

# Racial Prejudice, Intergroup Hate, and Blatant and Subtle Bias of Whites toward Blacks in Legal Decision Making in the United States

Adam R. Pearson\*, John F. Dovidio and Felicia Pratto

*University of Connecticut, USA*

## ABSTRACT

The present study examined the multidimensional nature of intergroup hate and the potential roles of hate and prejudice in expressions of White Americans' treatment of Blacks within the context of the U.S. legal system. White participants in the U.S. read about a provoked or unprovoked violent assault perpetrated by a Black assailant on a White victim. Emotional reactions and recommendations for punishment (prescribed sentencing and support for the death penalty) were assessed. Supportive of Sternberg's (2003) duplex model of hate, we found that explicit (self-reported) hate reflected separate components of negation of intimacy (e.g., disgust and repulsion), passion (anger and fear), and devaluation/commitment (e.g., attributions of evil and inhuman); these components, in turn, differentially mediated punitiveness toward the assailant. The results also revealed that although the direct effect of prejudice on retribution was mediated by self-reported hate, more subtle and indirect effects occurred independently of hate or its affective components. Practical and theoretical implications of these findings for biased decision making in legal contexts are considered. *Key words:* Blatant bias, Hate, Racial prejudice, Sternberg's duplex model of hate.

## RESUMEN

Este estudio examina la naturaleza multidimensional del odio intergrupar y el papel potencial del odio y del prejuicio, en las expresiones del tratamiento que los estadounidenses blancos dan a los negros dentro del contexto del sistema legal estadounidense. Participantes blancos leyeron un informe acerca de un asalto violento, provocado o no provocado, y perpetrado por un asaltante negro a una víctima blanca. Se evaluaron las reacciones emocionales y las recomendaciones de castigo (la sentencia a dictar y el apoyo a la pena de muerte). Se encontró apoyo para el modelo doble de odio propuesto por Sternberg (2003), según el cual el odio explícito (autoinformado) refleja componentes diferenciados de negación de la intimidad (por ejemplo, asco y rechazo), pasión (enfado y miedo) y devaluación/compromiso (por ejemplo, atribuciones de maldad y falta de humanidad). A su vez, estos componentes actuaban de forma diferente como mediadores de la punitividad hacia el asaltante. Los resultados pusieron también de manifiesto que, aunque el efecto directo del prejuicio sobre la retribución estaba mediado por el odio autoinformado, tenían lugar otros efectos más sutiles e indirectos, independientemente del odio o de sus componentes afectivos. Se analizan las implicaciones teóricas y prácticas de estos resultados para la toma sesgada de decisiones en los contextos legales. *Palabras clave:* modelo doble de odio de Sternberg, odio, prejuicio racial, sesgo manifiesto.

\*Correspondence concerning this article should be addressed to the first author: Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020 Storrs, CT 06269-1020, USA. E-mail: adam.pearson@uconn.edu

Although prejudice and hate have been commonly linked in popular vernacular and policy (for example, use of the label “hate crime,” to denote a criminal offense in the United States deemed motivated by intergroup bias), since Allport’s (1954) treatment in his seminal volume *The Nature of Prejudice*, only limited attention has been paid in the psychological literature to their relation (see Smith & Mackie, 2005). Recent work has focused on the psychology underlying hate crimes (Boeckmann & Turpin-Petrosino, 2002; Cogan & Marcus-Newhall, 2002; Craig, 2002; Torres, 1999) and mass violence (Staub, 1989; Sternberg, 2003), as well as to the experience of hate in interpersonal contexts (Fitness & Fletcher, 1993) and as individual pathology (Gaylin, 2003). Nevertheless, the precise role of hate in less extreme expressions of intergroup prejudice remains surprisingly understudied. The present research, drawing on recent work on intergroup emotions (Cottrell & Neuberg, 2005; Smith & Mackie, 2005) and on conceptualizations of hate (Sternberg, 2003), thus examined, first, the situational conditions that can arouse hate and its potential affective and cognitive underpinnings, and second, the potential roles of hate and prejudice, and their interrelationships, in expressions of White Americans’ treatment of Blacks.

The scarcity of research on hate and, in particular, its potential role as an affective mediator of intergroup prejudice may stem in part from the wide range of definitions, meanings, and contexts through which the term has been invoked by both the lay and scientific communities. Darwin (1872/1969) was perhaps the first to situate the roots of hate in prejudice, stating, “If we have suffered or expect to suffer some willful injury from a man, or if he is in any way offensive to us, we dislike him; and dislike easily rises into hatred” (p. 237). Baldwin (1901/1960) pointed to the destructive nature of hate and the importance of the situational context in its elicitation, defining hate as “an emotion characterized by the type of aversion which aims to damage or destroy, under conditions of more or less permanent restraint, limitation, or powerlessness, and the feeling-tone of intense anger” (pp. 442-443). Allport (1954) characterized hate as a sentiment, rather than emotion, and often used the term interchangeably with prejudice, characterizing the “prejudiced pattern” as one involving various degrees and kinds of hatred” (p. 366).

In his duplex theory of hate, Sternberg (2003) recently extended a conception of hate in its own right in a way that applies to both individuals and groups, infusing both cognitive and affective elements in its elicitation. He writes, “Typically, hate is thought of as a single emotion. But there is reason to believe that it has multiple components that can manifest themselves in different ways on different occasions” (p. 306). Sternberg identifies three primary components that separately, or in combination, may comprise hate: (a) the negation of intimacy, which is said to originate from feelings of disgust and repulsion; (b) passion, which is expressed in intense anger or fear arising from perceptions of threat; and (c) commitment, characterized by cognitions of devaluation and diminution through contempt. As Sternberg (2003) acknowledges, these components bear close relation to the family of other-condemning moral emotions -contempt, anger, and disgust (CAD)- identified by Rozin, Lowery, Imada, and Haidt (1999), which have been found to be primarily elicited in response to moral transgressions. The present research thus explored whether Sternberg’s (2003) proposed components of hate, and

thus hate itself, are similarly likely to be aroused in response to a perceived violation of moral standards.

In the present research, we first empirically investigated Sternberg's (2003) conceptualization of hate and its components and the situational antecedents that might trigger them, and then extended these findings by examining the role of hate in the expression of prejudice by Whites toward Blacks in the United States. We focus on the prejudice of White Americans toward Black Americans because of the central role that this phenomenon has had historically in social relations, policy, and politics in the United States, and because it is the most extensively researched prejudice within psychology (Dovidio & Gaertner, 2004). Specifically, we investigated the role of hate in both direct and indirect expressions of racial prejudice.

Although prejudice has been similarly varied in its conception, in contrast to the study of hate, prejudice has received substantially more empirical attention. Research on contemporary racial attitudes has identified different types of racial prejudice with different affective underpinnings, which suggests some manifestations of prejudice may bear closer relation to hate than others. For instance, Kovel (1970) distinguished between dominative and aversive racism. Dominative racism is said to reflect the "old-fashioned," blatant form. According to Kovel, the dominative racist is the "type who acts out bigoted beliefs -he represents the open flame of racial hatred" (p. 54). Aversive racists, in contrast, sympathize with victims of past injustice, support the principle of racial equality, and regard themselves as nonprejudiced, but at the same time possess negative feelings and beliefs about Blacks, which may be nonconscious. Whereas traditional prejudice is expressed directly, aversive racism is manifested indirectly, in subtle and rationalizable ways (Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986).

Prejudice has traditionally been conceptualized as a negative attitude toward a group and its members. Like other attitudes, it is assumed to have cognitive, affective, and behavioral orientation components (Zanna & Rempel, 1988). Supportive of the long tradition in the study of prejudice, Dovidio, Brigham, Johnson, and Gaertner's (1996) meta-analysis of racial attitudes and behavior revealed a significant moderate relationship ( $r = .32$ ) between the level of Whites' self-reported prejudice and their tendency to behave more negatively toward Blacks. One goal of the present research was to examine the degree to which the relationship between Whites' racial prejudice and recommended punishment of a Black assailant is mediated by particular affective responses, specifically hate. That is, to the extent that Whites' self-reported prejudice is fundamentally a form of antipathy toward Blacks, we hypothesized that the relation between Whites' level of prejudice and punishment endorsed for a Black assailant would be mediated by feelings of hate.

Although traditional, direct forms of prejudice have been directly linked to negative affect, contemporary forms of racism have been described as "cooler," more cognitively-based. The aversive racism framework (Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986), proposes that overtly negative feelings, such as antipathy and hate, may not accompany subtle forms of discrimination. In particular, the aversive racism framework proposes that Whites' prejudice will be manifested in negative behavior toward Blacks primarily in situations in which normative guidelines are weak or when it can be

justified on the basis of some factor other than race. Under these circumstances, aversive racists engage in behaviors that ultimately harm Blacks but in ways that allow Whites to maintain their self-image as nonprejudiced.

For instance, Johnson, Whitestone, Jackson, and Gatto (1995) found that in simulated juridic decisions White college students in the United States did not discriminate against Black defendants in their recommendations unless they could justify their decision on the basis of incriminating (though, in this case, inadmissible) evidence. More recently Hodson, Hooper, Dovidio, and Gaertner (2005) conceptually replicated these results in the United Kingdom. Additional research suggests that college students who score higher on explicit prejudice (Brigham, 1993) are more likely to respond negatively toward Blacks primarily when such behavior can be justified by some factor ostensibly unrelated to race (Hodson, Dovidio, Gaertner, 2002). In the present research, an unprovoked attack potentially provides this rationale. The present research extends previous work on aversive racism by exploring the potential role of hate and its hypothesized components in this process. Specifically, we expected that hate would be a more likely component of old-fashioned prejudice and associated direct forms of discrimination than contemporary and more subtle forms of bias.

The present study explored the relationship between self-reported prejudice and the experience of hate in the severity of punishment of a Black assailant who commits a violent crime against a White victim. We focused on recommendations for punishment within a legal context because, as Sidanius, Levin, and Pratto (1998) argue, the legal system can represent systematic aggression of dominant groups against subordinate groups in a socially acceptable form. Traditionally, Blacks and Whites have not been treated equally under the law in the United States (see Sidanius *et al.*, 1998). Across time and locations in the United States, Blacks have been more likely to be convicted of crimes and, if convicted, sentenced to longer terms for similar crimes, particularly if the victim is White. In addition, Blacks in the United States are more likely to receive the death penalty (Government Accounting Office, 1990). Baldus, Woodworth, and Pulaski (1990) examined over 2000 murder cases in Georgia and found that a death sentence was returned in 22% of the cases in which Black defendants were convicted of killing a white victim, but in only 8% of the cases in which the defendant and the victim were White. We propose that the aversive racism is particularly pertinent in the legal context because the body of evidence may offer nonracial justifications for actions and punishment that is formally endorsed and supported under these conditions, potentially leading to direct physical harm when the normative context supports it, as with a government sanctioned death penalty (see Dovidio, Smith, Donnell, & Gaertner, 1997).

In the present study, participants were asked to read one of two versions of a newspaper article that described an assault that had ostensibly occurred in a nearby city. The article described an assault perpetrated by a Black assailant, described as having been either provoked by the victim (the victim was reported to have cut in line and shoved the assailant prior to the assault) or unprovoked (police reports indicated the victim was an innocent bystander), that left a White victim in critical condition. Participants then completed a series of questionnaires that assessed the effectiveness of the Provocation manipulation, trait attributions and emotional responses related to the

assailant, and outcome measures related to recommended punishment of the assailant. The main dependent measures were participants' prescribed prison sentence (length of prison term and number of years required to serve before being offered parole) and support for the death penalty for the perpetrator of the crime.

Specifically, we explored both direct and more indirect effects of racial prejudice on the recommended punishment for the Black assailant. In terms of the direct effect, as suggested by the Dovidio *et al* (1996) meta-analytic review, White participants higher in explicit racial prejudice were expected to recommend more severe consequences for the Black assailant. We further examined whether explicit (self-reported) hate, and its hypothesized theoretical components (Sternberg, 2003), would mediate this relationship. Our predictions relating to the indirect effect were based on the Hodson *et al* (2002) research.

In the present study, the attack without provocation was hypothesized to provide justification for more severe recommendations for punishment. Thus, we predicted that higher levels of racial prejudice would relate to greater punishment of a Black assailant primarily in the unprovoked rather than the provoked condition. To the extent that subtle bias depends strongly on rationalization and justification, we anticipated that negative affective reactions, such as explicit hate and its hypothesized components (Sternberg, 2003), would be less related to subtle bias than the direct expression of prejudice. Pettigrew and Meertens (1995) found that whereas blatant prejudice was associated with negative emotions, subtle prejudice was not (see also Gaertner *et al.*, 1997). Thus, whereas we hypothesized that explicit hate would mediate the relationship between Whites' explicit prejudice and punitiveness toward a Black assailant, we expected the role of explicit hate and its theoretical components would play a much more minor role in accounting for the pattern of subtle bias associated with aversive racism.

## METHOD

### *Participants*

Participants were 183 self-identified White-American students (70% female) who were enrolled in introductory psychology classes. Students participated in small groups with 5-10 other students voluntarily in exchange for partial course credit.

### *Procedure*

Upon arrival, participants were informed that they would be asked to complete questionnaires for a series of different investigations. In a separate instruction sheet, one of these investigations (the present research) was described as having been designed to examine the impact of different forms of media (television, magazines, or newspapers) on people's reactions to local, national, and international news. In the instruction sheet provided, they were informed that they had been assigned to the newspaper condition and would be asked to give their reactions to the events described in a subsequent article. Participants were then administered one of two versions of a short (approximately

250-word) newspaper article. Immediately after reading the article, participants were asked to complete a series of questionnaires described as assessing their reactions to and opinions about the events in the article. In both the task instructions and in the initial information sheet, they were informed that their responses would remain anonymous and that there were no right or wrong answers.

*Editorial.* Two versions of the newspaper editorial ("Family, Friends Hold Vigil For Victim of Beating") were created and modeled after a regional urban newspaper. In both versions, a White male student was described as having been "attacked and brutally beaten" by a Black male after a confrontation outside of a local bar, leaving the student in critical condition at a nearby hospital. In both versions, the perpetrator of the attack was reported as having been "arrested two days after the attack and formally charged with the crime," and the victim and assailant were described as having been unacquainted prior to the assault. One version of the article (Provoked condition), suggested that the actions of the victim may have prompted the incident. In this version, police investigators reported that the attack appeared to have been provoked. Witnesses were reported to have seen the victim "cut in line just before the attack," and the lead investigator was quoted as remarking that the victim "apparently refused to move," and that witnesses had seen the victim shove the perpetrator just before he was attacked. The second version of the article (Unprovoked condition), quoted police investigators as stating that the attack appeared to have been unprovoked. All other content remained identical. Both versions of the article included reactions from family members and friends of the victim.

*Dependent measures.* As part of what was ostensibly one of a series of initial separate studies, participants were asked to complete Brigham's (1993) Attitudes Toward Blacks Scale (Cronbach alpha = .88), providing a pretest measure of explicit racial prejudice. To support this cover story, multiple consent forms were administered during the course of the experimental session.

After reading the article, participants were asked to complete a series of questionnaires containing a trait attribution task related to the perpetrator of the crime, an emotional response task reflecting the extent to which they experienced a variety of emotions when thinking about the perpetrator, and a series of questions designed to tap into the participants' opinions about how the perpetrator should be punished (sentencing measures and support for the death penalty, had the victim died after the attack). On the trait attribution task, participants were asked to indicate the extent to which they would characterize the perpetrator of the reported crime (1 = not at all to 7 = extremely) on a series of traits. The emotional response task asked participants to rate, on a scale from 1 to 7 (1 = not at all; 7 = extremely), the extent to which they personally felt a variety of emotions when they thought about the person who committed the reported crime.

Two trait attributions, moral and justified, were included to assess whether the Provocation manipulation related, as expected, to moral transgression. The two items designed to assess Sternberg's (2003) passion component consisted of the emotional experiences of anger and fear (Cronbach alpha = .60). Negation of intimacy was represented by the characterization of the offender as repulsive combined with the emotional experience

of disgust (Cronbach alpha = .76), and the devaluation/commitment component of Sternberg's (2003) model was assessed using the items evil and inhuman (Cronbach alpha = .85) on the trait attribution task. Explicit (self-reported) hate was directly represented by an item on the emotional response questionnaire.

After completing the attribution and emotional experience questionnaires, participants responded to a series of questions designed to assess punitiveness for the perpetrator of the crime. One measure of punitiveness involved socially-sanctioned ostracism in the form of recommendations for the length of sentencing for the perpetrator of the crime. Two open-response questions related to sentencing were asked: (a) "If the perpetrator were caught and found guilty, how many years would you recommend they be sentenced to jail?"; and (b) "What is the *minimum* number of years the perpetrator should serve in jail before being offered parole?" Scores were averaged across these two items to create a composite variable reflecting the prescribed prison sentence (Cronbach alpha = .75). A second outcome measure assessed participants' support for the death penalty for the perpetrator (-3 = not at all to 3 = very much), had the victim died after the assault. A modest correlation between support for the death penalty and the prescribed length of prison sentence,  $r = .38$ ,  $p < .01$ , indicated reasonable discriminant validity among these measures.

## RESULTS

The analyses investigated, in order, (a) the relation of Sternberg's (2003) theoretical hate components to explicit self-reported hate for the perpetrator of the violent crime, (b) the impact of the Provocation manipulation and participants' level of explicit prejudice on the two outcome measures (prescribed length of prison sentence and support for the death penalty), (c) and the potential role of self-reported hate and the hate components, respectively, as mediators of the effect of Provocation and prejudice on each of these outcomes. No systematic effects were found associated with participant gender across the dependent variables, thus, gender was not included as a variable in the analyses described here.

Consistent with Sternberg's (2003) duplex model, all three theoretical hate components, passion, negation of intimacy, and devaluation, were found to be strongly related to explicitly reported hate ( $\beta_s > .55$ ), as measured in the emotional response questionnaire (Table 1). However, when entered simultaneously in a standard regression, passion,  $\beta = .32$ ,  $t = 4.27$ ,  $p < .001$ , and devaluation,  $\beta = .35$ ,  $t = 4.14$ ,  $p < .001$ , but not negation of intimacy,  $\beta = .08$ ,  $t = .85$ ,  $p < .40$ , were found to uniquely predict explicit hate. The multiple correlation ( $R = .66$ ) was substantially higher than the unique contribution of any single component. The results of sequential regression analyses indicate this effect was due to the combined influence of passion and devaluation, rather than the individual effect of either component; when passion and devaluation were each entered into separate regression analyses along with negation of intimacy as predictors of explicit hate, negation of intimacy remained a unique predictor of hate ( $ps < .005$ ).

With respect to the manipulation check items, as expected, participants judged the assault as less justified when the attack was unprovoked than when it was provoked,

$M_s = 1.75$  vs.  $3.10$ ,  $F(1,180) = 38.70$ ,  $p < .001$ . In addition, they described the perpetrator as less moral when the attack was unprovoked than when it was provoked ( $M_s = 1.61$  vs.  $2.10$ ,  $F(1,179) = 6.10$ ,  $p = .015$ ). Again, these results support the assumption that an unprovoked attack is seen as a greater moral transgression than a provoked attack.

In terms of recommended sentencing, as expected, a main effect of Provocation was found,  $F(1,179) = 18.05$ ,  $p < .001$ ; participants in the Unprovoked condition recommended longer sentences than those in the Provoked condition ( $M = 9.66$  vs.  $4.99$  years). Prejudice (mean-centered) did not reliably predict sentencing ( $\beta = .04$ ,  $t = .54$ ,  $p < .59$ ), however, when entered simultaneously with its main effects, a significant Provocation x Prejudice interaction emerged ( $\beta = .17$ ,  $t = 2.43$ ,  $p < .02$ ).

With regard to support for the death penalty, a marginally significant main effect was found for Provocation,  $F(1,181) = 3.40$ ,  $p = .06$ , such that participants in the Unprovoked condition showed stronger endorsement for the death penalty than those in the Provoked condition ( $M_s = -.11$  vs.  $-.73$ ). A main effect of prejudice was also found ( $\beta = .23$ ,  $t = 3.13$ ,  $p < .003$ ), and when Provocation and prejudice were entered in a regression analysis along with its interaction term, a reliable Provocation x Prejudice interaction was found ( $\beta = .19$ ,  $t = 2.67$ ,  $p < .01$ ).

In the first set of analyses, we explored, separately, explicitly reported hate and each of the three theoretical components of hate (Sternberg, 2003) as potential mediators of the effect of Provocation, prejudice, and their interaction on respondents' punitiveness toward the perpetrator. Mediation is established when (a) the independent variable (Provocation) predicts the outcome variable (e.g., sentencing), (b) the independent variable predicts the potential mediator, (c) the mediator predicts the outcome variable when entered simultaneously with the independent variable, and (d) the effect of the independent variable on the outcome variable is weakened and potentially reduced to nonsignificance when controlling for the effect of the mediator (Baron & Kenny, 1986). Tests for mediation, including contrasts among multiple mediators, utilized bootstrapping procedures for estimating indirect effects of the criterion variables on the outcome variables through the proposed mediators (Preacher & Hayes, 2004). Bootstrapping is a nonparametric approach to effect-size estimation and hypothesis testing that, unlike other methods for testing mediation (e.g., Sobel test), imposes no assumptions about the shape of the distributions of the variables or the sampling distribution of the statistic (for reviews, see MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher & Hayes, 2004; Shrout & Bolger, 2002).

*Mediational role of explicit hate and hate components in prescribed length of sentence.* Because Prejudice did not directly predict sentencing, mediational analyses examined, in order, the main effect of Provocation, and then the Provocation x Prejudice interaction effect on the prescribed length of prison sentence.

In first examining the effect of Provocation on sentencing, all three hate components (passion, negation of intimacy, and devaluation) but not explicitly reported hate, were correlated with Provocation (see Table 1) and therefore, qualified as potential mediators. When each was entered with Provocation in separate regression analyses, all three components, passion ( $\beta = .34$ ,  $t = 4.76$ ,  $p < .001$ ), negation of intimacy ( $\beta = .33$ ,  $t = 4.52$ ,  $p < .001$ ), and devaluation ( $\beta = .27$ ,  $t = 3.67$ ,  $p < .001$ ), predicted sentencing. Consistent



with partial mediation, the effect of the Provocation manipulation on sentencing ( $r = .30, p < .001$ ) remained significant in each case, but was weaker than when it was considered as a lone predictor (passion equation:  $\beta = .20, t = 2.86, p < .006$ ; negation of intimacy equation:  $\beta = .18, t = 2.48, p < .02$ ; devaluation equation:  $\beta = .21, t = 2.86, p < .006$ ). Individual tests for mediation utilizing bootstrap estimation of indirect effects with 1000 replications (Preacher & Hayes, 2004) confirmed each component as a partial mediator, two-tailed  $ps < .01$  (99% confidence interval (CI) for passion indirect effect = .027, .184; negation of intimacy = .043, .198; and devaluation = .016, .170). Multiple mediation analyses considering all three components simultaneously revealed a significant partial mediational effect for only the passion component,  $p < .05$  (95% CI for passion indirect effect = .013, .131; negation of intimacy = -.031, .155; devaluation = -.054, .010). Further analyses examining only the items comprising the passion component (anger, fear) simultaneously as potential mediators revealed both as reliable partial mediators (95% CI for anger = .015, .124; fear = .004, .079) with pairwise contrasts showing no significant difference in the size of their indirect effects (95% CI = -.034, .094).

In a final set of mediational analyses, we tested the role of explicitly reported hate and its theoretical components as potential mediators of the effect of the Provocation x Prejudice interaction (subtle prejudice) on the prescribed length of sentence. As anticipated, neither explicit hate, passion, nor negation of intimacy were reliably predicted by the Provocation x Prejudice interaction when entered in regression analyses with its main effects (explicit hate  $\beta = .11, t = 1.53, p = .14$ ; passion  $\beta = .08, t = 1.16, p = .25$ ; negation of intimacy  $\beta = .10, t = 1.53, p = .13$ ). However, the interaction did reliably predict the devaluation component ( $\beta = .16, t = 2.21, p < .03$ ). As predicted, further analyses revealed partial mediation by the devaluation component: when devaluation was entered simultaneously with the interaction term and main effects as predictors of sentencing, the effect of devaluation remained significant ( $\beta = .25, t = 3.32, p < .001$ ), whereas the effect of the interaction, though still reliable ( $\beta = .15, t = 2.11, p < .04$ ), was significantly diminished,  $p < .05$  (95% CI for devaluation indirect effect = .004, .147).

*Mediational role of explicit hate and hate components in support for the death penalty.* In first examining the effect of Provocation on support for the death penalty, as in the previous analyses, all three hate components (passion, negation of intimacy, and devaluation) but not explicitly reported hate, were correlated with Provocation (see Table 1) and, thus, qualified as potential mediators. When each was entered with Provocation in separate regression analyses, all three components, passion ( $\beta = .30, t = 4.09, p < .001$ ), negation of intimacy ( $\beta = .28, t = 3.69, p < .001$ ), and devaluation ( $\beta = .34, t = 4.67, p < .001$ ), predicted support for the death penalty. In each case, the effect of Provocation on the outcome variable ( $r = .14, p < .06$ ) was reduced to non-significance when including the proposed mediator in the regression equation ( $\beta s < .05, ps > .48$ ), suggesting mediation. In separate tests, bootstrap estimation of indirect effects (1000 replications) confirmed reliable mediation by each hate component, all  $ps < .01$  (99% CI for passion indirect effect = .043, .406; negation of intimacy = .069, .467; devaluation = .100, .444). When considering all three hate components simultaneously as potential mediators, only devaluation remained as a reliable mediator,  $p < .05$  (95% CI for passion = -.006, .300; negation of intimacy = -.294, .171; devaluation = .054, .381).

Table 1. Correlations Among Racial Prejudice, Provocation, Explicit Hate, Sternberg's (2003) Hate Components, Prescribed Length of Sentence, and Support for the Death Penalty

Variable	2	3	4	5	6	7	Length of Sentence
1. Prejudice	-.06	.15*	-.10	-.09	.03	.23**	.04
2. Provocation <sup>1</sup>		.12	.28**	.37**	.31**	.14†	.30**
3. Explicit Hate			.57**	.56**	.59**	.26**	.31**
4. Passion				.65**	.56**	.57**	.40**
5. Negation of Intimacy					.75**	.56**	.40**
6. Devaluation						.59**	.34**
7. Death Penalty							.38**
Provocation x Prejudice <sup>2</sup>		.11	.08	.11	.16*	.19**	.17*

<sup>1</sup> p < .1; \* p < .05; \*\* p < .01 <sup>1</sup> Effect coded unprovoked assault = 1, provoked assault = -1.

<sup>2</sup> Row values are interaction term beta-weights, controlling for main effects.

Next, we examined the effect of prejudice on support for the death penalty. As can be seen in Table 1, only explicit hate was found to be correlated with prejudice and, thus, qualified as a potential mediator. When entered in a regression equation with prejudice (mean-centered), explicit hate remained a significant predictor of the outcome variable ( $\beta = .23$ ,  $t = 3.25$ ,  $p < .002$ ) and the effect of prejudice, though significant, was reduced ( $\beta = .23$  to  $.19$ ,  $ps < .01$ ). As predicted, consistent with partial mediation, bootstrap estimation (1000 replications) revealed a significant indirect effect of prejudice on support for the death penalty through explicit hate,  $p < .05$  (95% CI =  $.01$ ,  $.32$ ).

Finally, we examined the role of explicitly reported hate and its theoretical components as potential mediators of the effect of the Provocation x Prejudice interaction (subtle prejudice) on support for the death penalty. As reported previously, explicit hate, passion, and negation of intimacy were reliably not predicted by the Provocation x Prejudice interaction when entered in regression analyses with its main effects, leaving devaluation as the lone qualifying mediator. Consistent with partial mediation, when entered in a regression equation with the Provocation x Prejudice interaction term and its main effects, devaluation remained a significant predictor of the outcome variable ( $\beta = .31$ ,  $t = 4.27$ ,  $p < .001$ ) and the effect of the interaction, though still significant, was diminished ( $\beta = .19$  to  $.14$ ,  $ps < .05$ ). Follow-up tests for mediation controlling for main effects (bootstrap estimation using 1000 replications) showed a significant indirect effect on support for the death penalty through devaluation,  $p < .05$  (95% CI =  $.011$ ,  $.425$ ), consistent with our predictions and prior research (Hodson *et al.*, 2002).

## DISCUSSION

Although prejudice has traditionally been defined in terms of intergroup emotion as an antipathy toward a group and its members (Allport, 1954), with highly destructive behaviors toward a group and its members frequently characterized as actions of “hate,” relatively little empirical work has attempted to link overt and subtle prejudice with emotional reactions to group members. The present study attempted to address this issue by utilizing two conceptualizations of hate, explicit (self-reported) hate and Sternberg’s (2003) multidimensional theoretical model of hate. Our research provided an initial empirical test of Sternberg’s (2003) conceptualization.

In particular, supportive of Sternberg’s duplex model of hate, we found that explicit (self-reported) hate reflects the separate components of negation of intimacy (e.g., disgust and repulsion), passion (anger and fear), and devaluation/commitment (e.g., attributions of evil and inhuman); these components, in turn, differentially predicted different types of responses by Whites in the United States toward a Black individual arrested for a violent crime. Whereas passion was the strongest predictor of the length of sentencing when all three components were entered simultaneously in regression analyses, devaluation was the strongest predictor of support for the death penalty. These findings represent initial empirical tests of the theoretical and pragmatic implications of Sternberg’s (2003) multidimensional theory of hate.

Importantly, the present study also provides further evidence of the distinction between traditional, blatant prejudice and more subtle forms, such as aversive racism (Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986). Consistent with Allport’s (1954) definition of prejudice as antipathy, explicit (self-reported) hate mediated the direct (overt) relationship between racial prejudice and support for the death penalty. That is, White participants higher in blatant prejudice expressed higher levels of hate toward the Black assailant, which in turn mediated greater support for the death penalty, regardless of whether the attack was provoked or unprovoked.

However, we also found evidence consistent with the notion that contemporary forms of racism have emerged that are more subtle, indirect, and rationalizable (Dovidio & Gaertner, 2004; Kovel, 1970) and do not typically involve open feelings of antipathy or hate. As we have previously argued, contemporary racism is more subtle than traditional racism in the United States. It is not manifested directly, but instead it is expressed when a negative response can be justified on another, ostensibly nonracial, basis. In the present study, the description of an unprovoked attack by the assailant provided this justification. In particular, White participants higher in racial prejudice showed more negative responses toward Blacks, in terms of both recommending longer sentences and in showing greater support for the death penalty, primarily when the attack was described as unprovoked rather than provoked. Further, supportive of the aversive racism prediction that this form of bias is affectively “cooler,” involving largely cognitive processes to justify negative responses rather than conscious experience of antipathy or hate, the affective components of Sternberg’s (2003) model, negation of intimacy and passion, did not mediate this effect; only the more cognitive component, consisting of devaluing attributions, was found to be a reliable mediator.

The findings of the present study thus have valuable theoretical and practical implications. Theoretically, first, the current work demonstrates qualitative differences between the direct and indirect manifestations of racial prejudice. These forms of racism are fundamentally different, both in terms of how they are experienced from the perspective of the prejudiced individual as well as in the ways they are manifested in discrimination. Whereas explicit feelings of hate are integral to old-fashioned racism, they do not appear to be directly involved in the dynamics of subtle bias. Hate and racism are therefore not necessarily synonymous. Nevertheless, understanding the role of intergroup emotions helps to explain when and how more aggressive intergroup actions can occur.

Second, our work supports Sternberg's (2003) position that it is important to understand the different aspects that contribute to the experience and outcomes of even the most basic affective experiences, such as hate. The research provides an initial test of Sternberg's (2003) model of hate and offers direct evidence of the three basic components of hate that he hypothesized: passion, negation of intimacy, and devaluation. Each of these components related directly to explicit feelings of hate, but the different components related differentially to the different outcomes. As Sternberg (2003) proposes, recognizing the multidimensional nature of hate can illuminate the underlying affective and cognitive processes involved to distinguish different types of hate and to potentially understand their different action tendencies and, ultimately, consequences. However, our results also revealed that the experience of hate can represent "more than the sum of its parts." Specifically, undifferentiated, explicitly reported hate predicted both the length of sentencing and support for the death penalty, independent of any of the specific components either separately or in combination. In this respect, traditional prejudice may represent a general orientation of Whites toward Blacks that is characterized and driven by open and undifferentiated antipathy; there is little that is subtle or indirect about it.

Practically, the present research illustrates that, at least in the United States, justice is not necessarily color blind. Racial attitudes toward Black assailants mattered significantly in severity of sentencing and support for the death penalty, both directly and indirectly (moderated by provocation of the attack). However, regardless of whether prejudice operated in a more direct or indirect manner, the consequences were the same – more severe treatment of Blacks. Because the present research focused on the roles of racial prejudice, it did not include a condition in which the assailant was White. However, our findings extend previous research, laboratory and archival, indicating that Black defendants tend to be judged as more guilty and are punished more severely than are Whites (Sidanius *et al.*, 1998) by demonstrating the role of prejudice in responses to Blacks. Whereas the beliefs and negative affect associated with traditional prejudice may be detected in preliminary questioning of prospective jurors, in the *voir dire*, subtle prejudice, which is less obviously antipathy, can still exert a systematic influence that disadvantages Blacks in the U.S. legal system. The present research offers insights into the complex nature of these biases and can be instrumental in devising new ways to identify and potentially correct for the effects of contemporary racism.

In conclusion, the present research emphasizes the importance not only of understanding the beliefs associated with negative intergroup attitudes but also the

affective reactions involved. We propose that understanding the dynamics of Whites' reactions to crimes committed by Blacks against Whites, including the role of moderating and mediating factors such as the nature of the crime, prejudice, and particular affective responses, such as hate, can have important social implications. We returned to Allport's (1954) conception of prejudice as antipathy and found, consistent with this conceptualization, that explicit hate does indeed mediate more negative recommended treatment of Blacks within the legal system by Whites. However, both the nature of antipathy and the nature of prejudice may be more complex and differentiated than Allport suggested, in some instances incorporating devaluing cognitive attributions that may be largely devoid of negative affect.

#### NOTES

1. When all three components of hate were entered simultaneously in regression analyses to predict the length of sentencing, passion was the strongest predictor ( $\beta = .22$ ,  $t = 2.40$ ,  $p < .02$ ). A parallel analysis for support for the death penalty revealed that devaluation was the strongest predictor of ( $\beta = .28$ ,  $t = 2.64$ ,  $p < .01$ ).

#### REFERENCES

- Allport GW (1954). *The nature of prejudice*. New York: Addison-Wesley.
- Baldus D, Woodworth G & Pulaski C (1990). *Equal justice and the death penalty: A legal and empirical analysis*. Boston, MA: Northeastern University Press.
- Baldwin JM (1901/1960). *Dictionary of philosophy and psychology*. Gloucester, MA: Macmillan.
- Baron RM & Kenny DA (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Boeckmann RJ & Turpin-Petrosino C (2002). Understanding the harm of hate crimes. *Journal of Social Issues*, 58, 207-225.
- Brigham JC (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, 23, 1933-1967.
- Cogan JC & Marcus-Newhall A (2002). Introduction. Hate crimes: Research, policy, and action. *American Behavioral Scientist*, 46, 8-13.
- Cottrell CA & Neuberg SL (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to 'prejudice'. *Journal of Personality and Social Psychology*, 88, 770-789.
- Craig KM (2002). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. *Aggression and Violent Behavior*, 7, 85-101.
- Darwin C (1872/1969). *The expression of emotions in man and animals*. New York: Greenwood Press.
- Dovidio JF, Brigham JC, Johnson BT & Gaertner SL (1996). Stereotyping, prejudice, and discrimination: Another look. In CN Macrae, C Stangor & M Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 276-319). New York: Guilford.
- Dovidio JF & Gaertner SL (2004). Aversive racism. In MP Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 1-51). San Diego, CA: Elsevier Academic Press.
- Dovidio JF, Smith JK, Donnell AG & Gaertner SL (1997). Racial attitudes and the death penalty. *Journal of Applied Social Psychology*, 27, 1468-1487.

- Fitness J & Fletcher GJO (1993). Love, hate, anger, and jealousy in close relationships: A prototype and cognitive appraisal analysis. *Journal of Personality and Social Psychology*, 65, 942-958.
- Gaertner SL & Dovidio JF (1986). *The aversive form of racism*. San Diego: Academic Press, Inc.
- Gaertner SL, Dovidio JF, Banker B, Rust M, Nier J, Mottola G & Ward C (1997). Does racism necessarily mean anti-Blackness? Aversive racism and pro-Whiteness. In M Fine, L Powell, L Weis & M Wong (Eds.), *Off white* (pp. 167-178). London: Routledge.
- Government Accounting Office (1990). *Death penalty sentencing: Research indicates pattern of racial disparities* (Report to Senate and House Committee on the Judiciary, 101st Congress, 2nd Session). Washington, DC: Author.
- Hodson G, Dovidio JF & Gaertner SL (2002). Processes in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin*, 28, 460-471.
- Hodson G, Hooper H, Dovidio JF & Gaertner S L (2005). Aversive racism in Britain: Legal decisions and the use of inadmissible evidence. *European Journal of Social Psychology*, 35, 437-448.
- Johnson JD, Whitestone E, Jackson LA & Gatto L (1995). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin*, 21, 893-898.
- Kovel J (1970). *White racism: A psychohistory*. New York, NY: Pantheon.
- MacKinnon DP, Lockwood CM, Hoffman JM, West SG & Sheets V (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.
- Pettigrew TF & Meertens RW (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, 25, 57-76.
- Preacher KJ & Hayes AF (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers*, 36, 717-731.
- Rozin P, Lowery L, Imada S & Haidt J (1999). The cad triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76, 574-586.
- Sidanius J, Levin S & Pratto F (1998). Hierarchical group relations, institutional terror, and the dynamics of the criminal justice system. In J Eberhardt & ST Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 136-165). Newbury Park, CA: Sage.
- Smith ER & Mackie DM (2005). Aggression, hatred, and other emotions. In JF Dovidio, P Glick & LA Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 361-376). Malden, MA: Blackwell Publishing.
- Shrout PE & Bolger N (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422-445.
- Staub E (1989). *The roots of evil: The origins of genocide and other group violence*. New York: Cambridge University Press.
- Sternberg RJ (2003). A duplex theory of hate: Development and application to terrorism, massacres, and genocide. *Review of General Psychology*, 7, 299-328.
- Torres S (1999). Hate crimes against African Americans: The extent of the problem. *Journal of Contemporary Criminal Justice*, 15, 48-63.
- Zanna MP & Rempel JK (1988) Attitudes: A new look at an old concept. In D Bar-Tal & AW Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315-334). New York: Cambridge University Press.

Received, November 18, 2006  
Final Acceptance, April 3, 2007