Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Experimental Social Psychology 42 (2006) 491–499
Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

Are members of low status groups perceived as bad, or badly off? Egalitarian negative associations and automatic prejudice ☆

Eric Luis Uhlmann *, Victoria L. Brescoll, Elizabeth Levy Paluck

Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06520, USA

Received 24 February 2003; revised 4 October 2004

Abstract

Three studies explored the hypothesis that implicit measures of prejudice can tap negative, yet egalitarian associations. In Study 1, automatically associating African Americans with oppression predicted greater automatic prejudice. In Studies 2 and 3, classically conditioning associations between the novel group Noffians and words like *oppressed*, *maltreated*, and *victimized* led to greater automatic prejudice against Noffians. Results suggest that White Americans' negative automatic associations with African Americans may partly result from associating members of low status groups with unfair circumstances. Because automatic associations predict prejudiced behaviors, the burden of proof is on those wishing to argue that egalitarian negative associations complicate the assessment of automatic attitudes rather than contribute to prejudiced responses. Discussion focuses on the implications of egalitarian negative associations for the theory and measurement of automatic prejudice.
© 2006 Published by Elsevier Inc.

Keywords: Automatic attitudes; Implicit measures; Implicit attitudes; Prejudice

Contemporary theories of prejudice posit that despite the steady decrease in overt racism over the last 30 years, automatic prejudices remain pervasive (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Gaertner & Dovidio, 1986; Greenwald & Banaji, 1995). Unlike explicit prejudices, which involve the intentional endorsement of derogatory attitudes towards social outgroups, automatic prejudice is conceptualized as a negative automatic association with a target group (Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). These associations are proposed to be passively conditioned in us by our culture and shape our thoughts, judgments, and behaviors regardless of our intentions.

Supporting theories of automatic prejudice, studies employing reaction time based implicit measures find that the vast majority of White Americans automatically associate African Americans with "Bad" (as evidenced by relatively faster responses when African American faces are paired with negative words than positive ones; Dovidio et al., 1997; Fazio et al., 1995; Greenwald, McGhee, & Schwartz, 1998; Nosek, Banaji, & Greenwald, 2002; Nosek & Banaji, 2001; Wittenbrink, Judd, & Park, 1997).[1] This occurs despite White Americans' self-reported rejection of prejudiced attitudes towards African Americans.

How best to interpret such associations has been the subject of much recent debate. Among the issues discussed are whether automatic associations reflect personal attitudes or knowledge of cultural attitudes, irrational biases or rational base rates, and perceiving members of minority groups as bad or badly off (Arkes & Tetlock, 2004; Banaji, Nosek, & Greenwald, 2004; Olson & Fazio, 2004).

☆ Thanks to Geoffrey Cohen, William Cunningham, Daisy Grewal, T. Andrew Poehlman, and Gregory Walton for their comments on earlier versions of this paper.
* Corresponding author.
E-mail address: eric.uhlman@yale.edu (E.L. Uhlmann).

[1] Although it is important to note that evidence for automatic prejudice is also found using other paradigms. For example, stereotypes have a greater impact on social judgments when capacity for conscious thought is depleted (e.g., through distraction or time pressure; Shah, Kruglanski, & Thompson, 1998; van Knippenberg, Dijksterhuis, & Vermeulen, 1999).

0022-1031/$ - see front matter © 2006 Published by Elsevier Inc.
doi:10.1016/j.jesp.2004.10.003

The present research explores the latter of these issues—the possibility that one reason why White Americans associate African Americans with negativity is that they associate them with oppression, maltreatment, and victimization. Such, negative, yet egalitarian associations could lead many White Americans to associate African Americans with "Bad" on implicit measures of prejudice.

It is important to note that implicit measures have been conclusively shown to predict biased behavior against members of minority groups (Dovidio et al., 1997; Fazio et al., 1995; Florak, Scarabis, & Bless, 2001; Hugenberg & Bodenhausen, 2003, 2004; McConnell & Liebold, 2001). For example, automatic associations predict perceiving Black faces as more hostile than White faces (Hugenberg & Bodenhausen, 2004, 2003). Indeed, in Poehlman, Uhlmann, Greenwald, and Banaji (2006) meta-analysis, Implicit Association Test (Greenwald et al., 1998) measures predicted prejudiced behaviors more effectively than self-reported attitudes did. Thus, association-based measures are clearly valid means of assessing individual differences in automatic social cognition. The point of the present research is not to argue that implicit measures are invalid or necessarily contaminated by egalitarian negative associations. Rather, the purpose is to identify a significant source of automatic associations and discuss the implications of egalitarian negative associations for the theory and measurement of automatic attitudes.

Study 1

If implicit measures of prejudice reveal negative associations with African Americans in part because they measure egalitarian negative associations, then egalitarian negative associations should be predictive of automatic prejudice as measured by such tasks. Since viewing African Americans as oppressed is a cornerstone of racial egalitarianism (Jost & Banaji, 1994; Sears, 1988; Sidanius & Pratto, 1999), the extent to which White Americans automatically associate African Americans with oppression was assessed. It was hypothesized that this negative, yet egalitarian association would predict participants' tendency to automatically associate African Americans with "Bad" (the traditional operationalization of automatic prejudice; Dovidio et al., 1997; Fazio et al., 1995; Greenwald et al., 1998).

Method

Participants

72 White undergraduates (38 males, 34 females) participated in the study in return for monetary compensation ($5).

Materials and procedure

The implicit measure employed in this research was the Implicit Association Test (IAT; Greenwald et al., 1998), one of the most widely used (Karpinski & Hilton, 2001) and psychometrically reliable (Cunningham, Preacher, & Banaji, 2001) implicit measures of prejudice. The IAT uses reaction times when categorizing rapidly presented stimuli (e.g., African American faces, European American faces, good words, bad words) into paired categories (e.g., "African American and Bad"; "African American and Good") to gauge the extent to which those categories are automatically associated with one another. Greater speed when African American and Bad are paired (along with European American and Good) than when African American and Good are paired (along with European American and Bad) reflects negative associations with African Americans relative to European Americans.

Participants completed 3 IATs. A Flower vs. Insect, Good vs. Bad IAT (abbreviated as the flower–insect IAT) always came first and was intended to introduce participants to the IAT procedure. Participants next completed an African American vs. European American, Good vs. Bad IAT (henceforth abbreviated as the racial attitudes IAT) and an African American vs. European American, Oppressed vs. Privileged IAT (henceforth abbreviated as the oppression IAT). The words used to represent the category Oppressed were: *oppressed*, *brutalized*, *victimized*, and *mistreated*. The words used to represent the category Privileged were: *privileged*, *rulers*, *dominant*, and *powerful*. The words (and in the case of the racial groups, names) used to represent the categories of Flower, Insect, African American, European American, Good, and Bad were taken from sets of stimuli used in previous research on the IAT (e.g., Cunningham et al., 2001; Dasgupta, McGhee, Greenwald, & Banaji, 2000; Greenwald et al., 1998).

Following Nosek (2002), no practice blocks were used. Each IAT consisted of four critical blocks, two for each IAT task. For example, the oppression IAT consisted of two African American + Privileged/European American + Oppressed IAT tasks and two African American + Oppressed/European American + Privileged IAT tasks. The order of the oppression and racial attitudes IATs was counterbalanced between subjects, as was the order of blocks within each IAT. Because these counterbalancing variables did not interact with any of our effects of interest, they are not discussed further.

Results

The first five trials from each IAT block were discarded because they were intended to serve as a transition from the previous block, and because response latencies were typically longer. Following Greenwald et al. (1998), to correct for anticipatory responses and momentary inattention, latencies less than 300 ms and greater than 3000 ms were recoded as 300 and 3000 ms, respectively.

Response latencies on the IAT tasks are displayed in Fig. 1. To calculate the IAT effects, performance on the

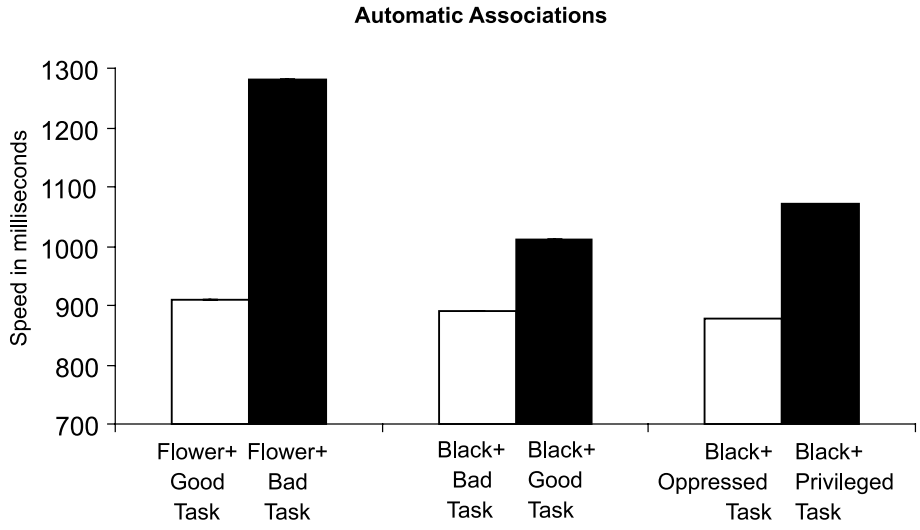

Fig. 1. Response latencies on Implicit Association Test (IAT) tasks. The *Y*-axis represents participants' mean response latency in milliseconds. Slower responses during the Flower + Bad task than during the Flower + Good task reflect positive associations with Flowers relative to Insects, slower responses during the African American + Good task than during the African American + Bad task reflect negative associations with African Americans relative to European Americans, and slower responses during the African American + Privileged task than during the African American + Oppressed task reflect an automatic association between African Americans and oppression.

Flower + Good, African American + Bad, and African American + Oppressed IAT tasks were subtracted from performance on the Flower + Bad, African American + Good, and African American + Privileged IAT tasks, respectively. Participants evidenced more positive automatic associations with Flowers than with Insects (flower–insect IAT effect = 372 ms; $t(70) = 13.65$, $p < .001$) and more positive automatic associations with European Americans than with African Americans (racial attitudes IAT effect = 123 ms; $t(71) = 4.05$, $p < .001$). Consistent with the idea that egalitarian negative associations are held by many if not most White Americans, participants strongly associated European Americans and African Americans with privilege and oppression, respectively (oppression IAT effect = 194 ms; $t(71) = 7.37$, $p < .001$). In fact, the egalitarian negative association evidenced between African Americans and oppression was significantly stronger than the association between African Americans and "Bad," $t(71) = 2.10$, $p = .04$.

Of particular interest was the relationship between automatically associating African Americans with oppression and associating them with "Bad." As hypothesized, higher scores on the oppression IAT (reflecting a strong association between African Americans and oppression) predicted higher scores on the racial attitudes IAT (reflecting a strong association between African Americans and "Bad"), $r(71) = .30$, $p = .012$. The flower–insect IAT was not significantly correlated with the oppression IAT, $r(71) = .14$, *ns*, or the racial attitudes IAT, $r(71) = -.15$, *ns*. Controlling for performance on the flower–insect IAT did not alter the correlation between the racial attitudes and oppression IATs, partial $r = .32$, $p = .007$.

Discussion

Research using implicit measures reveals that most White Americans have negative automatic associations with African Americans (e.g., Fazio et al., 1995; Greenwald et al., 1998; Nosek et al., 2002). According to theories of automatic prejudice (Dovidio et al., 1997; Gaertner & Dovidio, 1986; Greenwald & Banaji, 1995), this is the result of our culture conditioning us with prejudicial biases against members of stigmatized groups.

Part of the reason why White Americans automatically associate African Americans with negativity may be that they associate them with oppression, maltreatment and victimization—negative, yet egalitarian associations that acknowledge that African Americans are discriminated against. Consistent with this, Study 1 found that, in a sample of White college students, automatically associating African Americans with oppression correlated positively with associating them with "Bad."

It is interesting to compare these results using implicit measures of prejudice to results with explicit measures of prejudice. In a separate data collection, 22 White undergraduates were asked to rate African Americans on the dimensions *good*, *bad*, *negative*, *oppressed*, *victimized*, *like*, *positive*, *dislike*, and *maltreated*. Explicitly perceiving African Americans as oppressed (indexed by averaging the oppressed, victimized, and maltreated items) was negatively related to perceiving African Americans as bad (indexed by averaging the bad, negative, and dislike items), $r(21) = -.51$, $p < .05$, and positively related to seeing them as good (indexed by averaging the good, positive, and like items), $r(21) = .42$, $p < .05$. It appears that while explicitly perceiving African Americans as

oppressed is related to more *positive* explicit attitudes towards African Americans, automatically associating African Americans with being oppressed is related to *greater* automatic prejudice.

What accounts for this striking implicit–explicit dissociation? The explanation for explicit attitudes seems clear-cut: perceiving African Americans as oppressed and as good are both egalitarian sentiments, and it therefore makes sense that they correlate positively. In contrast, the associations tapped by implicit measures appear to lack much inferential complexity. They may simply reflect the sum total of negative and positive associations with the attitude object (Greenwald et al., 2002).

Study 2

Given that the data from Study 1 were correlational, Studies 2 and 3 employed an experimental manipulation in order to make causal inferences. A classical conditioning procedure similar to that used in previous research (Glaser, 1999; Karpinski & Hilton, 2001; Olson & Fazio, 2001) was used to show that creating egalitarian negative associations with a group can cause people to show more automatic prejudice towards that group on an association-based implicit measure. Participants were led to repeatedly associate novel groups (the Noffians and the Fasites; group stimuli borrowed from Glaser, 1999) with words related to either oppression or to privilege. It was expected that conditioning egalitarian negative associations with a group (e.g., Noffians = Oppressed) would lead people to score as automatically prejudiced towards that group.

Method

Participants and design

Forty-five White undergraduates (16 males, 29 females) participated in the study in return for monetary compensation ($5). Participants were randomly assigned to either (1) associate the novel group Noffians with words related to oppression and the novel group Fasites with words related to privilege or to (2) associate Noffians with words related to privilege and Fasites with words related to oppression.

Materials and procedure

Participants were told that they would first be learning the names of members of two fictional groups, then complete a memory task, and finally complete a reaction time task.

During the name learning task, participants were asked to categorize Noffian and Fasite names according to their group membership. The names used to represent the group Noffians were *Alnofka*, *Banofto*, *Cenofmi*, *Denofu*, and *Enofu*. The names used to represent the group Fasites were *Efason*, *Efasu*, *Gifason*, and *Hafaso*. Each name appeared on the screen one at a time and participants responded with the left key ("d") to categorize them as Fasites and the right key ("k") to categorize them as Noffians. A red "X" appeared when an incorrect categorization was made.

During the ostensive memory task, participants viewed 200 pairings of Fasite and Noffian names with words related to oppression and privilege. They were told to remember the number of times each group was paired with each word. In the Noffian = Oppressed condition, Noffian names were paired 25 times each with the words *oppressed, victimized, mistreated*, and *brutalized* and Fasite names were paired 25 times each with the words *privileged*, *rulers*, *dominant*, and *powerful.* In the Noffian = Privileged condition, these associations were reversed. Participants pressed the space bar when they were done studying each pairing, at which point the next pairing appeared.

Finally, all participants completed a Noffian vs. Fasite, Good vs. Bad IAT. To simplify the design, task order was not counterbalanced, such that all participants first completed the Noffian + Good/Fasite + Bad IAT task and then completed the Noffian + Bad/Fasite + Good IAT task.

Results and discussion

Data preparation for the IAT was identical to that in Study 1. To calculate the IAT effect, participants' performance on the Noffian + Bad IAT task was subtracted from their performance on the Noffian + Good IAT task.

The predicted interaction between IAT task and the classical conditioning manipulation was obtained, $F(1,43) = 10.00$, $p = .003$. As seen in Fig. 2, participants were faster to associate Noffians with "Bad" after being conditioned to associate Noffians with oppression, victimization, and discrimination. Consistent with Glaser (1999), there was also a main effect of relatively more positive associations with Fasites than with Noffians, $F(1,43) = 5.99$, $p = .019$. These results indicate that associating a group with oppression can cause people to display prejudice against that group on an implicit measure of attitude.

Study 3

Study 2 leaves open the question of how conditioning people to view a novel group as oppressed might affect their explicit attitudes towards the group. Therefore, Study 3 used a similar design to Study 2 but additionally assessed self-reported attitudes towards Noffians and Fasites. A number of interesting potential results with explicit attitudes could be anticipated. First, it may be that learning to automatically associate a group with negative concepts like oppression, discrimination, and maltreatment leads to explicit disliking of the group.

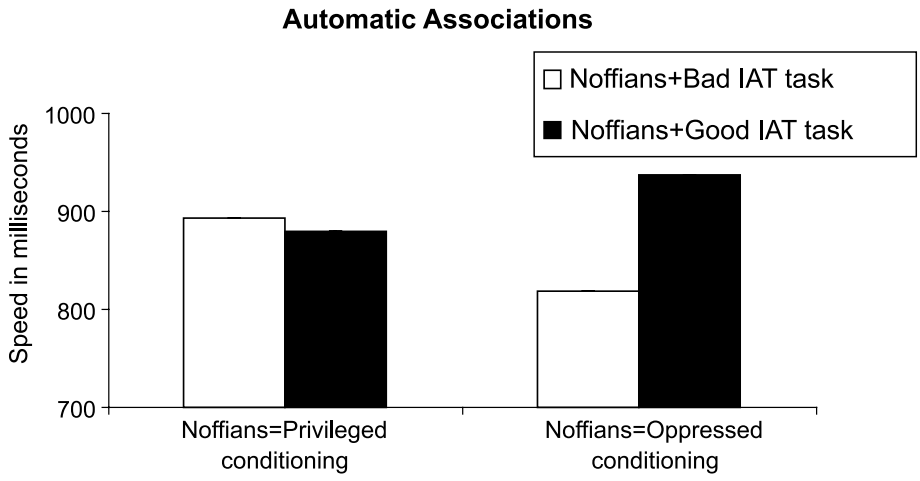

Fig. 2. Automatic associations with the fictional group Noffians after being classically conditioned to associate Noffians with either words related to oppression (an egalitarian negative association) or words related to privilege. The *Y*-axis represents participants' mean response latency in milliseconds for each IAT task. Slower responses during the Noffians + Good IAT task than during the Noffians + Bad IAT task reflect negative automatic associations with Noffians.

Although the negative correlation between perceiving African Americans as oppressed and explicit prejudice reported earlier casts some doubt on this hypothesis, attitudes towards novel groups may operate or develop differently from attitudes towards highly familiar groups. Second, to the extent that the person is aware of the conditioned association, finding out that a group suffers from oppression may elicit sympathy, leading to more positive explicit attitudes towards the group. Third, to the extent that self-reported attitudes develop via learning processes other than classical conditioning, they may show little to no effect of repeatedly associating a group with oppression.

Method

Participants

Fifty-two undergraduates (27 males, 25 females) participated in the study in return for course credit. The majority of participants were White ($n = 49$) while the remainder were Asian ($n = 1$) and African American ($n = 2$).

Design, materials, and procedure

The design, materials, and procedure were with a few exceptions identical to those in Study 2. Participants were again randomly assigned to either (1) associate the novel group Noffians with words related to oppression and the novel group Fasites with words related to privilege or to (2) associate Noffians with words related to privilege and Fasites with words related to oppression. Next, all participants completed a Noffian vs. Fasite, Oppressed vs. Privileged IAT (oppression IAT) and a Noffian vs. Fasite, Good vs. Bad IAT (attitude IAT) in that order. Finally, all participants filled out a self-report questionnaire of their judgments of Noffians and Fasites along the dimensions *good*, *bad*, *privileged*, *negative*, *oppressed*, *victimized*, *like*, *dominant*, *positive*, *dislike*, *maltreated*, and *powerful*.

Results and discussion

Some participants did not complete all of the measures and therefore degrees of freedom vary slightly for different comparisons. Data preparation for the IAT was identical to that in Studies 1 and 2. To calculate the oppression IAT effect, participants' performance on the Noffian + Oppressed IAT task was subtracted from their performance on the Noffian + Privileged IAT task. To calculate the attitude IAT effect, participants' performance on the Noffian + Bad IAT task was subtracted from their performance on the Noffian + Good IAT task.

Effects on the oppression IAT evidenced a strong influence of the conditioning manipulation, $F(1,48) = 37.44$, $p < .001$. Participants conditioned to associate Noffians with oppression and Fasites with privilege automatically associated Noffians with oppression on the IAT ($M = 287$ ms, $SD = 290$ ms), $t(24) = 4.95$, $p < .001$. Conversely, participants conditioned to associate Fasites with oppression and Noffians with privilege automatically associated Fasites with oppression on the IAT ($M = -155$ ms, $SD = 215$ ms), $t(24) = -3.61$, $p = .001$.

Effects on the attitude IAT closely replicated Study 2. A main effect of relatively more positive associations with Fasites than with Noffians again emerged, $F(1,48) = 7.12$, $p = .01$. More importantly, the critical interaction between IAT task and the classical conditioning manipulation was obtained, $F(1,48) = 5.13$, $p = .028$. Participants were again faster to associate Noffians with "Bad" after being conditioned to associate Noffians with

Table 1
Correlations between the oppression IAT, attitude IAT, explicit oppression ratings, and explicit group attitudes (Study 3)

	1	2	3	4
1. Oppression IAT				
2. Attitude IAT	.56*** (50)			
3. Explicit oppression ratings	.72*** (44)	.48*** (44)		
4. Explicit group attitudes	.25 (43)	.35* (43)	.21 (41)	

Note. Higher scores reflect associating Noffians with oppression on the oppression measures, negative associations with Noffians on the IAT, and negative self-reported attitudes towards Noffians on the explicit attitude measure.

* $p < .05$.

*** $p < .001$.

oppression than after being conditioned to associate them with privilege (Ms = 135 ms and 10 ms, SDs = 223 and 157, respectively).

Explicit measures were scored to make them as directly comparable to the IAT measures as possible (e.g., as difference score measures of attitudes towards Noffians relative to Fasites). The classical conditioning procedure led participants to explicitly perceive the group associated with oppression as oppressed and the group associated with privilege as privileged, $F(1,44) = 58.10$, $p < .001$. Noffians were explicitly associated with oppression in the Noffians = Oppressed condition ($M = 9.01$, $SD = 9.10$; $t(22) = 4.75$, $p < .001$), and Fasites were explicitly associated with oppression in the Fasites = Oppressed condition ($M = -9.93$, $SD = 7.70$; $t(22) = -6.19$, $p < .001$). In contrast, the conditioning procedure had no effect on explicit attitudes towards Noffians vs. Fasites, $F(1,42) = .06$, $p = .80$.

Table 1 displays the correlations between the oppression IAT, attitude IAT, explicitly perceiving Noffians vs. Fasites as oppressed, and explicit group attitudes. Notably, associating Noffians with oppression on the oppression IAT predicted more negative associations with Noffians on the attitude IAT, $r(49) = .56$, $p < .001$. However, explicitly associating Noffians with oppression did not predict explicit prejudice against Noffians, $r(40) = .21$, $p = .19$. Still, it is notable that this correlation, while not significant, was in the positive direction—in contrast to our earlier study in which explicitly perceiving African Americans as oppressed predicted *less* explicit racial prejudice. While speculative, it seems possible that more detailed knowledge of the historical circumstances that produced a group's low status is needed to evoke strong explicit sympathies with the group.

Mediation of the conditioning effect on the attitude IAT

Additional analyses tested whether changes in oppression associations mediated the effects of the conditioning manipulation on the attitude IAT, as hypothesized. As seen in Table 1, the oppression IAT and explicit perceptions of the groups as oppressed were highly correlated ($r = .72$, $p < .001$), suggesting that they tapped the same construct (presumably oppression associations). This is in no way a problem for the thesis of our paper. Our thesis is that the *attitude* IAT and explicit *attitude* measures tap different constructs that are differentially impacted by egalitarian negative associations. We do not claim that implicit and explicit measures of oppression associations tap separate constructs that differentially mediate the effects of egalitarian negative associations. That said, it was still of interest to examine how well the different measures of oppression associations mediated the effects of the conditioning manipulation on the attitude IAT.

There was a significant effect of the conditioning manipulation on both the oppression IAT, $\beta(48) = .66$, $p < .001$, and on the attitude IAT, $\beta(48) = .31$, $p = .028$. The oppression IAT and attitude IAT were significantly related, $\beta(48) = .56$, $p < .001$. Controlling for scores on the oppression IAT reduced the effect of condition on the attitude IAT to nonsignificance, $\beta(47) = -.11$, $p = .50$. However, the effect of the oppression IAT on the attitude IAT remained significant controlling for condition, $\beta(47) = .63$, $p < .001$.

A significant effect of the conditioning manipulation was also found on explicit oppression associations, $\beta(44) = .75$, $p < .001$. Explicit oppression associations and the attitude IAT were significantly related, $\beta(42) = .46$, $p = .001$. Controlling for explicit oppression associations reduced the effect of condition on the attitude IAT to nonsignificance $\beta(41) = -.22$, $p = .35$. However, the effect of explicit oppression associations on the attitude IAT remained significant controlling for condition, $\beta(41) = .65$, $p = .007$.

These analyses are consistent with the interpretation that both the oppression IAT and explicit oppression association measures mediated the effects of the conditioning manipulation.

Summary

To summarize the primary findings of Study 3, classically conditioning participants to associate a fictional group with oppression lead to more negative automatic associations with that group on an attitude IAT, but did not increase self-reported prejudice towards the group. This further suggests that part of the reason why people have negative associations with low status groups is that they associate such groups with unfortunate circumstances.

General discussion

Why do most White Americans have negative automatic associations with African Americans? One reason is that people have negative, yet egalitarian associations between members of low status groups and oppression, mistreatment, and victimization. There are a number of possible

interpretations of this finding. First, it is possible that perceiving a group as oppressed leads people to automatically dislike that group, a view that finds support in the predictive validity of implicit measures (Florak et al., 2001; Hugenberg & Bodenhausen, 2004; Poehlman et al., 2006). Second, while implicit measures capture the association between negativity and a social group, that association may not reflect dislike in all cases.

Egalitarian negative associations as a source of automatic prejudice

Egalitarian negative associations can be thought of as a potential source of automatic prejudice. This interpretation is strongly suggested by work on the predictive validity of implicit measures. As discussed earlier, not only do automatic association predict biased behaviors against members of minority groups, they significantly out-predict self-reported attitudes (Dovidio et al., 1997; Fazio et al., 1995; Florak et al., 2001; Gawronski, Ehrenberg, Banse, Zukova, & Klaur, 2003; Gawronski, Geschke, & Banse, 2003; Hugenberg & Bodenhausen, 2004, 2003; McConnell & Liebold, 2001; Rudman & Glick, 2001; see Poehlman et al., 2006, for a meta-analysis). Negative affect associated with a group may not be distinguished based on its source, and contribute to a feeling of negativity towards the group even though the affect was not initially encoded as dislike. Indeed, because automatic associations predict prejudiced behaviors, the burden of proof is on those wishing to argue that egalitarian negative associations complicate the assessment of automatic attitudes rather than contribute to prejudiced responses.

At the same time, people may automatically dislike oppressed groups as a consequence of rationalizing inequality. According to System Justification Theory (Jost & Banaji, 1994), people are motivated to uphold the status quo. Regardless of how we consciously feel about inequality, low status groups are automatically disliked simply by virtue of being low status. Thus, the fact that oppressed groups are automatically associated with "Bad" may be evidence for system justification operating in an automatic fashion.

It could also be the case that social judgments are based on an initial negative or positive emotional response, after which complex inferences imbue the initial, automatic response with deeper meaning (see also Schachter & Singer, 1962). For example, most of us would have a negative disgust response to the sight of a homeless person freezing to death on the sidewalk. But after realizing that the negative reaction is really an indicator of empathy and a signal to provide help (Pizarro, 2000), it may gain an altruistic meaning rather than a prejudicial one. Similarly, egalitarian negative associations may contribute to an initial, automatic negative response to members of oppressed groups that, after the opportunity to draw more complex inferences (e.g., "I had a negative reaction when I saw that Black person panhandling because I am against racial inequality"), leads to egalitarian behaviors.

If so, egalitarian negative associations may predict different behaviors depending on a person's capacity to reason about their initial, gut responses (Fazio, 1990). When people are not under cognitive load, egalitarian negative associations may predict antiracist acts such as support for civil rights laws or affirmative action. However, when the ability to make complex inferences is restricted (for example, by high degrees of environmental noise or preoccupation with other thoughts; Gilbert, Tafarodi, & Malone, 1993; Kruglanski & Webster, 1996), associating African Americans with negatively valenced concepts like oppression, victimization, and maltreatment may lead to prejudiced behaviors that are congruent with that negative valence.[2]

Negativity may not reflect dislike in every case

An automatic attitude is operationally defined as the sum total of positive and negative associations with a social target. While automatic associations measured in this way are important predictors of judgments and behaviors, negativity may not reflect dislike in every case. Due to egalitarian negative associations, some egalitarian individuals may appear more automatically prejudiced than they actually are.

Even if such an argument is found to be valid to some extent, the strong version of this argument should be viewed with skepticism. It is extremely implausible that implicit measures of prejudice mostly tap egalitarian negative associations. Other work confirms the validity of implicit measures as predictors of prejudiced judgments and behaviors (Poehlman et al., 2006). Thus, while they may tap egalitarian negative associations to a certain degree, implicit measures tap unambiguously prejudiced associations as well.

Moreover, even if egalitarian negative associations are an additional source of variability within implicit measures, this only points to the potential for such measures to become even more useful when it comes to predicting judgments and behaviors. Statistical methods can be developed to account for these components, improving the validity of implicit measures. For example, if the association between African Americans and "Bad" held by most White Americans stems partly from prejudiced negative associations and partly from egalitarian negative associations, then controlling for scores on the oppression IAT should increase

[2] Following this logic further, it seems possible that consciousness raising exercises aimed at increasing awareness of social oppression may, ironically, strengthen automatic prejudices. For example, viewing scenes of Jewish suffering in *Schindler's List* may increase the accessibility of egalitarian negative associations with Jewish people, ironically leading to more negative responses to Jewish people when conscious capacity is diminished.

the relationship between the racial attitudes IAT and behavioral measures of prejudice.

Concluding comment

The present research suggests that automatic associations have roots in both prejudiced (e.g., "African Americans are stupid, lazy, and violent") and egalitarian (e.g., "African Americans are oppressed, mistreated, and victimized") sentiments. Egalitarian negative associations may contribute to prejudiced reactions, as suggested by recent work on the predictive validity of association-based measures. To the extent that egalitarian negative associations turn out to complicate the assessment of automatic attitudes, controlling for their influence can only improve the assessment of individual differences in automatic social cognition. Investigating these interesting issues will shed additional light on the nature and function of social attitudes.

References

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?". *Psychological Inquiry, 15*, 257–279.

Banaji, M., Nosek, B., & Greenwald, A. (2004). No place for nostalgia in science: a response to Arkes and Tetlock. *Psychological Inquiry, 15*, 279–289.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: ruling out the familiarity explanation. *Journal of Experimental Social Psychology, 36*, 316–328.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.

Fazio, R., Jackson, J., Dunton, B., & Williams, C. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.

Fazio, R. H., Sanbonmatsu, D. M, Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: the MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York: Academic Press.

Florak, A., Scarabis, M., & Bless, H. (2001). When do associations matter?: the use of implicit associations toward ethnic groups in person judgments. *Journal of Experimental Social Psychology, 37*, 518–524.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). San Diego, CA: Academic Press.

Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klaur, K. C. (2003). It's in the mind of the beholder: individual differences in associative strength moderate category based and individuating impression formation. *Journal of Experimental Social Psychology, 39*, 16–30.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: associations influence the construal of individuating information. *European Journal of Social Psychology, 33*, 573–589.

Glaser, J. (1999). The relation between stereotypes and prejudice: a brief history and evidence from newly formed automatic associations. Unpublished doctoral dissertation, Yale University.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology, 65*, 221–233.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: the role of prejudice and facial affect in race categorization. *Psychological Science, 15*, 342–345.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640–643.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1–27.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774–788.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "seizing" and "freezing". *" Psychological Review, 103*, 263–283.

McConnell, A. R., & Liebold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442.

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition, 19*, 625–666.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics, 6*, 101–115.

Nosek, B.A. (2002). Moderators of implicit–explicit correspondence. Unpublished doctoral dissertation, Yale University.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*, 413–417.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653–667.

Pizarro, D. (2000). Nothing more than feelings? The role of emotions in moral judgment. *Journal for the Theory of Social Behaviour, 30*, 355–375.

Poehlman, T.A., Uhlmann, E., Greenwald, A.G., & Banaji, M.R. (2006). Understanding and using the Implicit Association Test: 3. Meta-analysis of predictive validity. Manuscript under review.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743–762.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379–399.

Sears, D. (1988). Symbolic racism. In P. Katz & D. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53–84). New York: Plenum.

Shah, J. Y., Kruglanski, A. W., & Thompson, E. P. (1998). Membership has its (epistemic) rewards: need for closure effects on in-group bias. *Journal of Personality and Social Psychology, 75*, 383–393.

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

van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgement and memory of a criminal act: the effects of stereotypes and cognitive load. *European Journal of Social Psychology, 29*, 191–201.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262–274.