Collective Harm-doing, Ingroup Identification, and Mental Health From the Perspective of the Perpetrator Group

Mengyao Li
Bernhard Leidner
University of Massachusetts Amherst, USA

Mengyao Li
Department of Psychological and Brain Sciences
University of Massachusetts Amherst
135 Hicks Way
Amherst, MA 01003 USA
E-mail: mengyao@psych.umass.edu
Tel: (+1) 917-836-4335

Author note: The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

Abstract

Collective harm-doing can have profound impact on the health and well-being of involved parties. Complementing prior research primarily focusing on the health consequences for victims of harm-doing, we examined the effects of collective harm-doing on the mental health of perpetrator group members, particularly those who were not directly involved in the harm-doing. In two experiments, we demonstrate that ingroup-committed harm has differential effects on group members’ health, depending on their identification with the ingroup (i.e., attachment and glorification). In both experiments, American participants were randomly assigned to read about the torture of Iraqi detainees in a prison either run by the U.S. or Australia. When the ingroup rather than the outgroup committed violence, low (but not high) glorifiers reported significantly more intense stress emotions, higher levels of distress, and more anxiety-related symptoms. These findings illuminate the complex role of social identity in collective harm-doing and health.

Word count: 146
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

Collective Harm-doing, Ingroup Identification, and Mental Health From the Perspective of the Perpetrator Group

From hate crimes, to armed conflicts, to ethnic cleansing and genocide, group-based harm-doing can have profound impact on the health and well-being of involved parties (Kellezi, Reicher, & Cassidy, 2009; for a review see Leidner, Li, & Kardos, 2015). One of the key features of collective harm-doing is that it can take a toll on group members who were not directly victimized by or responsible for the harm-doing (e.g., Doosje, Branscombe, Spears, & Manstead, 1998; Miller & Rasmussen, 2010). The American public, for instance, had visceral reactions to the massacres of civilians by U.S. forces in the Vietnam War, and, more recently, to the U.S. invasion of Iraq in 2003 and the systematic abuse of detainees in U.S.-run facilities in Guantanamo, Afghanistan, and Iraq. Although extensive prior work has documented the health consequences of collective harm-doing for victims and survivors, little attention has been paid to the health and well-being of members of perpetrating groups – especially those who were not in any way involved in the harm-doing (see Leidner et al., 2015).

Grounded within the social identity framework (Hornsey, 2008), the current research empirically examines the effects of intergroup harm-doing on the mental health of people who share group membership with the harm doers and the moderating role of ingroup glorification, a central dimension of recent multidimensional conceptualizations of ingroup identification (Roccas, Klar, & Liviatan, 2006). We propose that when faced with ingroup-committed (as opposed to outgroup-committed) wrongdoings, group members’ mental health will be negatively affected to the extent that they do not glorify, and hence are more critical toward, their group. In contrast, the mental health of those who unconditionally glorify their group will remain intact. In other words, glorification of the ingroup may serve as a psychological buffer against the
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

otherwise adverse effects of ingroup-committed wrongdoings on health. Although beneficial at the individual level, this proposed protective function of glorification can have quite negative consequences at the (inter-)group level, as shown elsewhere (e.g., Bilali, 2013; Leidner, 2015; Leidner, Castano, Zaiser, & Giner-Sorolla, 2010; Roccas et al., 2006). In this article, we aim to draw attention to the complex role of social identification (glorification, in particular) in perpetrator group members’ responses to collective harm-doing.

Collective Harm-doing and Health

Psychological research on health has traditionally focused on how factors within the individual or within the individual’s immediate environment can affect health (e.g., Lester, Smart, & Baum, 1994; McCrae, 1984). More recently, research has begun to explore the role of social connections such as interpersonal relationships, social networks, and group memberships in shaping people’s health and well-being (Christakis & Fowler, 2007; Jetten, Haslam, Haslam, Dingle, & Jones, 2014; Pietromonaco, Uchino, & Dunkel Schetter, 2013). In the context of intergroup conflict, it is therefore not surprising that direct and indirect experiences of collective harm-doing can have profound downstream impact on the physical and emotional well-being of group members on both sides of a conflict. The vast majority of this research has adopted a victim-centered approach, uncovering the link between victimization and mental health in conflict and post-conflict settings (e.g., Aoun et al., 2013; Al-Krenawi et al., 2011; Boltan & Betancourt, 2004; Miller & Rasmussen, 2010; Mollica et al., 1998; Moscardino, Scrimin, Capello, & Altoe, 2010). On the one hand, direct exposure to violence (e.g., physical assault, torture, the disappearance or death of loved ones) can lead to various mental health problems such as Post-Traumatic Stress Disorder (PTSD; Mollica et al., 1992). On the other hand, indirect daily stressors (e.g., poverty, social isolation, breakdown of basic services) in war-affected
regions are even more powerful than direct victimization in predicting distress and trauma (Miller & Rasmussen, 2010).

In contrast to the wealth of literature on victimization and health in the context of collective harm-doing, evidence on the link between perpetration and health is rather scarce. Moreover, perspectives on perpetrators have been largely confined to the health consequences of direct (rather than indirect) perpetration of intergroup violence (e.g., Betancourt, Agnew-Blais, Gilman, Williams, & Ellis, 2010; Danish & Antonides, 2013; Hasanovic & Pajevic, 2013; McMullen, O’Callaghan, Shannon, Black, & Eakin, 2013). Considerate efforts have been made to identify the mental health problems of U.S. military service members who returned from combat environments (Bonanno et al., 2012; Danish & Antonides, 2013). Common challenges faced by returning service members include stress, PTSD, depression, and difficulties of reintegrating into their families and communities. Other researchers examined the impact of war-related violence on mental health outcomes among child soldiers who not only are at high risk of PTSD, depression, and aggressive behavior, but also experience tremendous stress associated with social discrimination and stigma (Betancourt et al., 2010).

Considering that most of these acts of violence were committed on behalf a group or in the context of intergroup conflict, there are reasons to believe that collective violence may also affect the health of individuals who did not directly perpetrate violence but share the same group membership with those who did. While empirical research on this topic is virtually nonexistent, Leidner et al. (2015) recently offered a theoretical account of the potential health consequences of indirect involvement in collective violence. Drawing on this theoretical account, we argue that in order to understand how and why collective harm-doing influences the health of uninvolved perpetrator group members, it is crucial to consider the interplay between group membership,
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

social identity, and health. Specifically, we propose that the extent to which ingroup-committed violence negatively affects the health of ingroup members depends on their identification with the ingroup.

The Role of Group Identity in Collective Harm-doing and Health

Drawing upon social identity theory (SIT; Tajfel & Turner, 1979) and self-categorization theory (SCT; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; for an integrative review of both SCT and SIT see Hornsey, 2008), the social identity perspective on health suggests that as people derive their self-concept and self-esteem from their perceived membership in social groups, a shared identity with others can protect or harm their health and well-being (for a review see Jetten et al., 2014, and for a meta-analytic review see Steffens, Haslam, Schuh, Jetten, & Dick, 2016). The majority of this work has focused on the protective functions of social identity. According to this perspective, social identity provides a basis for cooperation, trust, security, and social support (Haslam, Jetten, Postmes, & Haslam, 2009; Khan et al., 2015; Levine, Prosser, Evans, & Reicher, 2005), which in turn have positive effects on both physical and mental health (e.g., Holt-Lunstad, Smith, & Layton, 2010; Jetten, Haslam, & Haslam, 2012).

Social identity also shapes how people appraise and respond to a wide range of life stressors. Shifts in the salience of identity, for example, can lead to changes in the way people evaluate illness and injury symptoms associated with the activated identity (Levine, 1999). Social support, one important psychological resource derived from social identity, has been shown to enable cancer patients to adopt accommodative coping strategies, which in turn result in more optimistic perceptions of cancer (Luszczynska, Mohamed, & Schwarzer, 2005). More recent research has demonstrated that social identity reduces depression by attenuating
depressive attribution style and fostering positive attribution style (Cruwys, South, Greenaway, & Haslam, 2015).

While there is considerable evidence demonstrating the health benefits of shared group identity, social groups are not always protective for health and well-being. In fact, belonging to stigmatized or socially devalued groups may provoke stressful experiences, leading to adverse psychological and physiological consequences (Allison, 1998; Clark, Anderson, Clark, & Williams, 1999; Matheson & Cole, 2004). This potential for group identity to also be harmful is particularly relevant to the case of collective harm-doing, as conflict threatens the psychological resources of both victim and perpetrator group members (Shnabel & Nadler, 2015; Shnabel, Nadler, Ullrich, Dovidio, & Carmi, 2009). Whereas members of the victimized group tend to experience threats to the power and status of the ingroup, members of the perpetrating group tend to experience threats to the moral image and reputation of the ingroup (Shnabel et al., 2009). It has been shown that membership in and identification with a victimized group can sustain and even amplify stress (Matheson & Cole, 2004; Wayment, 2004), especially when the victimized or devalued identity is salient (Major, Quinton, & McCoy, 2002). Other research, however, suggests that identification with a victimized group can sometimes protect individuals against stress, PTSD, and other psychological disorders (Kellezi & Reicher, 2011; Wohl & Van Bavel, 2011), especially when the victimization experiences are perceived as affirming rather than compromising group members’ identity (Kellezi et al., 2009). Given these mixed findings, it seems that vicarious victimization affects the health of group members differently depending on how they make sense of the experience(s) in relation to their social identity.

Much like the threat to the ingroup’s power and status, the threat to the ingroup’s moral standing should also affect group members’ health and well-being. Despite the lack of direct
evidence on the link between perpetrator group identity and health, research has demonstrated that shame and guilt – common emotional reactions to moral identity threats (Lickel, Schmader, & Barquissau, 2004) – have negative effects on both mental and physical health (e.g., Dickerson, Kemeny, Aziz, Kim, & Fahey, 2004; Scheff, 2001; Tangney, 1995; Quiles & Bybee, 1997). In the Israeli context, Bizman, Yinon, and Krotman (2001) demonstrate that people are emotionally distressed when they perceive that their group deviates from the ideal group image. However, not all group members experience compunction or distress when exposed to negative information regarding their group. On the contrary, people tend to deny (Cohen, 2013), disengage from (Bandura, 1999, 2002), or even justify and moralize (Giner-Sorolla, Leidner, & Castano, 2011; Leidner & Castano, 2012) immoral acts committed by their group against outsiders. Similar to vicarious victimization, vicarious perpetration is likely to have differential effects on group members depending on how they appraise the wrongdoing in relation to their group identity. We predicted that ingroup-committed wrongdoings should lead to negative health outcomes among group members who do not tend to deny, disengage from, or moralize ingroup-committed wrongdoings. Among group members who do tend to do so, on the other hand, these psychological defense mechanisms should provide a buffer against the adverse effects of ingroup-committed wrongdoings on health. These divergent perceptions of the ingroup and its moral infractions are associated with one particular mode of social identification: ingroup glorification.

Social Identification: Attachment and Glorification

Recent research on group processes and intergroup relations has shifted from unidimensional to bi-dimensional approaches to understanding ingroup identification, for instance distinguishing between attachment to and glorification of the ingroup (Roccas et al.,
Attachment refers to perceived importance of one’s group membership and commitment to the group, whereas glorification refers to beliefs in the superiority of the ingroup over other groups and emphasizes unconditional submission to ingroup norms and authorities. Despite the partial overlap between attachment and glorification – both tap the broad concept of identification with a group – they are distinct constructs that predict vastly different outcomes for intergroup relations (Roccas et al., 2006). In particular, there is now a fair amount of evidence demonstrating the destructive role of glorification, but not attachment, in intergroup relations. Glorification is associated with justification of ingroup-committed wrongdoings and low levels of collective guilt (Bilali, 2013; Roccas et al., 2006), reduced support for ingroup intervention in genocidal conflicts (Leidner, 2015), dehumanization of outgroup victims and decreased demands for justice (Leidner et al., 2010), a shift from endorsing violence-condemning towards violence-legitimizing moral principles (Leidner & Castano, 2012), as well as increased support for violence against previously uninvolved countries after being reminded of one’s country’s past interstate wars (Li, Leidner, Euh, & Choi, 2016). Group members who are relatively low on glorification, by contrast, are more critical toward their group and more willing to accept collective responsibility for ingroup-committed wrongdoings (Leidner, 2015; Leidner et al., 2010; Roccas et al., 2006).

Given the positive relationship between ingroup glorification and psychological defense mechanisms, we hypothesized that glorification serves a protective function for the self in terms of health and psychological wellbeing. Low glorifiers should be rather susceptible to the negative effects of collective harm-doing on mental health. High glorifiers, on the other hand, should be relatively protected from these effects, and their mental health should remain intact when confronted with negative information regarding their group and its moral infractions.
Past research regarding the link between ingroup attachment and collective harm-doing has produced somewhat ambiguous findings. On the one hand, attachment (when glorification is controlled for) allows people to criticize the immoral actions of their ingroup and positively predicts feelings of collective guilt (Roccas et al., 2006). On the other hand, attachment is unrelated to the dehumanization of outgroup members who are victimized by the ingroup (Leidner et al., 2010). Overall, attachment seems to increase the propensity toward ingroup criticism rather than justification of wrongdoings. Thus, if anything, strongly attached group members (when glorification is controlled for) should react similarly to those who weakly glorify their group, experiencing more distress and other mental health issues when confronted with the ingroup’s morally questionable actions. People who are not attached or committed to their group, in contrast, should be rather unaffected by the ingroup’s moral infractions due to their low level of concern for the ingroup to begin with. Thus, the health-related responses of weakly attached group members might appear to resemble those of high glorifiers, albeit for very different reasons. As the empirical evidence is less consistent regarding the role of attachment in intergroup relations compared to that of glorification, our predictions and analyses focused on glorification as a moderator of the effects of collective harm-doing on perpetrator group members’ mental health, while accounting for attachment (if not having strong expectations about its effects).

To examine the link between glorification (and attachment) and mental health in the context of collective harm-doing, we conducted two experiments in which we examined Americans’ emotional reactions and self-reported health status when exposed to violence either committed by the U.S. (ingroup) or an unrelated outgroup. We hypothesized that ingroup- rather
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

than outgroup-committed violence would negatively affect American participants’ mental health among those who do not glorify the U.S., but not among those who do.

Study 1

Method

Participants. We recruited 270 participants through Amazon Mechanical Turk (MTurk). Three participants were excluded because they were not born in the U.S., six because they were not native English speakers. When people participate in research online (rather than in the laboratory or other in-person settings), they tend to be less attentive to study materials (Goodman, Cryder, & Cheema, 2013). As it is imperative that participants attended to and understood the manipulation materials (an alleged news article), we adopted a screening procedure to ensure data quality (e.g., for employment of similar screening procedures, see Leidner, 2015 and Li et al., 2016; for an overview/discussion of online data screening procedures, see Chandler, Mueller, & Paolacci, 2014). As a result, 13 participants were excluded because they did not correctly identify the victims’ and perpetrators’ group membership as described in the manipulation materials (indicating that they had not paid sufficient attention to the materials), nine because they spent significantly longer time reading the manipulation materials or completing the entire survey (indicating that they were interrupted or distracted during the study; univariate outliers, see Tabachnick & Fidell, 2004), and another eight because they spent less than 30 seconds reading the manipulation materials (again indicating that they did not pay sufficient attention to the materials). In addition, we also excluded eight participants who had personal ties to the victim group, and therefore may not view them as the outgroup. Our data screening thus resulted in a sample of 223 American adults (130 female, age $M = 37.82$, $SD = 13.23$, range = 19-74) who were included in the subsequent data analyses. The percentage of the
sample excluded from data analysis (17.41%) was within range of the benchmarks for online studies (3-31%), and close to the average of 15% (Chandler et al., 2014).

**Procedure.** Upon consenting to participate in the study, participants read a fictitious, but allegedly real, news report on systematic torture of Iraqi detainees in a prison in Iraq. Participants were informed that the prison was either run by the U.S. army (ingroup-committed violence condition) or by the Australian army (outgroup-committed violence condition). The articles were identical across both conditions except for the national identity of the perpetrators. After the reading task, participants completed manipulation check questions and summarized the news article in their own words. They then filled out the following dependent measures in the order outlined below. Unless noted otherwise, all responses were measured on visual analog scales from 1 (strongly disagree) to 9 (strongly agree).

**Materials.**

*Mood/affect.* On the 20-item version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), participants indicated how intensely they felt each of the twenty emotions at the moment, from 1 (not at all) to 9 (very intensely). Lazarus and his colleagues have defined stress in terms of a subset of emotions, including guilt and shame (e.g., Folkman & Lazarus, 1985; Lazarus, 1993). According to this perspective, emotions are highly diagnostic of stress and how well individuals cope with stress (Folkman & Lazarus, 1980; 1985). Based on past research on emotional distress as an indicator of poor mental health and the effectiveness of coping (e.g., Epping-Jordan et al., 1999; Faller et al., 2013), as well as the mental health consequences of shame and guilt (e.g., Tangney, 1995), of particular interest to this research were the stress-related emotions such as *distressed, upset, guilty, ashamed.*
**Ingroup attachment and glorification.** Adapted from Roccas et al. (2006), attachment was measured with eight statements tapping the importance of the U.S. to participants’ identity (e.g., “Being American is an important part of my identity.”) and their commitment to the U.S. (e.g., “I am strongly committed to my nation.”). Glorification was measured with eight statements tapping participants’ belief in the superiority of the U.S. over other countries (e.g., “The U.S. is better than other nations in all respects.”) and their deference to American authorities (e.g., “There is generally a good reason for every rule and regulation made by our national authorities.”).

**Results**

**Ingroup attachment and glorification.** Neither attachment ($\alpha = .94, M = 6.35, SD = 1.99$) nor glorification ($\alpha = .89, M = 4.56, SD = 1.65$) was affected by condition, $F_s(1, 221) < .15, p_s > .700$, thus allowing us to use them, together with condition, as continuous independent variables (IVs) in the subsequent general linear models (GLMs) carried out in SAS 9.4. To this end, attachment and glorification were centered (Aiken & West, 1991; Cohen, Cohen, West, & Aiken, 2003).

**Mood/affect.** An exploratory factor analysis (EFA) with oblique rotation established a 4-factor solution: *positive affect* (e.g., excited, enthusiastic, proud; $\alpha = .87, M = 4.04, SD = 1.63$), *arousal* (attentive, alert, interested; $\alpha = .82, M = 6.69, SD = 1.45$), *unease* (scared, afraid, nervous, jittery; $\alpha = .91, M = 2.67, SD = 1.75$), and *stress emotions* (distressed, upset, guilty, hostile, irritable, ashamed; $\alpha = .89, M = 3.86, SD = 1.99$). The resulting four composite scores were entered as dependent variables (DVs) into a GLM with condition as the categorical IV, and attachment and glorification (centered) as continuous moderators.
**Positive affect.** The analysis with positive affect as the DV revealed a significant main effect of glorification, $F(1, 215) = 12.13, p < .001, \beta = .56, \eta^2_p = .05$. The three-way interaction of condition by glorification and attachment was also significant, $F(1, 215) = 4.51, p = .035, \eta^2_p = .02$. However, none of the simple effects underlying this three-way interaction was significant, $t(215) < 1.35, ps > .180$. No other effects reached significance, $Fs < 2.01, ps > .158, \eta^2_ps < .01$.

**Arousal.** The analysis revealed a marginally significant main effect of glorification, $F(1, 215) = 3.80, p = .053, \beta = .15, \eta^2_p = .02$, and a significant main effect of attachment, $F(1, 215) = 16.56, p < .001, \beta = -.28, \eta^2_p = .07$. There was also a significant two-way interaction of attachment by glorification, $F(1, 215) = 19.71, p < .001, \eta^2_p = .08$. Analyses of simple slopes revealed that when attachment was low, glorification was negatively associated with arousal, $\beta = -.65, p < .001$, and this negative relationship disappeared when attachment was high, $\beta = .09, p = .536$. No other effects reached significance, $Fs < 1.00, ps > .300, \eta^2_ps < .01$.

**Unease.** The analysis with unease as the DV did not yield any significant effect, $Fs < 1.50, ps > .200, \eta^2_ps < .01$.

**Stress emotions.** The analysis yielded a significant two-way interactions between condition and glorification (Figure 1), $F(1, 215) = 3.99, p = .047, \eta^2_p = .01$. Disentangling this interaction, low glorifiers (-1 SD below the mean) reported significantly more intense stress emotions in the ingroup- ($M = 4.62$) than in the outgroup-committed violence condition ($M = 3.62$), $t(215) = 2.00, p = .047$, whereas high glorifiers (+1 SD above the mean) did not report significantly more or less intense stress emotions depending on condition ($M_{\text{ingroup}} = 3.21$, $M_{\text{outgroup}} = 3.86$), $t(215) = -1.19, p = .234$. The interaction between condition and attachment was marginally significant, $F(1, 215) = 3.56, p = .060, \eta^2_p = .02$. Analyses of simple effects revealed that strongly attached participants reported significantly more intense stress emotions in the
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

ingroup- ($M = 4.33$) than the outgroup-committed violence condition ($M = 3.35$), $t(215) = 2.05$, $p = .042$, whereas weakly attached participants did not report significantly more or less intense stress emotions depending on condition ($M_{\text{ingroup}} = 3.48$, $M_{\text{outgroup}} = 4.14$), $t(215) = -1.11$, $p = .269$. No other effects reached significance, $F$s < 2.06, $ps > .155$, $\eta_p^2$s < .01.

**Discussion**

Study 1 provided preliminary evidence that collective harm-doing can have an adverse effect on the mental health, in terms of stress emotions, of perpetrator group members who are low on glorification. High glorifiers, by contrast, did not show differential levels of stress depending on whether the ingroup or an unrelated outgroup committed atrocities. These findings suggest that glorifying identification may be protective for the psychological well-being of group members who did not directly participate in the harm-doing. Additionally, we also obtained a marginally significant interaction between attachment and condition, suggesting that the emotional reactions of highly attached group members were similar to those of low glorifiers. This “mirroring effect” of attachment has emerged in other intergroup research as well (e.g., Leidner et al., 2010; Leidner, 2015; Li et al., 2016), and is consistent with the theorization of attachment as a more ingroup-critical mode of identification (Roccas et al., 2006). Given that we only measured stress-related emotions as a proxy for psychological well-being, it was important to extend the current findings to other more direct markers of mental health. We thus included additional measures of stress and anxiety in Study 2.

**Study 2**

The main goal of Study 2 was to replicate and extend the findings of Study 1 with more elaborative measures of mental health and well-being. In addition to using the 60-item PANAS-X instead of the 20-item PANAS, we also measured multiple dimensions of subjective state-level
stress using the Short State Stress Questionnaire (SSSQ; Helton, 2004) and the Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983), as well as subjective state-level anxiety using the State-Trait Anxiety Inventory (STAI; Spielberger, 1985).

Method

Participants. We recruited 295 participants through MTurk. Following the same data screening procedure as in Study 1, 13 participants were excluded because they were not born in the U.S., one because he/she was not a native English speaker, 18 because they did not correctly identify the victim and perpetrator group membership as described in the manipulation materials, seven because they spent significantly more time reading the manipulation materials or completing the entire survey, and ten because they spent less than 30 seconds reading the manipulation materials. The data from the remaining 246 American adults (151 female, age \( M = 36.72, SD = 12.49, \text{range} = 18-70 \)) were included in the subsequent data analyses.

Procedure. The procedure was identical to Study 1. The measures of ingroup attachment and glorification were also identical. The following measures were added and/or modified from the previous study. Unless noted otherwise, all responses were measured on visual analog scales from 1 (strongly disagree) to 9 (strongly agree).

Materials.

Mood/affect. We used the 60-item PANAS-X (Watson & Clark, 1999) to measure the mood/affect that participants experienced “at this moment.” Again, the “stress emotions” (Lazarus, 1993) were of particular interest to the current research.

Short State Stress Questionnaire (SSSQ). Developed by Helton (2004), the SSSQ assesses the multiple dimensions of subjective state response in stressful environments. Following Helton, three dimensions of subjective stress were measured: engagement (e.g., “I feel
confident about my abilities”, “Generally, I feel in control of things”), distress (e.g., “I feel dissatisfied”, “I feel angry”, “I feel depressed”), and worry (e.g., “I am worried about what other people think of me”, “I feel self-conscious”). Of particular interest was the distress subscale.

**Perceived Stress Scale (PSS).** The PSS was designed to measure the degree to which individuals appraise life situations as stressful, and has been shown to be correlated with depressive and physical symptomatology, utilization of health services, social anxiety, among other health-related outcomes (Cohen et al., 1983). Participants completed the 10-item version of the PSS (PSS-10; Cohen & Williamson, 1988; Taylor, 2015). Participants indicated to what extent they perceived themselves to be stressed, from 1 (no, absolutely not) to 9 (yes, absolutely). All items were anchored to the present moment (e.g., “Right now, do you find that…”, followed by the item, e.g., “… you are upset?”, “… you are unable to control the important things in your life?”).

**State-Trait Anxiety Inventory (STAI).** The STAI (Spielberger, 1985) is a self-report measure of anxiety-related symptoms. It is widely used by clinical researchers and has been evaluated in different clinical populations. Participants responded to the 20-item STAI Form Y-1, indicating how intensely they felt specific emotions (e.g., calm, secure, tense, strained) at the moment. An EFA with oblique rotation indicated a two-factor solution with positive (e.g., secure, satisfied, comfortable) and negative feelings (e.g., tense, strained, confused), respectively. Due to our interest in anxiety (rather than the lack thereof), the negative factor was of particular importance to the current research.

**Results**
Ingroup attachment and glorification. Again, neither attachment nor glorification was affected by condition, $F_{s}(1, 244) < .80, ps > .370$, thus allowing us to use them as moderators in subsequent analyses.

Mood/affect. To assure comparability to Study 1, we created the same four composite scores from the PANAS items as in Study 1. However, since Study 2 administered the long version of the PANAS, we also created composite scores in accordance with the established PANAS-X subscales (Watson & Clark, 1999). All of these were subjected to the same analysis as in Study 1.

PANAS-X subscales. Among the ten subscales, four (hostility, sadness, attentiveness, serenity) were affected by the interaction between glorification and condition. Of particular relevance to the current research, low glorifiers reported more hostility – possibly toward the ingroup (rather than outgroup), given their ingroup-critical tendencies – and more sadness in the ingroup- than outgroup-committed violence condition, whereas high glorifiers did not differ significantly depending on condition. Moreover, high glorifiers were more attentive in the ingroup- than outgroup-committed violence condition, whereas low glorifiers did not differ significantly depending on condition. Detailed results of each PANAS-X subscale are reported in the Supplementary Materials.

Positive affect. As in Study 1, the three-way interaction of condition by attachment and glorification was significant, $F(1, 238) = 6.36, p = .012, \eta^2_p = .03$. Yet, again as in Study 1, none of the simple effects underlying this interaction was significant, $t_{s}(238) < 1.60, ps > .100$. No other effects reached significance, $F_{s}(1, 238) < 2.04, ps > .150, \eta^2_p s < .01$.

Arousal. Only the main effect of attachment was significant, $F(1, 238) = 5.90, p = .016, \beta = .35, \eta^2_p = .02$. All other effects were not, $F_{s}(1, 238) < 2.60, ps > .100, \eta^2_p s < .01$. 
Unease. Only the main effects of attachment, $F(1, 238) = 7.67, p = .006, \beta = -.44, \eta^2_p = .03$, and glorification, $F(1, 238) = 5.44, p = .021, \beta = .36, \eta^2_p = .02$, were significant. All other effects were not, $Fs(1, 238) < 2.00, ps > .150, \eta^2_ps < .01$.

Stress emotions. The main effect of condition was marginally significant, $F(1, 238) = 2.80, p = .096, \eta^2_p = .01$, with participants reporting somewhat more intense stress emotions in the ingroup- ($M = 3.65$) than in the outgroup-committed violence condition ($M = 3.18$). As predicted, the main effect of condition was qualified by a significant interaction with glorification (Figure 2), $F(1, 238) = 5.20, p = .023, \eta^2_p = .02$. As in Study 1, low glorifiers (-1 SD below the mean) reported significantly more intense stress emotions in the ingroup- ($M = 3.98$) than in the outgroup-committed violence condition ($M = 2.73$), $t(238) = 3.02, p = .003$, whereas high glorifiers (+1 SD above the mean) did not report significantly more or less intense stress emotions depending on condition ($M_{\text{ingroup}} = 3.53, M_{\text{outgroup}} = 3.82$), $t(238) = -.61, p = .542$. The main effect of attachment was also significant, $F(1, 238) = 4.29, p = .039, \beta = -.37, \eta^2_p = .02$. No other effects reached significance, $Fs(1, 238) < 2.00, ps > .150, \eta^2_ps < .01$.

SSSQ distress. As expected, the analysis with SSSQ distress ($\alpha = .88, M = 3.51, SD = 1.73$) as the DV revealed a significant interaction of condition by glorification (Figure 3), $F(1, 238) = 7.49, p = .007, \eta^2_p = .03$. Low glorifiers reported significantly more distress in the ingroup- ($M = 3.92$) than in the outgroup-committed violence condition ($M = 2.84$), $t(238) = 2.91, p = .004$, whereas high glorifiers did not differ significantly depending on condition ($M_{\text{ingroup}} = 3.57, M_{\text{outgroup}} = 4.15$), $t(238) = -1.36, p = .175$. There was also a significant interaction between condition and attachment, $F(1, 238) = 3.95, p = .048, \eta^2_p = .02$. Strongly attached participants reported significantly more distress in the ingroup- ($M = 3.55$) than in the outgroup-committed violence condition ($M = 2.66$), $t(238) = 2.42, p = .063$, whereas weakly
attached participants did not report significantly more or less distress either way ($M_{\text{ingroup}} = 3.95$, $M_{\text{outgroup}} = 4.33$), $t(238) = -0.84$, $p = .400$. The main effect of attachment was also significant, $F(1, 238) = 10.47$, $p = .001$, $\beta = -.51$, $\eta^2_p = .04$. No other effects reached significance, $F_s(1, 238) < 2.60$, $p_s > .100$, $\eta^2_p s < .02$.

**SSSQ engagement.** Only the main effect of attachment on SSSQ engagement ($\alpha = .85$, $M = 6.50$, $SD = 1.30$) was significant, $F(1, 238) = 12.69$, $p < .001$, $\beta = .42$.

**SSSQ worry.** There was a marginally significant main effect of attachment on SSSQ worry ($\alpha = .83$, $M = 3.66$, $SD = 1.74$), $F(1, 238) = 3.33$, $p = .069$, $\beta = -.30$. No other effects reached significance, $F_s(1, 238) < 1.76$, $p_s > .180$, $\eta^2_p s < .01$.

**STAI negative feelings.** As predicted, the analysis with STAI negative feelings ($\alpha = .90$, $M = 3.20$, $SD = 1.62$) as the DV yielded a significant interaction between condition and glorification (Figure 4), $F(1, 238) = 4.43$, $p = .036$, $\eta^2_p = .02$. Low glorifiers reported significantly more symptoms indicative of anxiety in the ingroup- ($M = 3.60$) than in the outgroup-committed violence condition ($M = 2.68$), $t(238) = 2.61$, $p = .010$, whereas high glorifiers did not differ significantly depending on condition ($M_{\text{ingroup}} = 3.38$, $M_{\text{outgroup}} = 3.67$), $t(238) = -0.72$, $p = .470$. The main effect of attachment was also significant, $F(1, 238) = 6.72$, $p = .010$, $\beta = -.39$, $\eta^2_p = .03$. No other effects reached significance, $F_s(1, 238) < 2.40$, $p_s > .100$, $\eta^2_p s < .01$.

**Perceived stress.** The analysis with perceived stress ($\alpha = .91$, $M = 4.17$, $SD = 1.67$) as the DV again yielded a significant interaction between condition and glorification (Figure 5), $F(1, 238) = 7.54$, $p = .007$, $\eta^2_p = .03$. Low glorifiers perceived themselves to be significantly more

---

1 Because STAI positive feelings were of secondary interest to the current research, the results were reported in the Supplementary Materials. In a nutshell, high glorifiers reported significantly more positive feelings in the ingroup-than in the outgroup-committed violence condition, whereas low glorifiers did not differ depending on condition.
stressed in the ingroup- \((M = 4.70)\) than in the outgroup-committed violence condition \((M = 3.89)\), \(t(238) = 2.24, p = .026\), whereas high glorifiers perceived themselves to be marginally significantly less stressed in the ingroup- \((M = 3.79)\) than in the outgroup-committed violence condition \((M = 4.59)\), \(t(238) = -1.96, p = .051\). The interaction between condition and attachment was also significant, \(F(1, 238) = 3.92, p = .049, \eta^2_p = .02\). Strongly attached participants reported somewhat more stress in the ingroup- \((M = 4.28)\) than in the outgroup-committed violence condition \((M = 3.67)\), \(t(238) = 1.72, p = .087\), whereas weakly attached participants did not \((M_{\text{ingroup}} = 4.20, M_{\text{outgroup}} = 4.81), t(238) = -1.40, p = .163\). No other effects reached significance, \(Fs(1, 238) < 0.30, ps > .600, \eta^2_p s < .01\).

**Discussion**

Using various measures of stress and anxiety, Study 2 provided additional evidence for the hypothesis that ingroup-committed harm can have a negative effect on perpetrator group members’ mental health and that this effect depends on ingroup identification – glorification in particular. Replicating and extending Study 1’s main finding, Study 2 demonstrated that low (but not high) glorifiers experienced more intense stress emotions, higher levels of distress (measured by SSSQ and PSS), and more anxiety-related symptoms (measured by STAI) when confronted with ingroup- rather than outgroup-committed violence. It is worth noting that high glorifiers reported more attentiveness in the ingroup- than outgroup-committed violence condition. This finding suggests that the observed differences between low and high glorifiers’ health-related reactions were unlikely due to high glorifiers’ lack of, or low glorifiers’ heightened, attention to ingroup-committed violence. Glorification, thus, again served as a buffer against the adverse effects of ingroup’s wrongdoing on group members’ mental health. For two of the four DVs of interest (SSSQ distress and perceived stress by PSS), the responses of strongly attached
individuals resembled those of low glorifiers, again consistent with Study 1. While the moderating effect of attachment appeared to be less robust across DVs compared to that of glorification – supporting our decision to focus on glorification – this effect is in line with previous research regarding the role of attachment in intergroup conflict (Leidner et al., 2010; Li et al., 2016).

**General Discussion**

In two studies we sought to explore the mental health consequences of collective harm-doing for perpetrator group members who were not directly involved in the harm-doing. When faced with immoral acts (i.e., prisoner abuse) taken by the U.S. (as opposed to Australia) against Iraqi detainees, American participants exhibited divergent health-related responses depending on their modes of ingroup identification. In Study 1, participants who only weakly glorified the U.S. reported more intense stress emotions when the U.S. (rather than Australia), was responsible for the wrongdoing, whereas participants who strongly glorified the U.S. did not. Study 2 replicated the finding of Study 1 and extended it using additional health-related measures including the SSSQ, STAI, and PSS. We therefore provided convergent evidence that collective harm-doing has negative implications for the mental health of perpetrator group members who do not glorify their group. Attachment also played a moderating, albeit less central, role in the effects of ingroup-committed harm-doing on group members’ mental health, such that highly (but not weakly) attached group members tended to experience more stress when confronted with ingroup- than outgroup-committed wrongdoings. Our findings suggest that low glorifiers, compared to high glorifiers, are more prone to experiencing mental health problems. In other words, glorification may indeed serve a buffering function, protecting group members from the adverse effects of ingroup’s wrongdoings on health. Complementing prior research that has
focused almost exclusively on the health of victims and survivors of group-based violence, the current work thus serves as the first empirical step toward an understanding of how people respond to moral infractions committed by their own group in terms of health and psychological well-being.

**Social Identity Protects and Harms Health**

Consistent with existing theorizing of the link between social identity and health (e.g., Jetten et al., 2014; Kellezi & Reicher, 2011; Kellezi et al., 2009), the present research demonstrates that social identification can both protect and harm the mental health of group members. Whereas past work has focused largely on social connectedness and the availability of social support as either benefiting or, when lacking, compromising health (Jetten et al., 2014; Kellezi & Reicher, 2011), we investigated different modes of social identification and their implications for mental health. When exposed to the ingroup’s immoral acts, the divergent reactions of high and low glorifiers suggest that high levels of glorification may motivate defense mechanisms that protect group members’ mental health, whereas low levels of glorification can lead to worse mental health outcomes, at least in the short term as assessed in our experiments. Attachment, on the other hand, appears to predispose group members to experiencing more stress when faced with ingroup- rather than outgroup-committed wrongdoings. However, the moderating effects of attachment should be interpreted with caution, as they were not entirely consistent across the dependent variables.

Consistent with traditional health models such as the transactional model of stress (Lazarus & Folkman, 1984), we conceptualized social identification as a *moderator* in the effects of collective harm-doing on mental health. Other models of identity and health (e.g. the integrated social identity model of stress; Haslam, 2004; Haslam & Reicher, 2006), however,
conceptualize social identification as a mediator of the effects of life stressors on health. From the perspective of the integrated social identity model of stress (ISIS), for example, negative experiences of a stressor can be transformed into positive ones through increased shared identity and social support (Haslam & Reicher, 2006). Although we did not find direct empirical support for the latter approach as neither glorification nor attachment was affected by our manipulation, at a theoretical level we do not exclude the possibility that stressors caused by ingroup-committed wrongdoings can also shift the degree and/or quality of ingroup identification — a shift that can have downstream implications for health (see also Leidner et al., 2015).

**Glorification as a Double-Edged Sword**

At the individual level, ingroup glorification appears to be quite beneficial as it protects people from identity threat posed by the ingroup’s moral infractions. Individuals with low levels of glorification, as we show, are vulnerable to the negative effects of ingroup-committed wrongdoings on mental health and well-being. However, this by no means suggests that glorification should be re-considered the more “desirable” mode of identification, compared to (critical) attachment. The protective functions of glorification, on the contrary, can be rather detrimental at the intergroup level. As discussed earlier, there is now overwhelming evidence on the links between glorification and hostile intergroup attitudes and behaviors (Bilali, 2013; Leidner & Castano, 2012; Leidner et al., 2010; Leidner, 2015; Li et al., 2016; Roccas et al., 2006). In light of the findings of the current research, it seems plausible that high glorifiers are reluctant to restore justice or mend their relationships with outgroup victims partly due to a lack of emotional and psychological burden created by ingroup-committed wrongdoings. Although low glorifiers’ mental health is negatively affected by the ingroup’s immorality, feelings of stress and anxiety may motivate them to engage in non-defensive coping strategies such as ingroup
dissent and ingroup-critical collective action that aims to improve the moral image of their group (Leidner et al., 2015). Glorification is thus best understood as a double-edged sword from the perspective of the perpetrator group – while it is beneficial at the individual level, it is harmful at the (inter)group level.

**Limitations and Future Directions**

Several limitations of the current research should also be noted. First, we focused primarily on self-reported stress and anxiety as indicators of mental health and psychological well-being. While the stress- and anxiety-related measures employed in the two studies have been widely used in both clinical and nonclinical settings to assess mental health (Cohen et al., 1983; Helton, 2004; Spielberger, 1985), it remains unclear whether the current findings will generalize to other common mental health outcomes such as depression. Future research may also assess the physiological markers of stress and anxiety in addition to using self-report measures. Second, we measured participants’ immediate health-related responses to information about ingroup- versus outgroup-committed violence. In the context of intergroup conflict, however, people are often repeatedly exposed to negative portrayals of the ingroup both by its opponents and the international community. It would thus be important to examine the long-term effects of ingroup-committed wrongdoings on the mental health and well-being of group members. It is plausible that prolonged exposure to negative information regarding the ingroup will not only adversely affect group members’ mental health, but also shift their identification with the ingroup. Last but not least, the current research does not examine the underlying mechanisms of low and high glorifiers’ divergent responses to ingroup- (rather than outgroup-) committed violence. As mentioned earlier, the self-protective function of glorification may stem from the various defense mechanisms associated with glorification (e.g., the dehumanization of
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

outgroup victims; Leidner et al., 2010). Therefore, more research is needed to uncover the psychological processes underlying the link between glorification and mental health among members of the perpetrator group.

**Conclusion**

The present studies suggest that collective harm-doing has adverse effects on the health and well-being of perpetrator group members, and these effects are attenuated by ingroup glorification. Whereas low levels of glorification increase group members’ propensity toward experiencing stress and anxiety in response to the ingroup’s wrongdoings, high levels of glorification protect group members’ health. Bearing in mind the destructive role of glorification in intergroup relations, the current contribution illuminates the duality of glorification: as much as it is a cure at the individual level, it is a curse at the (inter)group level.
COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

References


http://doi.org/10.1111/ajop.12054


http://doi.org/10.1037/0022-3514.75.4.872


COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH


COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH

http://doi.org/10.1111/j.1464-0597.2008.00379.x

http://doi.org/10.1177/154193120404801107


COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH


COLLECTIVE HARM-DOING, IDENTIFICATION, AND HEALTH


http://doi.org/10.1177/0146167204271651


http://doi.org/10.1348/135910799168470


http://ir.uiowa.edu/cgi/viewcontent.cgi?article=1011&context=psychology_pubs.


http://doi.org/10.1177/0146167203261892

*Figure 1.* Stress emotions as a function of condition (ingroup- vs. outgroup-committed violence) and glorification, while taking into account the moderating effect of attachment (Study 1).
Figure 2. Stress emotions as a function of condition (ingroup- vs. outgroup-committed violence) and glorification, while taking into account the moderating effect of attachment (Study 2).
Figure 3. Distress (measured by SSSQ) as a function of condition (ingroup- vs. outgroup-committed violence) and glorification, while taking into account the moderating effect of attachment (Study 2).
Figure 4. Anxiety (measured by STAI negative feelings) as a function of condition (ingroup- vs. outgroup-committed violence) and glorification, while taking into account the moderating effect of attachment (Study 2).
Figure 5. Perceived stress (measured by PSS) as a function of condition (ingroup- vs. outgroup-committed violence) and glorification, while taking into account the moderating effect of attachment (Study 2).