

Hate Speech: Asian American Students' Justice Judgments and Psychological Responses

Robert J. Boeckmann*

Flinders University

Jeffrey Liew

Arizona State University

Two experiments using Asian American university student participants examined the distinctive characteristics of responses to racist hate speech relative to responses to other forms of offense. The studies varied the target of insulting speech (Asian, African, and Overweight person) or the nature of offense (petty theft vs. insulting speech). Participant variables included collective self-esteem and social identification. Results indicate that hate speech directed at ethnic targets deserves more severe punishment than other forms of offensive speech and petty theft. Hate speech also results in more extreme emotional responses and, in the case of an Asian target, has a depressing influence on collective self-esteem. Ethnic identification moderated punishment responses in study 1 only. The theoretical and practical implications of the results are discussed.

Intergroup relations have a history of tension and strife that predates society's efforts to document it. The many forms of prejudice have been an active area of research in the social sciences for decades, and efforts to understand their origins and to identify solutions to the seemingly intractable problems they create has produced an enormous body of theory and research (Brewer, 1994). Most of this research has focused on the perpetrators of prejudice and discrimination (Swim & Stangor, 1998). The research reported here is related to this tradition of research, but it is distinctive in two respects. First, it does not investigate the origins of prejudice

*Correspondence concerning this article should be addressed to Robert Boeckmann, School of Psychology, The Flinders University of South Australia, GPO Box 2100, Adelaide 5001, Australia [e-mail: Robert.Boeckmann@Flinders.edu.au]. We would like to thank Marlene McGovern for her contributions to measurement and design issues for study 2 and data collection for the same.

or the factors that influence its expression. Instead it focuses on one particular expression of prejudice—hate speech—and some of the issues surrounding legal sanction of this behavior. Second, this research focuses on the target of prejudice and discrimination rather than the perpetrators of these attitudes and behaviors. In particular, by presenting participants with a variety of offensive stimuli, we will investigate the psychological impact hate speech has on Asian American university students and how they judge perpetrators of hate speech relative to perpetrators of other nonnormative behaviors.

Hate crime is an old problem with a relatively new legal status (Jeness & Grattet, 1996; see also Levin, this issue). In essence, hate crime legislation recognizes that some crimes are motivated by intergroup hostilities and embodies the idea that there is sufficient societal consensus to augment the penalties assessed for such crimes. Often the expression of hateful sentiment is key to determining whether a crime is considered a hate crime and whether such augmenting is warranted (Gerstenfeld, 1992). Further, in some instances and jurisdictions the simple expression of hateful sentiments alone is considered criminally liable. Hate crime legislation is controversial for a variety of reasons, not the least of which is that it can often conflict with a core value of American democracy enshrined in the First Amendment: the freedom of speech.

Sanctioning communications with hateful intent (hate speech) is the focal point where this conflict is most clearly seen. Although the value of freedom of expression is important, it must be weighed against the right of individuals to be free from harm or from deprivations of liberty and to have equal protection of these rights. These values and rights are also embodied in the U.S. Constitution. Thus, there is a tension inherent in the Constitution concerning these values and therefore some basis for qualifying an absolute right to freedom of expression (see Cowan, Resendez, Marshall, & Quist, this issue; Nielsen, this issue). The First Amendment does allow for the censure of some forms of speech that are deemed sufficiently harmful (Lawrence, 1993). If hate speech is to be regarded as a special form of speech unworthy of First Amendment protection and worthy of censure, research is needed to understand how it is to be distinguished from other forms of offensive speech. The first aim of the research reported here is to make this distinction.

Arguably, the majority of criminal legislation is targeted at sanctioning behavior that creates some tangible harm in the form of physical or financial damage. Sanctions against hate speech are also controversial because the harm associated with it is less tangible than the harm associated with most other crimes. Hate speech is defined as “words that are used as weapons to ambush, terrorize, wound, humiliate, and degrade” (Lawrence, Matsuda, Delgado, & Crenshaw, 1993, p. 1). Hate speech is directed at injuring a person’s social identity and thus, even when the affront is successful, it is difficult to assess its impact. In addition, psychological harm and harm to status are often taken up as civil issues rather than as criminal

issues (Lawrence, 1993). The second primary aim of the present research then is to examine the extent to which hate speech compares to crime with more tangible material consequences.

In summary, the present studies are focused on understanding the harm of hate speech through the eyes of its intended targets and examining their retributive justice responses to hate speech. In particular we are interested in knowing how minorities respond to insults directed at their minority identity relative to other forms of insulting speech and how responses to hate speech differ from responses to antisocial behavior with more tangible consequences. Although our general research aim is to learn about how minorities respond to hate speech, our research efforts are focused on a convenience sample of Asian American university students. We examine three general types of responses made by these students. First, we determine whether they view hate speech as having a broader harmful impact on society than other forms of insult or property crime. Second, we determine whether they view hate speech as worthy of legal sanction and the relative severity of sanctions recommended for hate speech. Third, we examine their self-reported emotional responses to second-hand observations of hate speech. We now turn to reviewing theory and research relevant to making specific predictions for study 1 and study 2.

Protecting a Valued Identity

Social Identity Theory (Tajfel & Turner, 1986) argues that an important determinant of self-definition and self-worth is our cognitive and affective connection to particular social categories, and there is abundant research support for these basic assertions (Hogg & Abrams, 1988; Tajfel & Turner, 1986). Once a person has categorized herself as a member of a social group, she bases an aspect of her self-esteem (collective self-esteem) on the evaluative characteristics of that group (Luhtanen & Crocker, 1992). Through this process the positive features or negative features of the relevant group may be assimilated and contribute to the valence of collective self-esteem. Social Identity Theory also argues that people are motivated to maintain or enhance their impressions of the positive features of their group that provide collective self esteem. As a result people engage in cognitive and behavioral strategies that have an impact on intergroup perceptions, attitudes and behavior (Branscombe, Ellemers, Spears, & Doosje, 1999).

Social identity is most likely to become salient and actively influence behavior in contexts that emphasize group boundaries (Turner, 1999). Obviously as a stimulus that asserts group differences, hate speech is likely to make social identity salient. In addition, hate speech clearly is a threat to the positive features of a group and thus a threat to a potentially important aspect of self for group members. Thus, the degree of importance a person places on the threatened identity can be expected to moderate responses to hate speech. Specifically we hypothesize that

minorities who are strongly identified with their minority group will respond in a more extreme fashion to hate speech directed at their group than minorities who are less invested in their group identity. Recent evidence provides some mixed support for this assertion.

Hate Speech Research

In a series of studies Leets and Giles (1977) examined respondents' reactions to mild and severe hate speech stimuli directed at Asian Americans. Their samples were composed of Caucasian and Asian university students. Asian students were assumed to be more invested in the Asian identity than Caucasian students. Leets and Giles's primary hypotheses were as follows:

- Severe hate speech would be seen as more harmful than mild.
- Ingroup members (Asian participants) would attribute greater harm to the target of hate speech than outgroup members.
- Ingroup members would attribute higher speaker liability and be more supportive of victim compensation than outgroup members.

Leets and Giles's measure of harm contained 18 items that included assessment of harm to individuals, group, and society as well as 15 additional items assessing psychological and emotional distress (e.g., depressed, tense, angry). Factor analysis revealed the 33-item measure had a three-factor solution, yet an overall composite was used for analyses. Liability was assessed by asking respondents to rate the extent to which the target of hate speech was entitled to ask for compensation and whether the respondent would award compensation if he or she was serving as a jury member in a trial seeking compensation.

In the main, Leets and Giles's (1997) findings supported their hypotheses, however, there was one important unanticipated result. There was no main effect to indicate that Asians found hate speech more harmful than Caucasians. Instead, a marginally significant interaction ($p < .06$) suggested that Caucasians rated severe hate speech as more harmful than Asians rated either mild or severe hate speech, whereas Asians rated the mild hate speech as more harmful than Caucasians did. Study 2 and study 3 replicated this result at a conceptual level by replacing the mild/severe manipulation with an indirect/direct hate speech manipulation. Despite the observation that Caucasians tend to view severe or direct Asian-focused hate speech as more harmful than Asians, questions remain. Because of the multifaceted nature of the harm measure, it is difficult to fully understand the nature of this unexpected result, and some ambiguity remains concerning what was actually found. These results are also difficult to reconcile with findings by Cowan and Hodge (1996).

Cowan and Hodge (1996) conducted a study on hate speech that presented a diverse group of participants with hate speech examples that targeted a range of social groups, including racial minorities, sexual-orientation minorities, and women. In contrast to the Leets and Giles (1997) finding that the outgroup views hate speech as more harmful than the targeted ingroup, they found that relative to women, men rated hate speech targeting women as less offensive. Similarly, Caucasian men rated hate speech directed at minorities as less offensive than did minority respondents. In contrast, Caucasian women rated all forms of hate speech as more offensive and worthy of accountability than Caucasian men. Nor did the women discriminate between the different forms of hate speech in their responses. Cowan and Hodge suggest that women may not differentiate between the different targets of hate speech because they can empathize and identify with the targets' feelings because of similar experiences in their own lives.

Social Maintenance Model of Retributive Justice

In comparing the results of Leets and Giles (1997) and Cowan and Hodge (1996), it is clear that further research needs to be conducted that clarifies how minorities respond to hate speech directed at them and at other groups. It is also important to disentangle the different types of harm that may result from hate speech and how they may be related to liability judgments. In this latter regard a novel theory of reactions to deviance may be of use. The Social Maintenance Model of Retributive Justice (SMRJ) argues that retributive justice is ultimately motivated by a concern for rules that enable social groups to function (Boeckmann, 1996). It also argues that the relationship between rule violation and retributive justice responses is mediated by an evaluation of the social relevance of the rule broken and the extent of the violation's impact on society. Violation of norms that are essential to group functioning or violations that have a broad social impact lead to more negative evaluation and will warrant more severe punishment. Preliminary evidence suggests that offense evaluation and retributive justice concerns conform to this model (Tyler, Boeckmann, Smith, & Huo, 1997). From the SMRJ perspective, hate speech may be distinguished from other forms of offensive speech by the perceived impact it has on society relative to other forms of speech. In addition, the model predicts that offenses with this type of broad social consequence will be punished more severely than offenses with a more limited social impact.

Hate speech directed at racial or ethnic categories, by definition, has implications for a whole class or category of people, not just the individual member of the class that is being addressed. In addition, the long-standing problem of racism and civil rights initiatives in the United States signifies racism and intolerance as behaviors that have broad social impact and that violate widely held norms of social equality. Thus, we hypothesize that hate speech will be viewed

as having a broader social impact and as being more punishment worthy relative to other forms of insult. Further we proposed that the perceived societal impact of an offense will mediate the relationship between an offense and recommended punishment.

Study 1

Method

Participants and Design

Fifty self-identified Asian American university student volunteers participated in the study in partial fulfillment of their introductory psychology research participation requirement. The mean age of the participants was 20.06 years ($SD = 1.42$), and 76% of participants were female. Participants were presented with scenarios using a three-level repeated measures experimental design with full counterbalancing of the order of presentation as a between-participants factor.

Stimuli

The stimulus scenarios were designed to vary the nature of insult delivered in a public setting. The nature of insult was varied to focus on individual characteristics (overweight status) or collective ascribed characteristics (ethnic status). In the later case two levels of ascribed status were presented: Asian American status and African American status. Specifically, the scenarios described several people standing in a cashier's line in a grocery store. The target of the insult delays the line by using a credit card when he realizes he doesn't have enough money for his purchases. One of the people waiting mutters a slur—"What a fucking fat asshole" (Overweight condition); "What a fucking chink" (Asian American condition); or "What a fucking nigger" (African American condition)—and then when the target of insult makes a nonverbal surprise response, the person continues by shouting a more extensive degrading insult: "I hate you! [You fat asshole you (Overweight); All you chinks (Asian American); All you niggers (African American)] are forever holding up the entire line. If you didn't bring enough cash, you shouldn't come into the store at all! Go back home!"

Measures

The extent of harm. The Societal Impact Scale (Boeckmann, 1996) was used to measure the extent to which an insult's harm was perceived as focused narrowly on the individual target or having a broader societal impact. The scale is composed of six items that were rated on a 1 (*disagree strongly*) to 6 (*agree strongly*) Likert-type scale and then averaged to produce a single score for each response to a

scenario (alphas = .84 for Overweight, .78 for Asian American, and .76 for African American). Sample items: "The harm in a case like this is widespread" and "In a case like this the harm done primarily affects specific individuals" (reverse-scored). High scores indicate perceptions of a broad social impact.

Retributive justice response. Punishment severity was measured with a single item that asked participants to rate how severely the insulting person should be punished relative to a provided reference point on a 0 (*least severe*) to 100 (*most severe*) scale. The reference point was set at 70 (*punishment for public slander in the news media*).

Social identification. Crocker and Luhtanen's Collective Self-Esteem (CSE) scale was used as a measure of participants' psychological investment in their Asian American social identity (Crocker & Luhtanen, 1990). The scale is composed of 16 items that were rated on a 1 (*completely disagree*) to 7 (*completely agree*) Likert-type scale. The scale is composed of four subscales that are in turn composed of four items each. The subscales are public collective self-esteem (alpha = .55; sample item: "In general others respect the ethnic group that I am a member of"); private collective self-esteem (alpha = .89; sample item: "I often regret that I belong to the ethnic group I do belong to"); identification importance (alpha = .86; sample item: "Overall my ethnic group has very little to do with how I feel about myself"); and membership esteem (alpha = .74; sample item: "I am a worthy member of the ethnic group I belong to"). The four subscales are interrelated, and the alpha for the entire 16 items is .88.

Procedure

The participants completed the questionnaires in small groups of 5 to 7. Participants were led to believe they were participating in a University survey of student opinion about incidents of offensive speech on campus and freedom of expression on campus. The cover story emphasized that a wide variety of opinions are expressed on these issues and that there are no clear right or wrong answers. It then stressed the importance of responding with one's honest opinion and reminded participants that the survey was anonymous. Finally, it informed the participants that they would be reading scenarios of offensive speech that were based on incidents that actually occurred to students on their campus. Participants were then randomly assigned to one of the orders of presentation and provided with a questionnaire booklet containing the three stimulus scenarios and response measures. The CSE preceded the other materials and was presented as belonging to another researcher's study. Following administration, participants were debriefed and informed that the true nature of the research was to investigate reactions to and attitudes toward hate speech.

Results and Discussion

Repeated measures analyses of variance (ANOVAs) were conducted to examine the effects of insult on social impact evaluations and punishment recommendations. Assumptions of these analyses were met, and all analyses were first examined for the presence of order effects and none were found. Consistent with predictions, participants evaluated degrading speech directed at ethnic characteristics as having significantly greater societal impact (African American $M = 4.76$, $SD = .97$; Asian American $M = 4.78$, $SD = .87$) than speech directed at an individual characteristic (Overweight $M = 4.13$, $SD = 1.09$), $F(2, 98) = 12.70$, $p < .001$. As expected, participants also recommended more severe punishment for hate speech directed at ethnicity (African American $M = 76.20$, $SD = 13.98$; Asian American $M = 76.20$, $SD = 13.54$) than for insults directed at a person's overweight status ($M = 71.00$, $SD = 16.32$), $F(2, 98) = 16.78$, $p < .001$. Both of these findings are consistent with SMRJ predictions.

The mediation hypothesis was assessed by using standard mediation testing procedures (Baron & Kenny, 1986) and following recommended modifications of them (and aspects of the SPSS v. 10 output) to account for the repeated measures design (Cohen & Cohen, 1983). Refining Baron and Kenny's arguments, Holmbeck (1997) states that four conditions must be satisfied to demonstrate mediation. First, there must be a significant relationship between the independent variable and the mediator. Second, there must be a significant relationship between the independent variable and the dependent variable. Third, there must be a significant relationship between the mediator and dependent variable. Fourth, when the mediator is controlled for, the strength of the relationship between the independent and dependent variables must be reduced.

The repeated measures ANOVA reported above demonstrating the significant effect of insult type on social impact rating satisfies the first of Holmbeck's conditions for mediation. Second, using a dummy coding procedure, vectors were created for the insult variable, and the punishment variable was regressed on them using a repeated measures regression revealing a significant relationship, $R^2 = .29$, $F(2, 98) = 20.06$, $p < .01$, consistent with the results of the repeated measures ANOVA reported previously. Third and fourth, a hierarchical repeated measures regression analyses regressed punishment on social impact in the first step and revealed a significant relationship with punishment, $R^2 = .22$, $F(1, 98) = 18.57$, $p < .01$. In the second step the dummy variables representing insult were entered and revealed that insult explained significant additional variance beyond that explained by impact, R^2 changed = .18, $F(3, 98) = 15.04$, $p < .01$. The 11% reduction in variance explained by this equation is consistent with our theoretical model, which argues that evaluations of the social impact of an offense mediate the relationship between offense stimuli and punishment recommendations. An examination with the regression weights associated with the two dummy variables revealed that the

primary difference between insult conditions was between the overweight insult condition and the ethnic insult conditions, consistent with the ANOVA results.

To examine the impact of identification on punishment responses the CSE scale score was dichotomized by using a median split procedure (CSE $M = 5.51$, $SD = .73$, median = 5.63) and a 2 (high/low identification) \times 3 (insult type) split-plot ANOVA was conducted. This analysis revealed no significant interaction between identification and insult type in predicting punishment recommendations. However, an inspection of the means for high and low identifiers revealed that the pattern was as expected, with high identifiers punishing Asian-targeted hate speech more severely than low identifiers. A post hoc analysis was pursued using the following rationale.

The member esteem subscale of the CSE assesses the extent to which a person believes he or she is a valued member of his or her social group. Consistent with Smith and Tyler (1977), we reasoned that participants with a sense that they are valued by their group would be more inclined to act on its behalf. In this context, this inclination would manifest as more harsh sanctioning responses to threats to group identity. Thus, we dichotomized the membership esteem scale ($M = 5.51$, $SD = .88$, median = 5.63) and conducted comparisons between the high membership esteem and low membership esteem groups on punishment recommendations for each type of insult. As expected, participants with high membership esteem punished Asian hate speech more severely ($M = 80.40$) than participants with low membership esteem ($M = 72.00$), $F(1, 48) = 4.87$, $p < .05$. There were no significant differences between groups for the remaining types of insult.

The results of study 1 indicate that Asian American university student respondents can and do differentiate between insults directed at ethnic characteristics relative to those directed at an individual characteristic. Our interpretation is that the participants are distinguishing between insults directed at groups or ascribed characteristics and those directed at individual or earned characteristics. This distinction is reflected in the social impact measure and punishment recommendations and is consistent with the arguments made by the SMRJ.

The mediation analysis indicates that a considerable portion of the relationship between an offense and a person's retributive justice response to it is mediated by his or her evaluation of the impact the offense has on society. This supports the SMRJ as well as suggesting one basis upon which to distinguish hate speech from other forms of insult. This pattern suggests that reactions to antinormative behavior are motivated by a desire to support the norms that serve broad social aims. The pattern could also be interpreted as suggesting that insults directed at overweight people do not qualify as hate speech.

The preceding interpretation does agree with some definitions of hate speech. Matsuda (1993, p. 36) defines hate speech as consisting of "(a) messages of racial inferiority, (b) directed at historically oppressed groups, and (c) persecutory, hateful, and degrading." However, this definition is quite narrow, and we argue that

a defining feature is that the depreciating speech is directed at a group that an individual target represents rather than directed at a particular individual. Thus we would change Matsuda's first clause to read "(a) messages of group inferiority." The critical issue then becomes whether people perceive overweight individuals as representatives of a recognized social group. It appears that for our respondents evaluating this particular set of stimuli, overweight people were not viewed in this fashion. Efforts by activists to raise consciousness about weight discrimination may eventually lead to broad social consensus that overweight people constitute a stigmatized group and that weight insults constitute hate speech.

Study 2

In study 2 we aimed to replicate and extend the study 1 findings. Specifically we set out to conceptually replicate the societal impact mediation finding. In particular we again proposed that participants' evaluation of the extent of an offense's social consequences would mediate the relationship between the offense and recommended punishment. We also aimed to replicate and extend the social identification moderation effect observed in study 1 by again measuring CSE and by also measuring social identity with Brown and colleagues' (Brown, Condor, Matthews, Wade, & Williams, 1986) group identification scale. We expected that those participants with high membership esteem or with high social identification would respond to hate speech directed at their group more markedly than those with low membership esteem or identification.

Study 2 also extended study 1 by further distinguishing hate speech from other forms of offensive behavior by comparing it to petty theft. This comparison sheds light on how minorities view hate speech relative to a broadly recognized form of criminal behavior. The SMRJ (Boeckmann, 1996) argues that rules or principles promoting social harmony are more central to group functioning than rules governing minor economic exchanges. Thus, we predict punishment responses to hate speech will be more extreme than responses to petty theft. Finally, we also examined the psychological consequences of observing hate speech. Specifically we assessed Asian American students' emotional reactions to hate speech accounts and the impact that reading these accounts had on their collective self-esteem. We predicted that hate speech directed at the respondents' own social group would have a greater psychological impact than hate speech directed at other minority groups or petty theft.

Method

Participants and Design

Fifty three self-identified Asian American university student volunteers participated in the study in partial fulfillment of their introductory psychology research

participation requirement. The mean age of the participants was 21.21 years ($SD = 2.88$), and 51% of the participants were female. Participants were presented with scenarios using a three-level repeated measures experimental design with full counterbalancing of the order of presentation as a between-participants factor.

Stimuli

The stimulus scenarios were designed to vary the nature of offense (petty theft, insult to African American, insult to Asian American) occurring in a public setting. Specifically, the scenarios described two people standing in a cashier's line in a crowded campus store. The target of the offense delays the line by using an ATM card and getting \$10 cash back from her purchases. A man waiting in line behind her becomes impatient and angry and grumbles that she should be more prepared when she shops. The target is obviously upset by this behavior and hurriedly walks off with her purchase. In the petty theft condition she forgets the \$10, and the man quickly pockets the money before anyone notices. In the hate speech conditions she does not forget her money, and before she walks off the man loudly shouts: "Fucking [Nigger/Chink]! You mongrel people are always messing things up for everyone! You are all inferior and stupid—you can't even remember to bring money to the goddamn store! You should all just go back where you came from!" The target of the hate speech is then described as humiliated and walks away with her head down.

Measures

The extent of harm. As in study 1 the perceived extent of harm was measured using the Societal Impact Scale (Boeckmann, 1996). The range in alpha coefficients across the three stimulus conditions for the measure in this study were .91 (theft), .72 (Asian American), and .71 (African American). As in study 1 high scores indicate perceptions of a broad social impact.

Retributive justice responses. Punishment judgments were measured with two items. The first item was similar to the measure used in study 1; this item asked participants to rate the severity of recommended punishment on a 0 (*least severe*) to 100 (*most severe*) scale with a concrete reference point. That point indicated that the average punishment for slander in the news media on the scale would be 30. (We shifted the anchor point down from 70 [in study 1] to 30 because we thought that if participants believed petty theft was more severe than ethnic insults that the range between 70 and 100 might be too restricted.) The second item assessed punishment responses in a different way by asking the participants to express a behavioral intention. Specifically they were asked to indicate their level of support for a university suspension policy for people who behaved the antisocial way described in the scenario by making a rating on a 1 (*very strongly*

against) to 7 (*very strongly in favor*) scale. High scores on all three items indicate a more punitive response.

Emotion measures. Emotional response to the antisocial behavior scenarios was assessed by asking participants to reread the scenario and then rate a list of feelings to “give us a sense of how you are feeling right now.” The four emotions (anger, sadness, outrage, fear) were rated on a 1 to 9 scales. For example, the item “How much anger do you feel?” was rated on a 1 (*not angry at all*) to 9 (*very angry*) scale.

Collective self-esteem and social identification. As in study 1, Crocker and Luhtanen’s (1990) CSE scale was used as a measure of participants’ psychological investment in their Asian American social identity. The reliabilities for study 2 were public collective self-esteem $\alpha = .75$, private collective self-esteem $\alpha = .84$, identification importance $\alpha = .77$, and membership esteem $\alpha = .86$. The four subscales are interrelated, and the alpha for the entire 16 items was $.87$.

In this study we also measured two relevant aspects of CSE after participants read each of the offense scenarios. The public collective self-esteem aspect of CSE measures the extent to which people think the public views their group favorably. The private collective self-esteem aspect of CSE refers to internal evaluations of the group one belongs to. We predicted our Asian American participants would suffer a reduction in public and private aspects of CSE as a response to reading accounts of others disparaging their ethnic group but be relatively unaffected by reading about insults to other groups or a petty theft.

As an alternative measure of social identity the Brown et al. (1986) group identification scale was employed in this study. The original scale is comprised of 10 items tapping awareness of group membership (e.g., “I see myself as belonging to the Asian American group”) and evaluation (e.g., “I am a person who is glad to belong to the Asian American group”). These and like items were rated on Likert-type scales and averaged into a single score ($\alpha = .87$). One of the original items—“I am a person who tries to hide belonging to the _____ group”—was omitted, as it seemed inappropriate when assessing identification with a phenotypically obvious racial group.

Procedure

The participants completed the questionnaires in small groups of 5 to 9. Participants were led to believe that they were participating in a university survey of student reactions to and attitudes toward campus crime and antisocial behavior. The cover story emphasized that a wide variety of opinions are expressed on these issues and that there are no clear right or wrong answers. It then stressed the importance of responding with one’s honest opinion and reminded participants that the survey was anonymous. Participants were then randomly assigned to one

of the orders of presentations and provided with a questionnaire booklet containing the CSE, the group identification scale, three stimulus scenarios, and the response measures. Following administration, participants were debriefed and informed that the true nature of the research was an investigation of responses to hate speech.

Results

Three-level (nature of offense) repeated measures ANOVAs were conducted to evaluate the extent of harm, punishment response, and affect hypotheses. Assumptions of these analyses were met, and all analyses were first examined for the presence of order effects and none were found. Consistent with our predictions, participants rated hate speech directed at an Asian American ($M = 5.50, SD = .71$) and at an African American ($M = 5.51, SD = .74$) as having significantly greater societal impact than petty theft ($M = 4.02, SD = 1.25$), $F(2, 104) = 73.34, p < .001$. Participants also rated hate speech as deserving significantly more severe punishment (Asian American $M = 53.40, SD = 22.01$; African American $M = 52.45, SD = 20.00$) than petty theft ($M = 36.79, SD = 17.30$), $F(2, 104) = 22.56, p < .001$. Participants were also more supportive of suspension for hate speech (Asian American $M = 5.32, SD = 1.41$; African American $M = 5.23, SD = 1.48$) than for petty theft ($M = 3.70, SD = 1.80$), $F(2, 104) = 29.05, p < .001$. Simple planned comparisons on both variables indicated there were no significant differences between the hate speech conditions and that the hate speech conditions were significantly different from the petty theft condition, as the means suggest.

The mediation hypothesis was assessed using the same modified procedures described in study 1, and the conditions for mediation were satisfied as follows: First, the repeated measures ANOVA reported above demonstrating the significant effect of offense type on social impact rating satisfies the first condition for mediation. Second, using a dummy coding procedure, vectors were created for the offense type variable, and the punishment variable was regressed on them using a repeated measures regression, revealing a significant relationship, $R^2 = .31$, $F(2, 104) = 23.13, p < .01$, consistent with the results of the repeated measures ANOVA reported previously. Third and fourth, a hierarchical repeated measures regression analysis regressed punishment on social impact in the first step and revealed a significant relationship with punishment, $R^2 = .32$, $F(1, 104) = 17.00, p < .01$. In the second step the dummy variables representing insult were entered and revealed that insult explained significant additional variance beyond that explained by impact, R^2 changed = $.07$, $F(3, 104) = 3.91, p < .05$. The 24% reduction in variance explained by this equation is evidence for nearly complete mediation and is consistent with our study 1 finding and thus supportive of the SMRJ predictions. An examination with the regression weights associated with the two dummy variables revealed that the primary difference between offense type

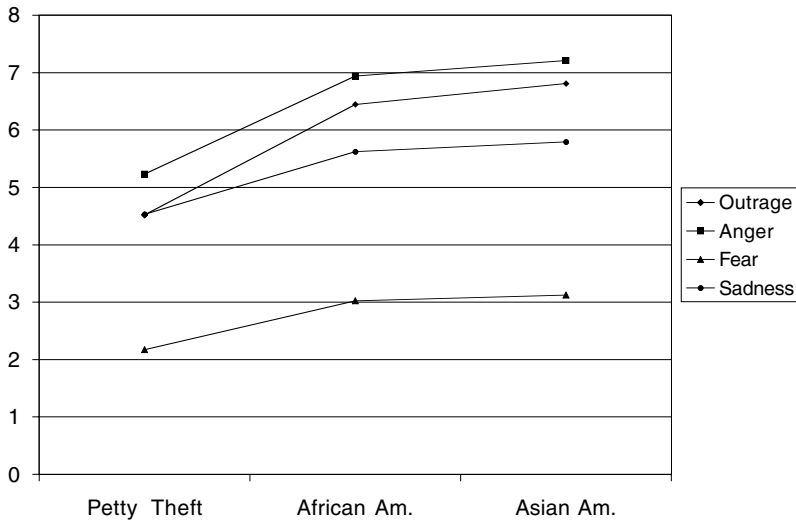


Fig. 1. Emotional reactions to hate speech.

conditions was between the petty theft condition and the ethnic insult conditions, consistent with the ANOVA results.¹

Expectations of the psychological responses to reading hate speech accounts were in the main confirmed. As predicted, emotional responses to reading the hate speech accounts were more extreme than responses following the petty theft scenario. This was true of all the emotional measures taken (e.g., outrage: $F(2, 104) = 29.31, p < .001$; anger: $F(2, 104) = 32.10, p < .001$; fear: $F(2, 104) = 10.75, p < .001$; and sadness: $F(2, 104) = 7.15, p < .001$). Repeated measures t -tests on these variables indicate that hate speech scenarios evoke significantly higher levels of all emotions relative to theft and that the hate speech scenarios do not differ from each other (see Figure 1).

The collective self-esteem reduction hypothesis was evaluated by using repeated measures t -tests to compare premeasures of the public and private CSE subscales to postmeasures of the same. Consistent with our predictions there was a significant reduction in the public aspect of CSE (premeasure $M = 5.50, SD = .91$) following exposure to the Asian American hate speech scenario (postmeasure

¹ One additional measure of punishment recommendations was used in this study. This measure simply required participants to indicate on a 1 (*least severe*) to 10 (*most severe*) scale what level of punishment should follow the offense. The pattern of results for this variable was identical to that for the 1–100 anchored punishment scale, except that the mediation analysis showed complete mediation. These analyses have not been presented in full detail because of space considerations but are available from the first author.

$M = 5.25$, $SD = 1.00$), $t(52) = 2.32$, $p < .05$. Thus our participants suffered a small but significant reduction in the aspect of collective self-esteem related to others' evaluation of their group. No significant differences were observed for comparisons involving the African American hate speech or petty theft scenarios. Examination of the pre- and postmeasures of the private CSE subscales also revealed no significant differences.

Contrary to our expectations, participants' level of social identification as measured by either the CSE scale or the Brown et al. (1986) measure did not moderate punishment responses, as found in study 1. Post hoc analyses were also conducted examining the four different aspects of CSE and no interactions were found.

Discussion

The results of the two studies support our hypotheses with one minor exception, discussed below. The results help shed light on how minorities perceive and respond to hate speech in terms of retributive justice judgments as well as in psychological and emotional terms. Two important considerations pertaining to legally restricting hate speech expression are distinguishing it from other forms of offensive speech and underscoring the less obvious harm it creates. In the first instance, study 1 clearly indicates that Asian American university students view insults directed at group characteristics as unique from insults directed at individual characteristics. The evidence suggests that one distinguishing feature is the perceived scope of harmful impact created by group-based insults. These perceptions also appear to be related in predictable ways to justice judgments. Offensive speech with broad social consequences appears to warrant more severe punishment. This is consistent with the basic assertions of the SMRJ, which argues that evaluations of the social impact of an offense mediate the relationship between offenses and the punishment recommendations (Boeckmann, 1996). It is also consistent with the general tendency observed in the justice literature for people to interpret collective issues in justice terms relative to individual issues (Tyler et al., 1997). Study 2 results are also consistent with this pattern and indicate that Asian Americans view hate speech as a more serious crime than petty theft. This may reflect social change that is both cause and consequence of legislative changes that now formally sanction racist acts.

Studies 1 and 2 both indicate that people believe hate speech has a broad social impact. What of hate speech's impact on the individual? Study 2 shows that our participants were emotionally affected by second-hand accounts of hate speech and suffered a (presumably) temporary reduction in collective self-esteem as a consequence of reading about their own group being disparaged. Direct experience with being the target of hate speech would no doubt result in more extreme and enduring consequences (but see Crocker, 1989, for qualifications of this view).

Indeed, recent field research using the Racist Schedule of Events (Landrine & Klonoff, 1996), which includes items pertaining to hate speech, indicates that the frequency and stressfulness of experiences of racism are negatively correlated with clinical measures of psychological well-being (Bailey & Boeckmann, 2001). Together the evidence on the perceived extent of harmful impact and psychological impact of second-hand exposure to hate speech adds to the ongoing debate on hate speech legislation.

The present findings also bear on theoretical issues. First, the pattern of punishment recommendations provides support for the basic proposition of the SMRJ (Boeckmann, 1996) expressed in the introduction and is also consistent with arguments by Miller and Vidmar (1981). In both cases the proponents argue that retributive justice responses are importantly determined by social concerns rather than individual or monetary concerns. These arguments and their relationship to hate speech, in particular, are noteworthy. Hate speech could be viewed as an offensive act unworthy of sanction because it ostensibly is an attack on an individual with no material or economic consequence. However, Boeckmann and Miller and Vidmar both argue that punishment responses are forthcoming so long as the nonmaterial consequences of deviance have important social implications, as is clearly the case with hate speech. In addition, critical race theorists argue that group-based insults directed at any particular individual have broad aims of oppression and a wide range of consequences that include social, material, and economic deprivation (Lawrence et al., 1993).

Second, the results of the present study could be seen as challenging some propositions of social identity theory. Social identification appears to moderate responses to hate speech, but the data are inconclusive. As expected, strength of social identification as indicated by the membership esteem aspect of the CSE-moderated responses to threats to the participants' social identity. Participants with high membership esteem reacted more strongly to threats to their group than low identifiers. This is consistent with basic assertions of social identity theory, which argues that simply being a member of a group is not sufficient to motivate group behavior. Rather, subjective identification is required to motivate group-enhancing or -protective behavior (Tajfel & Turner, 1979). However, study 2 failed to support the moderation hypothesis suggested by Social Identity Theory.

Inspection of the pattern of results from study 1 and 2 suggests one reason for the inconsistent results. We note that in both studies our participants did not discriminate between hate speech directed at Asian American targets and that directed at African American targets. This is also inconsistent with Social Identity Theory. However, these apparent inconsistencies are resolved by suggesting that making racism salient may cause participants to shift their social identity to a higher level of abstraction to simply "minority." Thus our use of a measure of identification with Asian Americans may have been too specific in this context, where the participants were inadvertently primed to think of themselves simply as

a minority. The social identity moderation hypothesis is still relevant, however, we argue that imprecision of measurement may be responsible for the inconsistent results seen here.

What do the results of the present studies mean in practical terms? Lawrence (1993) notes that the justification for infringing on first amendment protections in favor of the regulation of hate speech lies in distinguishing hate speech from other forms of speech. Our study adds to the debate for maintaining hate speech regulation where it exists and justifies extension of the policy into other jurisdictions. It does so by clearly indicating that hate speech is distinctive from other forms of insult and by evaluating it in the context of other criminalized behavior. The extent of harm results suggest that an essential strategy for overweight people is to frame their grievances about insulting behavior in group-based terms. Clearly support for retributive sanctions is linked to perceptions of group or societal impact. Sanctions for hate speech are also justified on the basis of the significant emotional and psychological effects that must be considered as an aspect of its harmful impact.

Further research is needed to address inconsistencies and shortcomings in the present studies. In particular, future research should measure social identity at an appropriate level of abstraction or utilize a between-participants design in order to better assess the social identification moderation hypothesis. To better understand emotional and esteem effects of hate, research should also be conducted in which minorities observe hate speech first-hand. Obviously ethical considerations call for careful design involving minority confederates and thorough debriefing. The use of only Asian Americans as participants in this study is one last and important limitation of the present research. Asian American students at the University of California, Berkeley, no doubt represent Asian Americans with educational and possibly financial privilege.

Research suggests that assertive reactions to injustice may be moderated by the vulnerability of the victim of injustice. Victims who are particularly vulnerable to further injury or exploitation are less likely to assert their grievances for fear of further harm (Tyler et al., 1997). The results of these studies must be generalized to other Asian Americans and other minorities with considerable caution. Young, educated Asian Americans at an elite university are considerably less vulnerable to injustice and may be more assertive in their responses to hate speech than other Asian Americans and other minority groups without similar advantages. Generalizations to nonminority groups are probably wholly inappropriate for similar reasons. Cowan and Hodge's (1996) research also makes it clear that non-minority reactions to hate speech deserve further attention. Finally, the present findings regarding extent of consequences and punishment reactions, in particular, must be replicated with a more diverse sample of participants. Despite these limitations the present research makes important contributions to a nascent body of research on responses to hate speech and suggests strategies for overweight people to employ to have verbal assaults recognized as hate speech.

References

- Bailey, M. E., & Boeckmann, R. J. (2001). *The effects of racial prejudice on psychological well being and the coping mechanisms discrimination targets employ*. Unpublished manuscript, Flinders University of South Australia.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.
- Boeckmann, R. J. (1996). *An alternative conceptual framework for offense evaluation: Implications for a social maintenance model of retributive justice*. Unpublished doctoral dissertation, University of California, Berkeley.
- Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers & R. Spears (Eds.), *Social identity: Context, commitment, content* (pp. 35–58). Oxford, UK: Blackwell Science Ltd.
- Brewer, M. (1994). The social psychology of prejudice: Getting it all together. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario symposium* (Vol. 7, pp. 315–329). Hillsdale, NJ: Lawrence Erlbaum.
- Brown, R., Condor, S., Matthews, A., Wade, G., & Williams, J. (1986). Explaining intergroup differentiation in an industrial organization. *Journal of Occupational Psychology*, *59*(4), 273–286.
- Cohen, J., & Cohen, P. (1983). *Applied MRC analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cowan, G., & Hodge, C. (1996). Judgments of hate speech: The effects of target group, publicness, and behavioral responses of the target. *Journal of Applied Social Psychology*, *26*(4), 355–374.
- Crocker, J. (1999). Social stigma and self-esteem: Situational construction of self-worth. *Journal of Experimental Social Psychology*, *35*(1), 89–107.
- Crocker, J., & Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology*, *58*(1), 60–67.
- Gerstenfeld, P. B. (1992). Smile when you call me that: The problems with punishing hate motivated behavior. *Behavioral Sciences and the Law*, *10*(2), 259–285.
- Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. Florence, KY: Taylor and Francis/Routledge.
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, *65*(4), 599–610.
- Jeness, V., & Grattet, R. (1996). The criminalization of hate: A comparison of structural and polity influences on the passage of “bias-crime” legislation in the United States. *Sociological Perspectives*, *39*(1), 129–154.
- Landrine, H., & Klonoff, E. A. (1996). The schedule of racist events: A measure of racial discrimination and a study of its negative physical and mental health consequences. *Journal of Black Psychology*, *22*(2), 144–168.
- Lawrence, C. R. III. (1993). If he hollers let him go: Regulating racist speech on campus. In M. J. Matsuda, C. R. Lawrence III, R. Delgado, & K. W. Crenshaw (Eds.), *Words that wound: Critical race theory, assaultive speech, and the First Amendment* (pp. 53–88). Boulder, CO: Westview.
- Lawrence, C. R. III, Matsuda, M. J., Delgado, R., & Crenshaw, K. W. (1993). Introduction. In M. J. Matsuda, C. R. Lawrence III, R. Delgado, & K. W. Crenshaw (Eds.), *Words that wound: Critical race theory, assaultive speech, and the First Amendment* (pp. 1–13). Boulder, CO: Westview.
- Leets, L., & Giles, H. (1997). Words as weapons: When do they wound? Investigations of harmful speech. *Human Communication Research*, *24*(2), 260–301.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one’s social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318.
- Matsuda, M. J. (1993). Public response to hate speech: Considering the victim’s story. In M. J. Matsuda, C. R. Lawrence III, R. Delgado, & K. W. Crenshaw (Eds.), *Words that wound: Critical race theory, assaultive speech, and the First Amendment* (pp. 17–52). Boulder, CO: Westview.

- Miller, D. T., & Vidmar, N. (1981). The social psychology of punishment reactions. In M. J. Lerner & S. C. Lerner (Eds.), *The justice motive in social behavior* (pp. 145–172). New York: Academic Press.
- Smith, H. J., & Tyler, T. R. (1997). Choosing the right pond: The impact of group membership on self-esteem and group-oriented behavior. *Journal of Experimental Social Psychology*, 33(2), 146–170.
- Swim, J. K., & Stangor, C. (1998). Introduction. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 1–8). San Diego, CA: Academic Press.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–12). Chicago, Nelson Hall.
- Turner, J. C. (1999). Some current issues in research on social identity and self categorization theories. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 6–24). Oxford, UK: Blackwell.
- Tyler, T. R., Boeckmann, R. J., Smith, H. J., & Huo, Y. J. (1997). *Social justice in a diverse society*. Boulder, CO: Westview Press.

ROBERT BOECKMANN received his PhD from the University of California, Berkeley, in 1996 and is an Assistant Professor of Psychology at Flinders University of South Australia. His primary research interest is in the area of social justice, with an emphasis in retributive justice. His research on hate speech represents a conjunction of this emphasis and an interest in social identification. His recent research has appeared in *Psychology, Public Policy and the Law: Law in Context* and in *Law & Society Review*. He is also the author, with Tom Tyler, Heather Smith, and Yuen Huo, of *Social Justice in a Diverse Society*.

JEFFREY C. LIEW completed his BA in Psychology at the University of California, Berkeley, in 1995. He is now a graduate student in the Department of Psychology at Arizona State University. His research interests are in understanding the factors (such as socialization and social cognition) that may mediate the relationship between emotional self-regulation and social adjustment.