A Threat-Emotion Profile Approach to Explaining Active versus Passive Harm in Intergroup Relations

Brian M. Johnston[1]

Demis E. Glasford[12]

[1]Graduate Center, City University of New York

[2]John Jay College, City University of New York

Word count

Abstract: 106

Main body text (excluding figure, references, & footnotes): 4669

Accepted for publication at the *Journal of Social Psychology*. Please do not cite or distribute without permission of first author.

Author Note

Correspondence concerning this article should be addressed to Brian M. Johnston, Department of Psychology, CUNY Graduate Center, 365 5th Avenue, Room 6304.01, New York, NY 10016. Email: bjohnston1@gc.cuny.edu

Abstract

Research on the sociofunctional threat approach (STA) illustrates that people have distinct emotional reactions to different forms of threat from outgroups, such that there are distinct threat-emotion profiles. Drawing on emotion-appraisal theory, the present research investigated whether three threat-emotion profiles (obstacle-anger, contamination-disgust, and safety-fear) would be differentially related to active versus passive harm. In two studies, participants were randomly assigned to evaluate different outgroups and completed threat, emotion, and harm measures. Whereas the obstacle-anger profile was more likely to be associated with active, but not passive, harm, contamination-disgust and safety-fear were more likely to be associated with passive harm. Implications for prejudice and prejudice-reduction are discussed.

Keywords: threat, emotion, prejudice, active harm, passive harm

A Threat-Emotion Profile Approach to Explaining Active versus Passive Harm in Intergroup Relations

Threat and emotion are integral to understanding intergroup prejudice. Indeed, a sociofunctional threat approach (STA) to prejudice suggests that qualitatively distinct intergroup threats can lead to functionally unique emotional responses (Cottrell & Neuberg, 2005). At the heart of this approach is the notion of a threat-emotion profile, which suggests that particular threats elicit specific emotional responses (Neuberg, Smith, & Asher, 2000). Three common threat-emotion profiles are obstacle-anger, contamination-disgust, and safety-fear (Cottrell & Neuberg, 2005). STA research finds, for example, that to the extent that a person perceives an obstacle threat (e.g., job loss), they experience anger, but to the extent that a person feels a contamination threat (e.g., to cultural values), they experience disgust (Cottrell & Neuberg, 2005). Finally, to the extent that a person perceives a safety threat (e.g., to physical well-being), they experience fear. Despite evidence of unique threat-emotion profiles (Cottrell & Neuberg, 2005), relatively few studies have investigated whether different threat-emotion profiles are more strongly associated with specific forms of bias toward the group perceived as a threatening. Drawing on the STA, the present research investigated the extent to which obstacle-anger, contamination-disgust, and safety-fear profiles were differentially related to active versus passive harm biases.

Although there are several forms of bias, the present research is primarily concerned with active versus passive harm. Recent work suggests that one way that bias can be differentiated is based on whether it is *active* (e.g., harassment) or *passive* (e.g., exclusion; Cuddy, Fiske, & Glick, 2007). For example, host citizens may become violent and actively harm immigrants, whereas heterosexuals may support exclusionary policies and passively harm LGBT individuals. Active and passive harm have been shown to be unique forms of prejudice that are explained by independent and diverse predictors (Cuddy, Fiske, & Glick, 2008). Importantly, research demonstrates that distinct emotions often differentially explain active and passive harm (Fiske, Cuddy, & Glick, 2007). For instance, whereas intergroup envy

explains active, but not passive, harm, pity explains passive, but not active, harm (Cuddy et al., 2007). Past work, however, has not yet examined whether threat-emotion profiles are differentially associated with active versus passive harm.

From an STA perspective, prejudice, such as active and passive harm, is rooted in the psychological experience of intergroup threat, which may stem from outgroup stereotypes or the intergroup context. More specifically, STA research suggests that specific threats are associated with unique emotions. For example, past STA work illustrates that among a population of white college students, based on stereotypes, active feminists were associated with obstacle threat, gay men were associated with contamination threat, and Mexican-Americans were associated with safety threat (Cottrell & Neuberg, 2005). Importantly, from an STA perspective, as long as a person experiences a threat, they should experience the unique emotion linked to that particular threat. Thus, across groups, to the extent that a person experienced a specific threat (irrespective of the particular stereotypes associated with the target group) they also experienced the emotion linked to the respective threat; those that experienced obstacle, contamination, and safety threat were more likely to experience anger, disgust, and fear, respectively (Cottrell & Neuberg, 2005). Therefore, there is evidence that specific threats often lead to distinct emotions. Much of the research on the STA, however, has focused primarily on linking threats to emotions (Cottrell & Neuberg, 2005), with some work investigating their ability to predict policy attitudes (Cottrell, Richards, & Nichols, 2010). We examined the extent to which three threat-emotion profiles might be differentially related to two forms of prejudice: active versus passive harm.

In seeking to understand which threat-emotion profiles should lead to active versus passive harm we draw on emotion-appraisal theory, which proposes that emotions can be appetitive or aversive (Roseman, 2001). Emotion-appraisal theory classifies emotions based on the motivational states they elicit and suggests that emotions can be classified as either appetitive (emotions that produce an action-

oriented state to maximize desirable events) or aversive (emotions that produce an avoidance-oriented state to minimize undesirable events; Roseman, 2001). Whereas anger, guilt, and joy are common appetitive emotions, disgust, contempt, and distress are common aversive emotions. Anger, for example, has been shown to lead to more active-oriented actions, such as confrontation of an outgroup (Mackie, Devos, & Smith, 2000), but disgust has been shown to lead to more passive-oriented actions, such as excluding the culture of an outgroup (Matthews & Levin, 2012). This is in line with physiological research on emotions noting anger as an approach motivating emotion and disgust and fear as avoidance motivating emotions (Carver, & Harmon-Jones, 2009; Lang, Davis, & Öhman, 2000; Vrana, 1993). Drawing on an emotion-appraisal framework, it would be expected that threat-emotion profiles that include an appetitive emotion should be more likely to lead to active harm, but threat-emotion profiles that include an aversive emotion should be more likely to lead to passive harm. Therefore, we hypothesized that three threat-emotion profiles, obstacle-anger (active harm), contamination-disgust (passive harm), and safety-fear (passive harm), would be differentially related to active versus passive harm.

First, we posit that an obstacle-anger profile should be associated with active harm. Emotion research argues that anger results from appraising stimuli as an injustice or a blockage to some desirable goal (Roseman, 2001). Within the STA, this goal blockage is the product of appraising an outgroup as an obstacle threat (e.g., active feminists threatening social coordination; Cottrell & Neuberg, 2005), leading to anger toward that outgroup. Consistent with emotion-appraisal theory, anger is an appetitive emotion, and thus often motivates attack behaviors (Mackie et al., 2000). Additionally, anger has been associated with active harm more consistently than passive harm (Cuddy et al., 2007). Finally, in line with our hypotheses, there is some evidence illustrating that anger mediates the relation between an obstacle threat and confrontation (Kamans, Otten, & Gordijn, 2011). Thus, we expected the obstacle-anger threat-emotion profile to be more strongly associated with active, rather than passive, harm.

Second, a contamination-disgust profile should be strongly associated with passive harm. Disgust, an aversive emotion (Roseman, 2001), can arise from perceived violation of cultural values or norms (Rozin, Lowery, Imada, & Haidt, 1999). Indeed, within an STA framework, disgust has been shown to be associated with perceptions of cultural contamination (e.g., the sexual orientation of gay men as a contamination of some Christian ingroup morals; Cottrell & Neuberg, 2005). This might explain why traditional sexual values, such as agreeing that sex is primarily for procreation, is associated with increased "cold" attitudes toward non-heterosexual outgroups (Herek, 2002). Finally, disgust has been shown to be associated with outgroup exclusion (Matthews & Levin, 2012). We therefore expected a contamination-disgust threat-emotion profile to be more strongly associated with passive, rather than active, harm.

Third, we expected a safety-fear profile to be associated with passive harm. Fear, as a typically aversive emotion, has been shown to elicit an escape orientation or distancing actions (Roseman, 2001). Within the STA, fear is hypothesized to occur when appraising a target outgroup to be a safety threat (e.g., perceiving Mexican Americans to be a violent group; Cottrell & Neuberg, 2005), which often leads to actions to distance from groups eliciting fear (e.g., segregation from conflict in Northern Ireland; Hughes, Campbell, Hewstone, Cairns, 2008). Similarly, fear has been shown to mediate the relation between physical threat and avoidance (Kamans et al., 2011). Thus, the safety-fear profile was expected to be more strongly associated with passive, rather than active, harm.

Drawing on theory and research on threats, emotions, and prejudice, the present work explores the extent to which threat-emotion profiles are related to two different forms of bias: active versus passive harm. Past work examining threats has primarily focused on the emotions threats elicit (Cottrell & Neuberg, 2005) and research on active and passive harm has predominantly focused on the antecedents to these two forms of bias (Cuddy et al., 2008). The present research expands on past prejudice research work by examining how particular threat-emotion profiles may be associated with

specific forms of bias. Drawing on emotion-appraisal theory (Roseman, 2001), we argue that threat-emotion profiles should be especially likely to lead to active or passive harm based on the motivational nature (i.e., appetitive versus aversive) of the emotion involved in the profile. Taken together, this work therefore builds not only on the threat-emotion literature (Cottrell & Neuberg, 2005) by testing for how threat-emotion profiles lead to alternative forms of bias, but also complements and builds on previous work exploring active and passive harm (Cuddy et al., 2008) by linking threats, as antecedents, to active and passive harm.

Overview

Consistent with recent work suggesting a need for greater replication within the social sciences (Simmons, Nelson, & Simonsohn, 2011), we conducted two studies to test the same four hypotheses. More specifically, participants in both studies were randomly assigned to evaluate an outgroup; we then assessed threats (obstacle, contamination, safety), emotions (anger, disgust, fear), and biases (active and passive harm) toward the respective target outgroup. Because our hypotheses were concerned with group-specific and across-group effects, we included three target outgroups (active feminists, gay men, or Mexican-Americans) in both studies. These three groups have been shown to be associated with obstacle, contamination, and safety threat, respectively (Cottrell & Neuberg, 2005), thus allowing us to test not only a group-specific hypothesis, but also enabling a stronger, more generalizable, test of whether threat-emotion profiles lead to active versus passive harm, irrespective of stereotypes. Drawing on past STA work (Cottrell & Neuberg, 2005), we expected active feminists (obstacle-anger), gay men (contamination-disgust), and Mexican-Americans (safety-fear) to elicit qualitatively unique threat-emotion profiles (*Hypothesis 1*). However, controlling for target outgroup, we also expected threat-emotion profiles to be differentially associated with active versus passive harm. Specifically, we expected that obstacle threat would be associated with active harm, which would be explained by anger (*Hypothesis 2*). Additionally, we expected that contamination threat would be related to passive harm,

which would be explained by disgust (*Hypothesis 3*). Finally, we expected that safety threat would be related to passive harm, which would be explained by fear (*Hypothesis 4*).

Study 1

Method

Participants. Eighty-four American participants were recruited to complete an online questionnaire via Amazon's "Mechanical Turk" (MTurk) website, which increases diversity (e.g., age) compared to traditional college students, and allows participants to complete tasks for monetary compensation (Buhrmester, Kwang, & Gosling, 2011). In addition, MTurk allows researchers to restrict access to the task based on the ISP address of computer. All participants in the present work participated from within the United States. The sample was mostly white (88.1% white / Caucasian, 11.9% black / African American), had a fairly even gender split (48.8% female, 48.8% male, 1.2% other, 1.2% missing data), was mostly heterosexual (84.5% heterosexual or straight, 6.0% homosexual, gay, or lesbian, 7.1% bisexual, 2.4% other), and had a wide range in age (M = 35.12, SD = 12.25).

Procedures. Participants were first randomly assigned to a target group questionnaire. All three versions were identical with the exception of the target group (active feminists, gay men, or Mexican-Americans).[1] The questionnaire involved Likert ratings of how "Americans" perceive the threats (obstacle, contamination, and safety), emotions (anger, disgust, and fear), and biases (active and passive harm) toward the target group. To reduce social desirability, we used "Americans" for the wording of items (Fiske, Glick, & Xu, 2002). Lastly, demographic variables were measured and participants were debriefed.

Measures. For each of the following scales, participants were asked to responded on a 7-point Likert scale (1 = *not at all* and 7 = *extremely*) based on how "Americans" would respond. The scales were presented in the following order, with items presented in a random order within each scale, and coded such that higher scores represent more of each construct.

Threats. Items for obstacle (three items, e.g., "threaten social coordination and functioning," "threaten personal rights and freedom"), contamination (four items, e.g., "threaten American values," "are immoral"), and safety threat (three items, e.g., "endanger physical safety," "are harmless") were adapted from Cottrell and Neuberg (2005). A confirmatory factor analysis (CFA) was run for these items. After removing a reverse-coded obstacle threat item with a poor factor loading ("trustworthy"), a three factor CFA revealed moderate support, χ^2 (24, N = 86) = 34.10, p = .083, *RMSEA* = 0.07.[2] These items were averaged separately for obstacle (α = .74, M = 3.83, SD = 1.55), contamination (α =.89, M = 3.95, SD = 1.61), and safety (α = .62, M = 3.54, SD = 1.24) threat.

Emotions. Anger (M = 4.13, SD = 1.64), disgust (M = 4.48, SD = 1.69), and fear (M = 3.58, SD = 1.74) were assessed using single-item measures among several emotion indices (similar to Cottrell & Neuberg, 2005).

Active and passive harm. Three items each were adapted for active ("attack," "harass," "fight") and passive harm ("exclude," "demean," "ignore"; Cuddy et al., 2007). After removing one item from each scale ("fight" and "ignore"), a two factor CFA demonstrated strong factor loadings and fit the data well, χ^2 (1, N = 86) = 0.66, p = .42, *RMSEA* < 0.001. These items were averaged separately for active (α = .80, M = 3.56, SD = 1.51) and passive harm (α = .76, M = 4.14, SD = 1.42).

Independent samples t-tests did not indicate any participant demographic differences on any of the above measures (all ts < |1.5|).

Results

Threat-emotion profiles: specific target groups. Consistent with our first hypothesis, MANOVAs with target group as the independent variable revealed significant differences on both the threat, Wilks' Lamba = 0.59, F (6, 158) = 8.04, p < .001, and emotion, Wilks' Lamba = 0.79, F (6, 158) = 3.39, p = .004, outcomes. Whereas gay men, for example, were rated higher on contamination threat and disgust, Mexican-Americans were rated higher on safety threat. Although space constraints limit the full

description of these findings, we found support that specific target groups elicited unique threat-emotion profiles (Cottrell & Neuberg, 2005).[3]

Threat-emotion profiles and active versus passive harm. Our primary hypotheses (2-4) were concerned with the extent to which threat-emotion profiles would be differentially associated with active versus passive harm across groups. Thus, controlling for target group, we expected the obstacle-anger profile to be more strongly associated with active harm than passive harm (*Hypothesis 2*), but the contamination-disgust (*Hypothesis 3*) and safety-fear (*Hypothesis 4*) profiles to be more strongly associated with passive harm than active harm. To test these hypotheses, the threat predictor, three emotion mediators, bias outcome, and dummy covariates for the target group were input into mediation models (Preacher & Hayes, 2008). This was done for six models, one for each of the three threats with both active and passive harm outcomes. Emotion indirect effects were estimated within a bias corrected 95% confidence interval using 5,000 bootstrap samples. Below we report the results primarily from the three hypothesized harm models, as displayed in panels A1, B1, and C1 of Figure 1, rather than the alternate harm outcome.[4]

Obstacle threat. As expected, obstacle threat was positively associated with anger and active harm (see Panel A1 of Figure 1). Additionally, there was a significant indirect effect for anger, .19, 95% CI [.05, .38], but not fear, .09, 95% CI [-.05, .15]. However, an unexpected indirect effect for disgust also emerged, .15, 95% CI [.03, .30]. Consistent with hypotheses, the obstacle-anger model was more strongly associated with active harm, relative to passive harm, as there was no indirect effect of anger in a passive harm alternate model, .09, 95% CI [-.06, .27].

Contamination threat. Contamination threat was positively associated with disgust and passive harm as hypothesized (see Panel B1 of Figure 1). Furthermore, as expected only the indirect effect of disgust explained this relationship, .23, 95% CI [.10, .40], rather than anger, .06, 95% CI [-.04, .21], or fear, .02, 95% CI [-.03, .10]. In support of our hypotheses, there was some evidence suggesting that the

contamination-disgust profile was more strongly associated with passive harm, relative to active harm, as there was a weaker indirect effect for disgust with an active harm alternate model, .18, 95% CI [.03, .38], which was coupled with anger, .12, 95% CI [.02, .28].

Safety threat. Safety threat was positively associated with fear and passive harm (see Panel C1 of Figure 1). However, our mediation hypothesis regarding the indirect effect for fear was not supported, .04, 95% CI [-.10, .19]. Results revealed that disgust, .24, 95% CI [.10, .45], but not anger, .11, 95% CI [-.09, .31], explained this relationship. An alternate model with active harm also did not have a significant indirect effect for fear, .06, 95% CI [-.08, .25].

Study 2

Method

Participants. Using the same MTurk online recruiting procedures as Study 1, ninety-eight American participants were recruited for the Study 2 test for replication (92.9% white / Caucasian, 7.1% black / African American; 53.1% female, 44.9% male, 2% other; 88.8% heterosexual or straight, 2.0% homosexual, gay, or lesbian, 8.2% bisexual; age M = 33.19, SD = 12.27).

Procedures and measures. Identical procedures and measures as Study 1 were utilized.

Threats. The same threat item was dropped as per similar CFA results, χ^2 (24, N = 97) = 25.52, p = .378, *RMSEA* = 0.026, though with the same limitations as Study 1. Reliability estimates were a bit lower for obstacle (α = .69, M = 3.80, SD = 1.39) and contamination (α = .82, M = 3.89, SD = 1.28), but higher for safety threat (α = .71, M = 3.54, SD = 1.22).

Emotions. Anger (M = 4.22, SD = 1.38), disgust (M = 4.26, SD = 1.39), and fear (M = 3.64, SD = 1.61) had similar distributions.

Active and passive harm. The same harm items were dropped as per similar CFA results, χ^2 (1, N = 96) = 0.12, p = .728, *RMSEA* < 0.001. Active harm had roughly the same reliability (α = .83, M = 3.80, SD = 1.40), while passive harm was somewhat lower (α = .69, M = 4.52, SD = 1.25).

The only participant demographic differences on these measures were with race / ethnicity (e.g., black / African Americans expressing greater contamination). This was not adjusted for as the results below were substantively identical with or without a race / ethnicity covariate.

Results

Threat-emotion profiles: specific target groups. Providing further support to our first hypothesis, MANOVAs revealed significant differences on both the threat, Wilks' Lamba = 0.65, F (6, 184) = 7.37, $p < .001$, and emotion, Wilks' Lamba = 0.83, F (6, 186) = 3.07, $p = .007$, outcomes. Results were similar, such as Mexican-Americans being rated higher on safety threat.

Threat-emotion profiles and active versus passive harm. Panels A2, B2, and C2 of Figure 1 present the results from the mediation analyses. Bootstrap estimates were again utilized to assess indirect effects, but marginal indirect effects were also discovered (i.e., within a 90% confidence interval) and are described below (note that marginal indirect effects were not found in Study 1).

Obstacle threat. Obstacle threat was again positively associated with anger and active harm (see Panel A2 of Figure 1). The hypothesized anger indirect effect was marginal, .11, 90% CI [.01, .26], while there were not (marginal or significant) indirect effects for disgust, .08, 90% CI [-.02, .20], or fear, .06, 90% CI [-.003, .15], as expected. Consistent with Study 1, the obstacle-anger model was more strongly associated with active harm, as there was no indirect effect of anger in a passive harm alternate model, .06, 90% CI [-.03, .20].

Contamination threat. Contamination threat was found to be positively associated with disgust and passive harm again (see Panel B2 of Figure 1). There was moderate support for this model with a marginal disgust indirect effect, .12, 90% CI [.003, .26], and no indirect effect for anger, .08, 90% CI [-.04, .24]. However, there was an unexpected indirect effect for fear, .11, 95% CI [.03, .22]. Nonetheless, providing further support beyond Study 1, an active harm alternate model did not have any indirect effects for disgust, .12, 90% CI [-.01, .26].

Safety threat. As before, safety threat was positively associated with fear and passive harm (see Panel C2 of Figure 1). Furthermore, the hypothesized indirect effect for fear was significant in the present study, .13, 95% CI [.04, .27], but not for anger, .11, 90% CI [-.02, .28]. However, there was an unexpected indirect effect for disgust, .13, 95% CI [.02, .29]. An active harm alternate model provided further support though, as there was not a significant indirect effect for fear, .09 90% CI [-.01, .22].

General Discussion

Across two studies, the present work provides initial evidence for an association between threat-emotion profiles and specific forms of harm. Consistent with past work (Cottrell & Neuberg, 2005), we found that specific threat-emotion profiles (e.g., contamination-disgust) were more strongly associated with specific groups (e.g., gay men). Moreover, irrespective of target group, there was some evidence that threat-emotion profiles were also differentially related to active versus passive harm. Complementing intergroup emotions theory work (Mackie et al., 2000), to the extent that an outgroup elicited an obstacle-anger profile, participants were more likely to report active harm. Additionally, adding to previous work on disgust-exclusion (Matthews & Levin, 2012), to the extent that an outgroup elicited a contamination-disgust profile, participants were more like to report passive harm. Finally, the safety-fear profile was more strongly associated with passive harm, comparable to segregation of outgroups perceived to be violent (Hughes et al., 2008).

This research illustrates the critical interplay between threats and emotions in explaining prejudice. Whereas threats have been shown to predict emotions (Cottrell & Neuberg, 2005) and prejudice (González, Verkuyten, Weesie, & Poppe, 2008), and emotions have been shown to predict active and passive harm (Cuddy et al., 2008), the current results demonstrate the relationship among threats, emotions, and bias: specific threat-emotion profiles were differentially and more strongly related to active versus passive harm. Past STA work has focused on action tendencies (Kamans et al., 2011), but the current work illustrates the usefulness of the STA framework for explaining active and

passive harm. A threat-emotion profile framework may also help to explain different forms of sexism (e.g., hostile versus benevolent; Glick & Fiske, 2001) or racism (e.g., old-fashioned versus modern; Gawronski, Peters, Brochu, & Strack, 2008). For example, whereas old-fashioned, hostile racial hate-crimes are likely to be associated with an obstacle-anger profile, modern cultural exclusion may be more likely to be associated with a contamination-disgust profile. Finally, previous work has explored the link between threats and emotions (Cottrell & Neuberg, 2005), stereotypes, emotions, and bias (Cuddy et al., 2007), and threats, emotions, and bias (the current research), but more work should focus on the interrelations between stereotypes, threats, emotions, and forms of bias. Indeed, although not tested in the present work, it may be the case that stereotypes serve as an antecedent to threat-emotion profiles (Stephan, Ybarra, & Morrison, 2009) and subsequent bias. For example, stereotypes (e.g., beliefs that Mexican Americans are violent) could elicit specific threat-emotion profiles (e.g., safety-fear) that in turn are associated with a particular form of bias (e.g., passive harm).

A threat-emotion profile framework may also have implications for prejudice-reduction interventions. To the extent that prejudice-reduction interventions are framed around not only emotions, but also threats, interventions should be more likely to reduce prejudice. Whereas intergroup contact, for example, might be better able to reduce prejudice stemming from contamination-disgust or safety-fear, emotion regulation strategies might be better able to reduce obstacle-anger prejudice. Because contact reduces anxiety (Pettigrew & Tropp, 2008), which is often linked to disgust and fear, intergroup contact might be a useful way to reduce passive harm explained by contamination-disgust or safety-fear. Conversely, preventing or reducing active harm from an obstacle-anger profile might be best achieved by using outgroup knowledge (Hill & Augoustinos, 2001) as a means of cognitively reappraising the outgroup obstacle perception to down-regulate anger (Gross, 2008). Thus, prejudice-reduction interventions may benefit from a more nuanced and comprehensive approach rooted in a threat-emotion framework.

It should be noted that we are not suggesting that threat-emotion profiles will necessarily always lead to the hypothesized forms of bias. Indeed, it is very likely that situational or group-based factors may moderate the hypothesized relations between threat-emotion profiles and active versus passive harm. For example, it is very likely that under certain situational conditions (e.g., instability within intergroup relations), such as severe conflict or war (Letendre, Fincher, & Thornhill, 2010), one threat-emotion profile (e.g., contamination-disgust) would be associated to the alternative form of bias (i.e., active harm), than the one hypothesized in the present work (i.e., passive harm). Similarly, group power within an intergroup context may moderate preference to engage in active versus passive harm. Group members with little or no power may be less likely to engage in active harm in response to an obstacle-anger profile, based on their perceived ability to take action (Galinsky, Gruenfeld, & Magee, 2003). Emotion-appraisal theory should therefore be viewed as a framework which helps to explain when a threat-emotion profile is *likely* to lead to active versus passive harm, but acknowledge there are certainly situational or group-based factors that moderate the relation between threats, emotions, and preference for active versus passive harm.

Across the two studies, there were some inconsistent findings. First, we expected disgust and fear to independently explain their respective threat-emotion profiles (contamination-disgust; safety-fear), but both in Study 1 (the safety-fear model was only explained by disgust) and in Study 2 (the contamination-disgust model was also explained by fear; the safety-fear model was also explained by disgust) fear and disgust were correlated. These findings complement past work, which finds that fear and disgust are sometimes associated with one another (Olatunji, Sawchuk, Lohr, & de Jong, 2004). Second, contrary to hypotheses, the obstacle-anger model was also explained by disgust (in Study 1, but not Study 2). Anger and disgust explaining active harm is consistent with work that has demonstrated that anger and disgust are often associated with active moral judgments (Solerno & Peter-Hagene, 2013). These results suggest that when appetitive emotions (e.g., anger) and aversive emotions (e.g., disgust)

co-occur, appetitive emotions may be more likely to determine the likelihood of active, over passive, harm. Indeed, all active harm mediation results implicated anger, even when other emotions explained a portion of the indirect effects. Of course, we acknowledge that some of the results related to the threat-emotion profiles may be due to the high correlation between the threat measures or the fluctuating reliabilities of these scales across studies. Thus, to provide a stronger test of the present hypotheses, future research could develop novel measures of, or manipulate, threats.

A cautious interpretation of the findings is warranted based on some study limitations. Because minority groups have different motivations regarding the status quo (Sidanius & Pratto, 1999), it may be the case that the present results with a predominantly white majority group sample are less likely to generalize to minority groups. Future research should therefore examine the extent to which a threat-emotion profile framework generalizes to more diverse samples and intergroup contexts. In the present work, there was some asymmetry in the active and passive harm predictions. This is because we explored three STA threat-emotion profiles that had the strongest support in past work (Cottrell & Neuberg, 2005) and is not necessarily a reflection of asymmetry in threat-harm relationships in general. Additional research is also needed to investigate other threat-emotion profiles and put more emphasis on examining active harm. The present work, though, provides initial evidence in support of the hypotheses that threat-emotion profiles lead to preferences for active and passive harm.

The present findings provide preliminary support for a threat-emotion profile framework for explaining active and passive harm biases. As illustrated in examples, such as host citizens becoming violent and actively harming immigrants perceived to be stealing jobs, or heterosexuals supporting exclusionary policies and passively harming LGBT individuals perceived to be a threat to cultural values, the threat-emotion experience of perpetrators may help to explain a person's likelihood of engaging in active or passive harm. A more nuanced threat-emotion framework that explicitly makes connections

among threats, emotions, and forms of bias may help lead to more effective prejudice reduction strategies.

References

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3-5.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135*(2), 183-204.

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reaction to different groups: A sociofunctional threat-based approach to "prejudice". *Journal of Personality and Social Psychology, 88(5),* 770-789.

Cottrell, C. A., Richards, D. A. R., & Nichols, A. L. (2010). Predicting policy attitudes from general prejudice versus specific intergroup emotions. *Journal of Experimental Social Psychology, 46*(2), 247-254.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92*(4), 631-648.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Competence and warmth as universal dimensions of social perception: The stereotype content model and the BIAS map. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (vol. 40, pp. 61-149). New York, NY: Academic Press.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology, 85*(3), 453-466.

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin, 34*(5), 648-665.

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109-118.

González, K. V., Verkuyten, M., Weesie, J., & Poppe, E. (2008). Prejudice towards Muslims in the Netherlands: Testing integrated threat theory. *British Journal of Social Psychology, 47*, 667-685.

Gross, J. J. (2008). Emotion regulation. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 497-512). New York, NY: Guilford Press.

Herek, G. M. (2002). Heterosexuals' attitudes toward bisexual men and women in the United States. *The Journal of Sex Research, 39*(4), 264-274.

Hill, M. E., & Augoustinos, M. (2001). Stereotype change and prejudice reduction: Short- and long-term evaluation of a cross-cultural awareness programme. *Journal of Community and Applied Social Psychology, 11*(4), 243-262.

Hughes, J., Campbell, A., Hewstone, M., & Cairns, E. (2008). "What's there to fear?" – A comparative study of responses to the out-group in mixed and segregated areas of Belfast. *Peace and Change, 33*(4), 522-548.

Kamans, E., Otten, S., & Gordijn, E. H. (2011). Power and threat in intergroup conflict: How emotional and behavioral responses depend on amount and content of threat. *Group Processes and Intergroup Relations, 14*(3), 293-310.

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders, 61*, 137-159.

Letendre, K., Fincher, C. L., & Thornhill, R. (2010). Does infectious disease cause global variation in the frequency of intrastate armed conflict and civil war? *Biological Reviews, 85*, 669-683.

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79*(4), 602-616.

Matthews, M., & Levin, S. (2012). Testing a dual process model of prejudice: Assessment of group threat perceptions and emotions. *Motivation and Emotion, 36*(4), 564-574.

Neuberg, S. L., Smith, D. M., & Asher, T. (2000). Why people stigmatize: Toward a biocultural framework. In T. Heatherton, R. Kleck, M. Hebl, & J. Hull (Eds.), *The social psychology of stigma* (pp. 31-61). New York, NY: Guilford Press.

Olatunji, B. O., Sawchuk, C. N., Lohr, J. M., & de Jong, P. J. (2004). Disgust domains in the prediction of contamination fear. *Behaviour Research and Therapy, 42*, 93-104.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology, 38*(6), 922-934.

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.

Roseman, I. J. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 68-91). London: London University Press.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between the three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology, 76*(4), 574-586.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359-1366.

Solerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage and judgments. *Psychological Science*. doi: 10.1177/0956797613486988

Stephan, W. G., Ybarra, O., & Morrison, K. R. (2009). Intergroup threat theory. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 43-60). New York, NY: Psychology Press.

Vrana, S. R. (1993). The physiology of disgust: Differentiating negative emotional contexts with facial EMG. *Psychophysiology, 30*, 279-286.

Footnotes

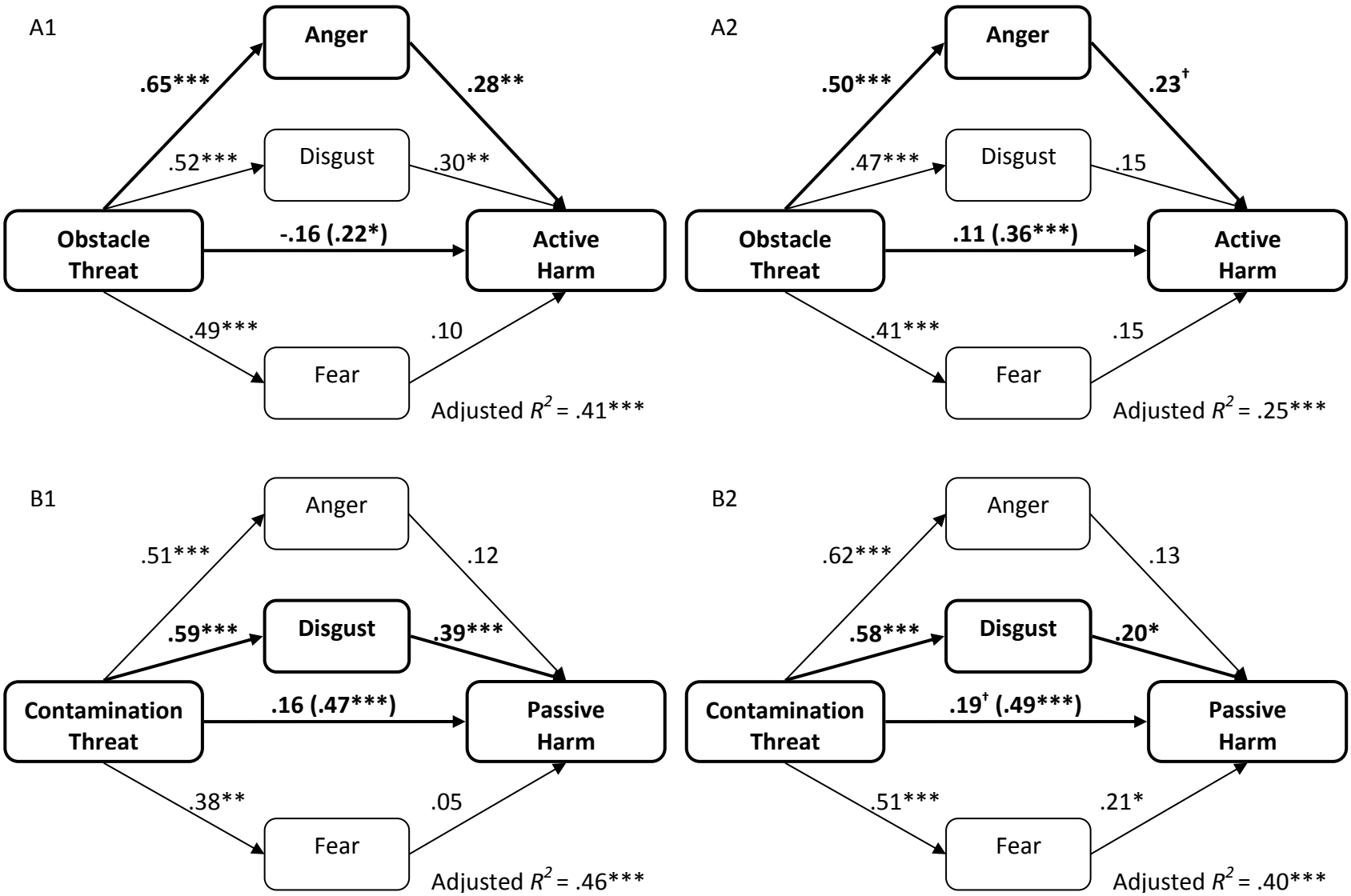
[1] In Study 1, based on demographic data we are certain with the Mexican-American questionnaire that all participants were rating an outgroup and with the gay men questionnaire that all but three participants were rating an outgroup (analyses with and without these participants are nearly identical and therefore were included in all analyses). Demographic items did not address whether or not participants self-identified as an active feminist, making it a possibility that those with an active feminist questionnaire were responding toward an ingroup. Nonetheless, there were no gender differences discovered for any of the measures, which provides some indirect support of these participants rating an outgroup.

[2] This three factor threat CFA was limited as there was a not positive definite error in the phi matrix, resulting in possible inaccurate fit statistics and factor loadings. It would appear that this error is due to high correlations between the latent threat variables ($\varphi_{Obstacle,\ Contamination}$ = .98, $\varphi_{Obstacle,\ Safety}$ = .96, $\varphi_{Contamination,\ Safety}$ = .72), rather than an issue of negative or zero error variances (all $\vartheta_{\delta ii} > 0.27$). Nonetheless, of various models tested, this three factor model provided the best fit. For example, constraining correlations to be zero resulted in an inability to estimate standard errors and a severe lack of fit, χ^2 (27, N = 86) = 141.03, $p < .001$, *RMSEA* = 0.23. Furthermore, exploratory factor analysis (EFA) results indicated that a three factor threat EFA model fit the data well, χ^2 (12, N = 86) = 10.83, p = .54, with Eigenvalues of 4.92 (obstacle), 1.16 (contamination), and .73 (safety). Overall, the CFA limitation of high correlations between threats can only work against our hypotheses attempting to make distinctions between threat-emotion profiles. The high correlation of the threat measures and inconclusive CFA results are similar to the results of past work investigating threat-emotion profiles (Cottrell & Neuberg, 2005). The overall pattern of findings suggests a need for additional work to utilize more nuanced and novel threat measures with improved psychometric properties. For readers interested in exploring these issues further, variance-covariance matrices of data are available here

https://sites.google.com/site/psychologybrianmjohnston/johnston-glasford-2014-supplemental-materials-pdf (see Tables 1-3). Further details of these analyses and full datasets are also available from the first author upon request.

[3] Detailed MANOVA results for both Study 1 and Study 2 are available here https://sites.google.com/site/psychologybrianmjohnston/johnston-glasford-2014-supplemental-materials-pdf (see Tables 4 and 5).

[4] Space constraints limit a full explanation of the three alternate harm models; a figure of these models for both Study 1 and Study 2 is available here https://sites.google.com/site/psychologybrianmjohnston/johnston-glasford-2014-supplemental-materials-pdf (see Figure 2). Additionally, while several intergroup emotions theories (e.g., Cuddy et al., 2007) posit that outgroup perceptions (e.g., threats, stereotypes) elicit emotions that motivate bias behaviors, another potential alternate explanation is that threats predict emotions with bias as the mediator (i.e., a threat-bias profile to explaining intergroup emotions). However, additional analyses largely did not support this alternate explanation. These additional models were run and there was only one significant indirect bias effect in Study 1 (with contamination threat predicting disgust and passive harm as the mediator) and there were not any significant indirect bias effects in Study 2. In comparison to the several significant threat-emotion indirect effects reported in Study 1 and 2, the hypothesized models seem more plausible. Detailed results are available for these alternate bias mediator models from the first author upon request.


A1
Anger
Disgust
Fear
Obstacle Threat
Active Harm
.65***
.28**
.52***
.30**
-.16 (.22*)
.49***
.10
Adjusted R^2 = .41***
A2
Anger
Disgust
Fear
Obstacle Threat
Active Harm
.50***
.23†
.47***
.15
.11 (.36***)
.41***
.15
Adjusted R^2 = .25***
B1
Anger
Disgust
Fear
Contamination Threat
Passive Harm
.51***
.12
.59***
.39***
.16 (.47***)
.38**
.05
Adjusted R^2 = .46***
B2
Anger
Disgust
Fear
Contamination Threat
Passive Harm
.62***
.13
.58***
.20*
.19† (.49***)
.51***
.21*
Adjusted R^2 = .40***

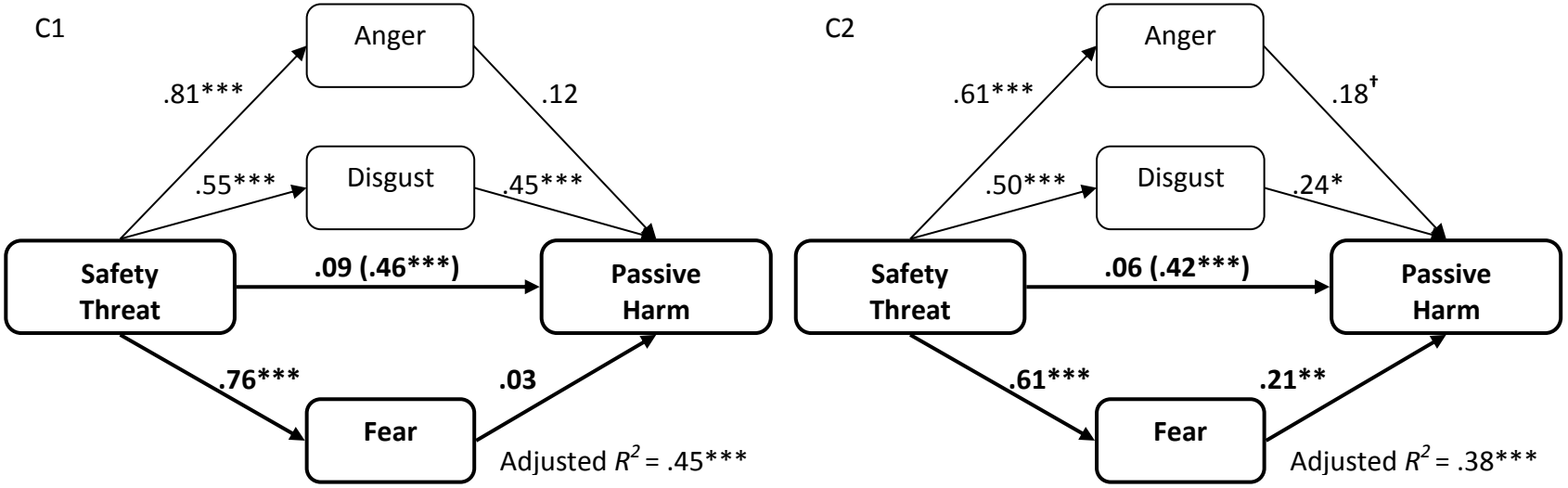


Figure 1. Hypothesized mediation models for Study 1 (panels A1, B1, and C1) and Study 2 (panels A2, B2, and C2). Coefficients are unstandardized, and multiple emotion indirect effects were estimated simultaneously within each model across three target groups (active feminists, gay men, and Mexican-Americans); groups were controlled for in all of the above estimates (covariate estimates not presented). The coefficients in parentheses represent the total effects of the threat on the outcome (without the emotions in the model), while the coefficients to the left of this represents the direct effects (with the emotions in the model). The bolded paths represent the hypothesized relationships. The adjusted R^2 refers to the overall dependent variable model (i.e., the model with the threat and emotion estimates), for which F statistics for panels A1, B1, and C1 are respectively F (6, 77) = 10.50, $p < .001$, F (6, 77) = 12.81, $p < .001$, and F (6, 77) = 12.13. $p < .001$, and panels A2, B2, and C2 are respectively F (6, 91) = 6.33, $p < .001$, F (6, 90) = 11.55, $p < .001$, and F (6, 91) = 10.81. $p < .001$.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.