a) social image concerns and (b) moral reproach.

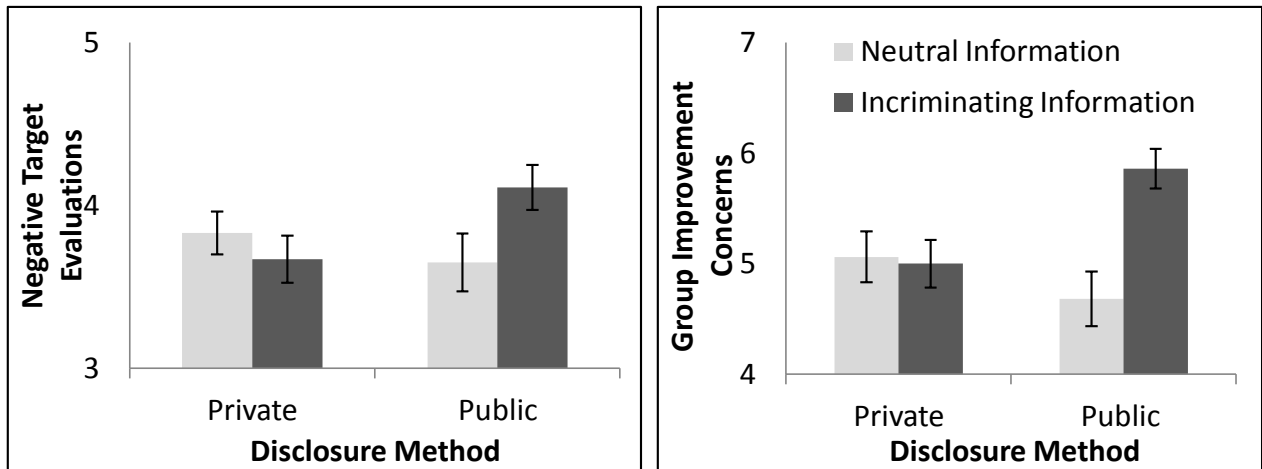


Figure 2. Study 6 target evaluations (left panel) and group improvement concern (right panel) as a function of target disclosure and disclosure method.