# Recognizing discrimination explicitly while denying it implicitly: Implicit social identity protection ☆

Jennifer M. Peach [a,*], Emiko Yoshida [a], Steven J. Spencer [a], Mark P. Zanna [a], Jennifer R. Steele [b]

[a] Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada
[b] Department of Psychology, York University, Toronto, Ontario, Canada

A B S T R A C T

Past research suggests that members of devalued groups recognize their groups are discriminated against. Do the implicit responses of members of these groups demonstrate the same pattern? We argue that they do not and that this is due to a motivated protection of members of devalued groups' social identity. Study 1 demonstrates that, at an explicit level African-Canadians recognize that their group is discriminated against, but at an implicit level African-Canadians think that most people like their group to a greater extent than do European-Canadians. Study 2 replicates this implicit finding with Muslim participants while demonstrating that, when affirmed, this group difference disappears. Study 3 demonstrates that implicit normative regard can predict collective action over and above implicit attitudes and explicit normative regard. The implications for changing the status quo are discussed.

© 2010 Elsevier Inc. All rights reserved.

Members of devalued groups are often aware of the discrimination that their groups face (Crocker, Luhtanen, Blaine, & Broadnax, 1994) and the recognition of this pervasive discrimination can negatively impact their self-esteem (Branscombe, Schmitt, & Harvey, 1999). Yet, according to social identity theorists (Tajfel & Turner, 1979), members of these groups are motivated to have a positive social identity. How do members of devalued groups reconcile their desire to have a positive social identity with the social reality that their group is not regarded positively? We suggest that members of devalued groups' motivation to maintain a positive social identity is limited by the reality of the inequality that their group faces, making it difficult for members of such groups to believe that others view their group positively at an explicit level. We reason, however, that this motive might be evident when others' regard is measured implicitly (what we call implicit normative regard) and predict that it will be a potent predictor of behavior.

Theorizing and research on maintaining a positive social identity has almost exclusively focused on maintaining this identity at a conscious level. Such maintenance strategies have included exiting one's group to join a higher status group (Tajfel & Turner, 1979), forming a positive attitude towards one's group, perhaps by devaluing domains in which one's group does not perform well, or engaging in collective action in an effort to improve the actual social standing of one's group (Tajfel & Turner, 1979). Because using these methods of social identity maintenance may require consciously recognizing that one's group has low status in society, however, they may all be restricted by reality. For example, it may be difficult for Blacks to engage in collective action to improve their group's condition if they believe such actions can never lead to advancement of their group due to the prevalence of discrimination. Similarly, it is difficult for members of devalued groups to explicitly believe that most people value their group while also recognizing that their group is discriminated against, making it challenging to use this method to maintain a positive social identity at an explicit level.

In contrast, when beliefs about how most people view one's devalued group are measured outside of awareness (i.e., implicitly, Yoshida, Peach, Spencer, & Zanna, 2010), we believe that there may be little conflict between explicitly believing that one's group is devalued and implicitly believing that one's group is valued. But can the motive to maintain a positive social identity affect implicit processes? There is reason to believe that it can. According to Bargh's automotive model (Bargh, 1990, 2007; Chartrand & Bargh, 2002) whenever people repeatedly pursue a goal in the presence of a social cue, this goal can be automatically activated (i.e., through implicit processes). Based on this theorizing we reason that if members of devalued groups repeatedly pursue the goal of perceiving their group as valued, for example, in the presence of members of their group, then the goal to maintain a positive social identity would be automatically activated (i.e., through implicit processes) whenever they encounter members

of their group.[1] This reasoning suggests that people may develop positive implicit normative regard about their groups without this positive implicit normative regard conflicting with negative explicit normative regard.

We refer to this construct as implicit normative regard because we wish to draw a similarity between our concept and public regard (e.g., Luhtanen & Crocker, 1992) but also want to highlight the normative nature of our measure. In other research (Spencer, Peach, Yoshida, & Zanna, 2010; Yoshida, et al., 2010) we have suggested that in order to function in society, individuals need to have a readily-accessible (or, implicit) sense of how others react to social groups, their own included. Based on this theorizing and research, we reason that implicit normative regard is related to, but not redundant with implicit attitudes, which measure personal associations with social groups. Implicit normative regard and implicit attitudes should be related because they both assess automatic associations with the same social groups. For example, individuals' implicit associations between what society likes and feminists should be related to individuals' implicit associations between what they personally like and feminists because both are likely to be influenced by character-istics of feminists. In addition, over time implicit normative regard and implicit attitudes are likely to influence one another, thus increasing the extent to which they are related.

Our theorizing suggests that these constructs will not be redun-dant, however, for two reasons. First, implicit attitudes and implicit normative regard may have different antecedents. Based on the research of Fazio and his colleagues (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Fazio & Williams, 1986; Olson & Fazio, 2004) we argue that implicit attitudes often form from direct experience with members of social groups. In contrast, we argue that implicit normative regard is derived from exposure to how groups are depicted and treated in society. For example, when participants were exposed to a negative joke about Muslims at which the audience laughed they had more negative implicit normative regard towards Muslims than when the audience did not laugh. The audience's reaction, however, did not influence participants' implicit attitudes towards Muslims (Yoshida et al., 2010). We are not arguing, however, that implicit normative regard is formed simply by passively receiving cultural media depictions of one's group. Instead, we theorize that people actively develop associations between their group and society's view of their group and that their social identity influences the development of these associations. From this perspective cultural portrayals of one's group provide the clay, but social identity related motives sculpt the final shape of implicit normative regard.

Second, we suggest that implicit normative regard and implicit attitudes are related but not redundant because, at least at times, individuals may be motivated to not conform with society (McClelland, Atkinson, Clark, & Lowell, 1953; Thrash, Elliot, & Schultheiss, 2007). This motivation may lead them to adopt different perspectives and behavior, creating different individual associations with social groups than societal associations with these same groups. For example, an individual who is allergic to flowers may know others like flowers (i.e., they may have positive implicit normative regard towards flowers), but may have negative implicit attitudes towards flowers themselves.

In order to measure a more personalized implicit attitude, past researchers (Olson & Fazio, 2004) have modified the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998) to include the category labels "I like" and "I don't like." Following this assessment technique, to measure implicit normative regard, we have

modified the Implicit Association Test to assess the association between what most people like or do not like and social groups (Spencer et al., 2010; Yoshida et al., 2010). In this measure, stronger associations between a social category (such as "Black") and the category "most people like" (and between the categories "White" and "most people don't like") indicate that an individual has relatively positive implicit normative regard towards the social category "Black." Specifically, we believe that when two concepts are associated in memory and these two concepts share a response key, participants are faster to respond, and when these concepts are not associated, participants are slower to respond. Thus, our IAT merely measures associations between the concept "most people like" and "most people don't like" and corresponding social groups.

Although we cannot review all validating data here, we have evidence that this measure is influenced by normative information, predicts behavior, and is distinct from implicit attitudes and explicit normative regard. As previously mentioned, in other research (Yoshida et al., 2010), we have found that whether or not an audience laughed at a racist joke about Muslims impacted participants' implicit normative regard towards Muslims. Specifically, when the audience laughed at the joke, implicit normative regard towards Muslims became more negative than when the audience did not laugh at this joke, indicating that normative information can influence this measure of implicit normative regard. In addition, the more negative their implicit normative regard towards Muslims became, the less money participants allocated to a Muslim student organization, indicating that implicit normative regard can predict meaningful behaviour. Moreover, audience reactions did not influence partici-pants' implicit attitudes, which also did not predict discriminatory behavior, indicating that implicit attitudes and implicit normative regard are distinct (Yoshida et al., 2010).

In the present research, we utilize this measure to test our prediction that members of devalued groups will show more positive normative regard on implicit measures than on explicit measures. Based on social identity theory (Tajfel & Turner, 1979) we argue that members of devalued groups are motivated to maintain positive implicit normative regard about their group, but the reality of discrimination impedes this motivation at an explicit level. Stating explicitly that most people like one's group when often confronted with discrimination is a difficult proposition to maintain when consciously considered. In contrast, at an implicit level when such contradictions are not consciously considered, maintaining a belief that one's group is viewed positively should be easier. We test this possibility in Study 1.

Although the reality of discrimination may make it difficult for members of devalued groups to maintain positive explicit normative regard about one's group, it should not affect their ability to like their group. We therefore suggest that members of devalued groups will have positive explicit attitudes, as well as positive implicit attitudes towards their group. We also test this possibility in Study 1.

We suggest that people maintain positive implicit normative regard in order to protect their social identity, but there may be many ways that a positive social identity can be maintained. Because positive implicit normative regard is influenced by members of devalued groups' motivation to see their group positively and consequently to see themselves positively, when this motivation is satisfied using another means (such as through an affirmation (Steele, 1988)), we predict that members of these groups will have less positive implicit normative regard toward their group. We test this possibility in Study 2 by affirming participants' group identity and then assessing their implicit and explicit normative regard. We expect that, because implicit normative regard is used to affirm members of devalued groups' social identity, an affirmation will actually lead to *less* positive implicit normative regard.

Finally, we suggest that implicit normative regard can predict meaningful behavior, specifically collective action. We further suggest

---

[1] Of course, there may be extreme situations in which members of some groups (such as Jews during the Holocaust) may be unable to deny the discrimination towards their group even at an implicit level, and so it may be the case that not all devalued groups show this effect, but in general we believe the pattern outlined above will be true.

that implicit normative regard can predict collective action above and beyond implicit attitudes and explicit normative regard. We test these predictions in Study 3.

## Study 1

In this study, we chose African-Canadians as our target group because they are often the targets of prejudice and discrimination (Taylor, 1997). We assessed African-Canadians' implicit and explicit attitudes and implicit and explicit normative regard towards African-Canadians. We predicted that their implicit normative regard would be more positive than their explicit normative regard. Furthermore, we compared their responses to those of European-Canadians in order to test our prediction that African-Canadians would have more positive implicit and explicit attitudes, and more positive implicit normative regard, than European-Canadians.

### Method

#### Participants

Eighty-nine undergraduate students at the University of Waterloo and York University participated in exchange for course credit towards their introductory psychology class, or for payment of $8.00 Canadian. Our African-Canadian sample included participants who were born in Canada (33) and participants who were not (37).[2] Nineteen participants were European-Canadians born in Canada. In total, there were 21 men, 66 women, and 2 who did not indicate their gender (aged 17–45, mean age 22.61).

#### Materials

*Implicit attitudes.* Implicit attitudes were assessed using a modified version of the Implicit Association Test (IAT, Olson & Fazio, 2004). Participants were asked to distinguish between "things you might like or dislike," and the category labels were "I like/I don't like." Participants were also asked to distinguish between faces of Black and White targets (Gunz, 2004) and the category labels were "Black/White."

The IAT contained five blocks of trials. The first block was a practice block, in which participants categorized words such as "party" and "disease" to the evaluative category labels "I like/I don't like." The second block was also a practice block, in which participants categorized the faces of Black and White targets using the category labels "Black/White." The third block was a critical block, in which the category labels, "I like" and "Black" and "I don't like" and "White," shared the same response key. The fourth block was another practice block, in which participants had to categorize Black and White targets, and the category labels were reversed from Block 2 and 3. The fifth block was also a critical block, in which the category labels, "I like" and "White" and "I don't like" and "Black," shared the same response key.

*Implicit normative regard.* The measure of implicit normative regard was the same as the implicit attitude measure except that participants were asked to distinguish between "things that most people like or dislike," and the words were characterized as referring to "what people in North America actually like, not what they should like." The category labels "most people like/most people don't like" replaced the category labels "I like/I don't like." All other aspects of the task were the same as the implicit attitude measure (Yoshida et al., 2010).

*Explicit attitudes and normative regard.* Participants completed several semantic differentials. They completed a semantic differential assessing their attitudes towards Black people, from 1 (*I like extremely*) to 7 (*I dislike extremely*), and completed a comparable semantic differential assessing their attitudes towards White people. They also completed two normative semantic differentials assessing their perceived normative regard towards Black people, and towards White people, from 1 (*most people like extremely*) to 7 (*most people dislike extremely*).

*Perceived bias in the assessment of merit (Son Hing, Bobocel, & Zanna, 2002).* This 22-item scale assessed participants' perception that visible minorities are disadvantaged in the workforce. An example item from this scale is "historically, subtle personal biases of job interviewers disadvantaged visible minority job applicants in the assessment of their qualifications." The item responses ranged from 1 (*strongly disagree*) to 7 (*strongly agree*) (Cronbach's $\alpha = .89$).

#### Procedure

Participants completed two online sessions, each 4 to 7 days apart, in a counterbalanced order. They were told the purpose of the study was to assess their reactions to various objects and social groups. IATs assessing implicit attitudes and implicit normative regard were completed during separate online sessions, to reduce potential carry-over effects. Explicit attitudes and explicit normative regard were collected in the same session as the comparable IATs. Not all participants completed all measures, and so we included all possible data in each analysis and did not exclude participants with missing data. Once participants completed the sessions, they received partial course credit or payment, and were debriefed and thanked for their participation.

### Results and discussion

#### Calculation of IAT scores

We calculated scores on the IAT following procedures described by Jordan, Jordan, Spencer and Zanna (2005).[3] We recoded scores below 300 ms to 300 ms, and scores above 3000 ms to 3000 ms. Errors trials were excluded from calculations. We then subtracted the reaction times obtained in the second critical block (Block 5) from the first critical block (Block 3). If the category labels "I like" and "White" and "I don't like" and "Black" are highly associated, participants will tend to be faster in Block 5 than Block 3. We then reverse scored this measure so that high scores indicated a stronger association between the positive evaluative label (i.e., "I like") and "Black," and between the negative evaluative label (i.e., "I don't like") and "White."

#### Relation of group membership to implicit and explicit attitudes

Recall we predicted that African-Canadians would have both more positive explicit and implicit attitudes towards Blacks than would European-Canadians. Consistent with our predictions, African-Canadian participants had significantly more positive implicit attitudes towards Blacks ($M = 42.89$, $SD = 164.67$) than did European-Canadian participants ($M = -104.17$, $SD = 129.19$), $F(1, 87) = 12.95$, $p = .001$. Is it the case, however, that this difference in implicit attitudes is driven by the relation between this implicit attitudes measure and another construct, such as participants' explicit attitudes, or their implicit or explicit normative regard? To test whether this difference was independent of participants' implicit and explicit normative regard and their explicit attitudes we controlled for

---

[2] Of the participants not born in Canada, 6 were born in African countries, 4 from Caribbean countries, and 1 from Italy. The rest of the participants indicated they were not born in Canada, but did not indicate in which country they were born. There were no significant differences between African-Canadian participants who were born in Canada and those not born in Canada on any measures, so we collapsed across country of origin when analyzing data.

[3] If these data are scored using Greenwald et al. (2003) scoring method, the results of the studies are essentially the same, except that the interaction in Study 2 is no longer significant, $F < 1$, although the means show the same pattern as the results in Study 2.

these variables and found that this difference in implicit attitudes remained significant, $\beta = .37$, $t(79) = 3.50$, $p = .001$.

To make the explicit measures more comparable to the implicit measures we computed scores so that higher scores indicated more positive attitudes towards Blacks.[4] African-Canadian participants had marginally more positive attitudes towards Blacks vs. Whites ($M = .46$, $SD = 1.06$) than did European-Canadian participants ($M = .05$, $SD = .52$), $F(1, 85) = 2.57$, $p = .11$. Thus, African-Canadians tended to have more positive implicit and explicit attitudes towards Blacks than did European-Canadians.

We next tested whether there was a significant difference between implicit and explicit attitudes among African-Canadians and among European-Canadians. Because implicit and explicit attitudes (and implicit and explicit normative regard) are on different scales, we recoded each measure to standard deviation units.[5] Further, because we were comparing these measures within race we computed the standard deviation for each race separately. We then conducted a 2 (attitude: implicit vs. explicit) × 2 (race: African-Canadian vs. European-Canadian) mixed measures ANOVA. There was a main effect of type of attitude, such that explicit attitudes ($M = .34$) were more positive towards Blacks than implicit attitudes ($M = .03$), $F(1, 85) = 10.40$, $p = .002$. There was also a main effect of race, such that African-Canadian participants ($M = .34$) had more positive attitudes than European-Canadian participants, ($M = -.35$), $F(1, 85) = 11.73$, $p = .001$. These main effects were qualified by a two-way interaction, $F(1, 85) = 5.41$, $p = .02$. African-Canadian participants did not have significantly different implicit and explicit attitudes ($M(\text{implicit}) = .27$; $M(\text{explicit}) = .41$), $F < 1$. European-Canadian participants, however, had more positive explicit ($M = .10$) than implicit attitudes towards Blacks ($M = -.81$), $F(1, 85) = 9.85$, $p = .002$.

*Relation of group membership to implicit and explicit normative regard*

We predicted that, at an implicit level, African-Canadians would have more positive implicit normative regard towards Blacks (vs. Whites) than would European-Canadians, but that, due to the reality of discrimination that African-Canadians face, at an explicit level the opposite would be true. Consistent with our predictions, African-Canadian participants had more positive implicit normative regard towards Blacks ($M = 42.71$, $SD = 168.87$) than did European-Canadian participants ($M = -90.01$, $SD = 203.81$), $F(1, 83) = 8.01$, $p = .006$. To test whether this difference in implicit normative regard could be due to African-Canadians and European-Canadians' different implicit or explicit attitudes or explicit normative regard, we controlled for these variables and found that the differences in implicit normative regard still held, $\beta = .26$, $t(80) = 2.19$, $p = .03$. We calculated explicit normative regard towards Blacks using the same methodology used for explicit attitudes described above. African-Canadian participants reported marginally more negative explicit normative regard towards Blacks vs. Whites ($M = -1.21$, $SD = 1.70$) than did European-Canadian participants ($M = -.44$, $SD = 1.04$), $F(1, 84) = 3.28$, $p = .07$. African-Canadian participants also perceived more discrimination against visible minorities ($M = 5.21$, $SD = .81$) than did European-Canadian participants ($M = 4.39$, $SD = .66$), $F(1, 84) = 15.38$, $p = .0001$. Our results supported our predictions. Implicit normative regard towards Blacks versus Whites was more positive for African-Canadians than it was for European-Canadians, but the

opposite was true for explicit normative regard. Although implicit attitudes and implicit normative regard showed a similar pattern of results across participants and were significantly correlated, $r(83) = .23$, $p = .03$, based on the magnitude of the relation it is clear that the two measures were not redundant.

We again assessed whether the pattern of results for implicit normative regard was significantly different from that of explicit normative regard by calculating standard deviation units for implicit and explicit normative regard within each racial group and then conducting a 2 (normative regard: implicit vs. explicit) × 2 (race: African-Canadian vs. European-Canadian) mixed-measures ANOVA. There was a main effect of type of normative regard, such that implicit normative regard ($M = .10$) was more positive than explicit normative regard ($M = -.66$), $F(1, 83) = 7.26$, $p = .009$. This main effect, however, was qualified by the predicted two-way interaction, $F(1, 83) = 7.72$, $p = .007$. African-Canadian participants had more positive implicit ($M = .25$) than explicit normative regard ($M = -.72$), $F(1, 83) = 35.34$, $p < .0001$. The implicit and explicit normative regard of European-Canadian participants did not differ ($M(\text{implicit}) = -.44$; $M(\text{explicit}) = -.43$), $F < 1$.

Thus, our results are consistent with our hypotheses–African-Canadian participants had more positive implicit normative regard about African-Canadians than did European-Canadian participants–supporting our assertion that African-Canadians are motivated to have a positive social identity as assessed by their responses on an implicit measure. Contrary to these implicit findings, but consistent with our hypothesis that the reality of discrimination constrains expressions of explicit normative regard, we found that African-Canadian participants tended to have more negative explicit normative regard about African-Canadians than did European-Canadians. African-Canadians also had more positive explicit and implicit attitudes towards African-Canadians than did European-Canadians, as predicted.

*Replication of findings*

We sought to replicate our implicit findings with a different group in order to increase our external validity. We chose feminists because they are not a visually identifiable group but are still devalued (Haddock & Zanna, 1994). We reasoned that such a non-visually identifiable group might be a particularly rigorous test of our hypothesis. Seventy-seven women participated for course credit.[6] We assessed participants' implicit normative regard about feminists by using the category labels "most people like/most people don't like" (to which participants categorized the same words as the IATs above) and "feminist/housewife" (to which participants categorized stereotyped words such as activist to feminist and domestic to housewife). We found that identification with feminism among women (assessed using three items, Reid & Percell, 2004)[7] was correlated with implicit normative regard, $r(73) = .40$, $p = .0001$, such that the more women identified with feminism the more they implicitly believed that most people like feminists. Identification with feminism was also correlated with implicit attitudes towards feminists, $r(74) = .41$, $p = .001$.[8] Thus,

---

[4] We calculated the explicit measures by subtracting the scores for Blacks from the scores for Whites. Because low numbers indicates greater liking for the group, higher numbers on these measures indicates more positive explicit attitudes and explicit normative regard for Blacks relative to Whites.

[5] These scores were in standard deviation units meaning that one scale point represents one standard deviation unit for each measure, and the mean was not centered at zero but rather was the difference between the arithmetic mean and zero expressed in standard deviations. We did not standardize these variables (i.e., recoded in standard deviation units with a mean of zero) because comparing two measures with means of zero would not be a meaningful comparison.

[6] Two participants did not complete all measures.

[7] As in Study 3, we assessed women's identification with feminism using the items "I am a feminist," "I am strongly identified with being a feminist," and "being a feminist is central to who I am" (Reid & Percell, 2004). Because few women in our sample were highly identified with feminism we over-sampled women who scored above the midpoint on all three identification items.

[8] We could not control for implicit attitudes when determining whether identification with feminism predicted implicit normative regard (and vice versa) because identification with feminism and implicit attitudes interacted when predicting implicit normative regard, $\beta = 2.91$, $t(70) = 3.15$, $p = .002$. For women who were highly identified with feminism, implicit attitudes and implicit normative regard were not related, $\beta = .16$, $t(70) = 1.09$, $p = .28$, but for women who did not identify with feminism their implicit attitudes and implicit normative regard were highly related, $\beta = .70$, $t(70) = 5.74$, $p < .0001$. This indicates that, as we might predict, there is a looser connection between implicit attitudes and implicit normative regard for members of a devalued group.

the responses of women who voluntarily identify with a devalued group also support our hypothesis that members of devalued groups are motivated to protect their social identity, replicating our findings with African-Canadian participants.

In addition we replicated this finding among members of another devalued group, that is, among Asians.[9] Using a slightly different measure of implicit normative regard (in which the category labels were "people should like/people shouldn't like" and "Asian/White") we found that Asian-Canadian participants had more positive implicit normative regard towards Asians than did European-Canadian participants, ($M$ (Asian-Canadian) $= -51.24$, $SD = 158.57$; $M$(European-Canadian) $= -196.77$, $SD = 257.71$), $F(1, 103) = 9.76$, $p = .002$. This difference held when we controlled for implicit and explicit attitudes towards Asians, stereotypes towards Asians, and explicit normative regard, $\beta = .33$, $t(82) = 3.12$, $p = .003$, indicating that this difference in implicit normative regard is not due to participants' differential implicit or explicit attitudes or explicit normative regard towards their group. Asian-Canadians also had more positive implicit attitudes towards Asians than did European-Canadians, ($M$(Asian-Canadian) $= -16.76$, $SD = 160.51$; $M$(European-Canadian) $= -199.10$, $SD = 139.43$), $F(1, 106) = 38.17$, $p < .0001$. This difference held when we controlled for implicit and explicit normative regard, explicit attitudes, and stereotypes towards Asians, $\beta = .34$, $t(82) = 3.43$, $p = .001$, again indicating that this difference is independent of differences in these variables. Interestingly, members of devalued groups do not differ from members of the majority on implicit normative regard towards all social groups. On an IAT with the category labels "old/young" Asian-Canadian and European-Canadian participants did not have different implicit attitudes, ($M$(European-Canadian) $= -146.64$, $SD = 149.35$; $M$(Asian-Canadian) $= -157.68$, $SD = 142.83$), $F < 1$, or different implicit normative regard towards the elderly, ($M$(European-Canadian) $= -167.21$, $SD = 172.57$; $M$(Asian-Canadian) $= -195.03$, $SD = 209.79$), $F < 1$.

Thus, we have replicated our implicit findings among members of three different devalued groups; African-Canadians, feminists, and Asian-Canadians. These groups had more positive implicit attitudes and implicit normative regard about their group than did members of non-devalued groups. Participants' implicit normative regard was consistent with their explicit attitudes towards their group, but discrepant from their explicit normative regard. We have suggested that this is motivated by a desire to protect their social identity. These studies have provided correlational evidence supporting our claim but have not provided a direct test of this motivation. Thus, in Study 2 we assess whether an affirmation (Steele, 1988) can reduce identity protection.

## Study 2

We have suggested that members of devalued groups maintain positive implicit normative regard in order to maintain a positive social identity. If this is the case, when this motive is met through an affirmation (Steele, 1988), they should no longer show positive implicit normative regard. We chose to affirm participants' group identity because we hypothesized that this type of affirmation would most effectively meet the motivation to protect their social identity.

To test this possibility, we selected members of a devalued group (i.e., Muslims) and non-Muslims, and randomly assigned them to a group-based affirmation or a control condition. Our first hypothesis was that, when they were not affirmed, Muslims would have more positive implicit normative regard about Muslims than would non-

Muslims. Our second hypothesis was that, when affirmed, Muslims and non-Muslims would no longer have different implicit normative regard about Muslims.

Because maintaining the belief that one's group is discriminated against at an explicit level might help members of devalued groups to maintain a positive sense of self, we thought that affirmation might have a different effect on explicit normative regard. Indeed, based on research by Crocker and Major (1989) it is likely that members of devalued groups use the knowledge that their group is discriminated against to cope with threats to their self-concept. If this is true, we would expect that an affirmation would reduce reports of discrimination against one's group (i.e., would lead to more positive explicit normative regard for Muslim participants). We tested this prediction in this study as well.

### Method

#### Participants

A total of 67 participants (32 Muslim, 35 non-Muslim; 35 women, 30 men, 1 did not indicate gender; $M$(age) $= 20$ years old) participated in exchange for course credit or for $8. Non-Muslim participants were born in Canada, were White, and either identified as Christian or did not identify as religious.

#### Procedure

Participants were brought into the lab in groups of one to three. They were randomly assigned to the group affirmation or control condition. In the group affirmation condition, participants were asked to think about a social group that was most important to them (Luhtanen & Crocker, 1992), and then to think about a value that was most important to their group. They could choose from: business/economics, social life/relationships, art/music/theatre, science/pursuit of knowledge, or helping those in need (Fein & Spencer, 1997). Next, participants were asked to come up with three reasons why this value was important to their group, and an example of how their group has demonstrated this value (Sherman, Kinias, Major, Kim, & Prenovost, 2007).

In the control condition, participants were not asked to think about a group that was important to them. Instead, they were asked to choose the value that was least important to them and were asked to list three reasons why this value might be important to someone else, and how someone might demonstrate this value. We did not have participants think about a group that is important to them and then rate the value that is least important to their group because we reasoned that thinking about an important group membership might be affirming.

Next, to establish the same norm for both Muslims and non-Muslims all participants watched a series of comedy clips, the last one containing a joke about racial profiling of Muslims (Yoshida et al., 2010). Participants then completed the implicit normative regard measure and then the explicit measure of normative regard towards Muslims, and finally reported what social group they had thought about earlier in the study.

### Materials

*Implicit normative regard.* Similar to Study 1, this IAT had the category labels "most people like/most people don't like." In this study, however, the other category label was "Muslim/neutral."[10] In the

---

[9] We selected Asian-Canadian participants who were born in Canada or had lived in Canada for at least 5 years. In total there were 139 participants (87 European-Canadian, 52 Asian-Canadian, 35 male, 104 female, ages ranged from 17–46, mean age $= 20.07$).

[10] The neutral items used in this IAT were pictures of a stapler, a table, a fork, and a chair.

IATs in previous studies, we measured participants' associations with two groups at once (i.e., Black and White). Some researchers have suggested that measuring associations with two groups at once renders it unclear to which group the associations are made (Blanton, Jaccard, Gonzales, & Christie, 2006). To make it clear in the present study that associations are with Muslims we changed the other category label from another social group (i.e., White) to neutral. This change in procedure makes it clear that any evidence of identity protection in this study is due to ingroup favoritism (i.e., about associations with Muslims for the Muslim participants) rather than outgroup derogation. Higher scores indicated a stronger association between "most people like" and "Muslim."

*Explicit normative regard toward Muslims.* Six questions assessed the extent to which people in Canada or North America like, respect, and support Muslims, and the extent to which Muslims face discrimination in Canada. An example item is "most Canadians support Muslims." Agreement with these items was assessed on a 1 *(strongly disagree)* to 7 *(strongly agree)* scale (Cronbach's $\alpha = .83$).

### Results and discussion

#### Group selected for group affirmation

In this study we did not direct participants to think about their religious identity because we were concerned that such instructions might alert participants to the purpose of our study. At the end of the study, however, we did ask them what social group they had thought about when asked to think about a group that was important to them. Only 4 of the 17 Muslim participants in the affirmation condition listed Muslim as their most important group.

#### Implicit normative regard

Our first prediction was that Muslim participants who were affirmed would show less positive implicit normative regard for their group. To test this hypothesis we conducted a 2 (group: Muslim vs. non-Muslim) × 2 (condition: affirmation vs. control) between-subjects ANOVA, with implicit normative regard as the dependent variable. There was a main effect of group, $F(1,61) = 15.59$, $p = .0001$, such that Muslims had more positive implicit normative regard about Muslims than did non-Muslims ($M = 65.81$, $SD = 191.84$; $M = -106.46$, $SD = 172.61$ respectively). As depicted in Fig. 1, this main effect was qualified by the predicted interaction, $F(1, 61) = 4.63$, $p = .04$. In the control condition, Muslim participants had more positive implicit normative regard towards Muslims ($M = 127.01$, $SD = 192.47$) than did non-Muslim participants ($M = -143.83$, $SD = 181.80$), $F(1, 61) = 18.18$, $p = .0001$, but when both groups were affirmed this group difference was no longer significant ($M(Muslim) = 15.41$, $SD = 181.45$; $M(non-Muslim) = -64.41$, $SD = 156.61$), $F(1, 61) = 1.65$, *ns*.
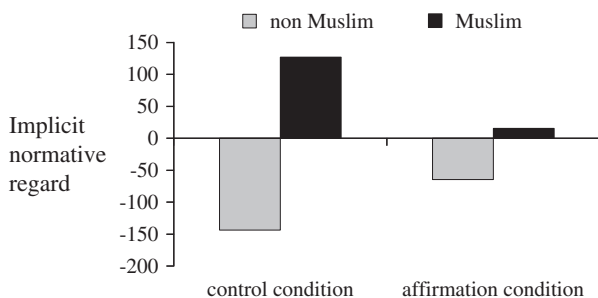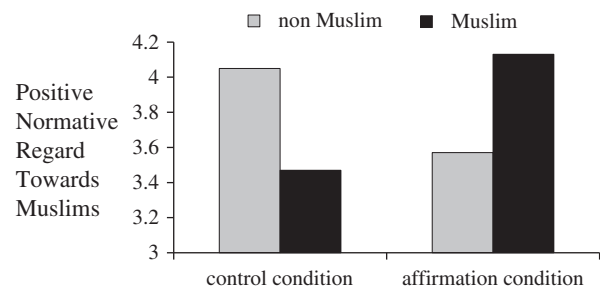
Looking at the results in a different way, when Muslims were affirmed, they tended to have more negative implicit normative regard towards Muslims than when they were not affirmed, $F(1, 61) = 3.01$, $p = .09$. The implicit normative regard of non-Muslim participants did not differ whether they were affirmed or not, $F(1, 61) = 1.68$, *ns*.

#### Explicit normative regard towards Muslims

We assessed the effect that affirmation had on explicit normative regard about how Muslims are viewed in North America. We again conducted a 2 (group: Muslim vs. non-Muslim) × 2 (condition: affirmation vs. control) between-subjects ANOVA. There was a significant interaction between group and condition, $F(1, 60) = 5.85$, $p = .02$. As depicted in Fig. 2, in the control condition the results showed the opposite pattern from the implicit results, such that when they were not affirmed, Muslims tended to have more negative explicit normative regard towards Muslims ($M = 3.47$, $SD = 1.34$) than did non-Muslims ($M = 4.05$, $SD = .61$), $F(1, 60) = 2.89$, $p = .09$. When they were affirmed, Muslims tended to have more positive explicit normative regard towards Muslims ($M = 4.13$, $SD = 1.08$) than did non-Muslims ($M = 3.57$, $SD = .55$), $F(1, 60) = 2.96$, $p = .09$.

Analyzed in a different way and consistent with research by Crocker and Major (1989), when Muslim participants had not been affirmed, they tended to think that norms towards Muslims were more negative than when they had been affirmed ($M = 3.47$, $SD = 1.34$; $M = 4.13$, $SD = 1.08$ respectively), $F(1, 60) = 3.67$, $p = .06$. When non-Muslim participants had not been affirmed, they thought that norms towards Muslims were more positive than when they had been affirmed ($M = 4.05$, $SD = .61$; $M = 3.57$, $SD = .55$ respectively), $F(1, 60) = 2.22$, $p = .14$, although this difference was not significant.

To assess whether the pattern of results was statistically different between implicit and explicit normative regard, we conducted a 2 (measure: implicit vs. explicit) × 2 (group: Muslim vs. non-Muslim) × 2 (condition: affirmation vs. control) mixed-measures ANOVA using the standard deviation measures. Non-Muslim participants had more negative normative regard overall ($M = -.45$) than did Muslims ($M = .11$), $F(1, 60) = 3.87$, $p < .05$. There was a significant interaction between measure and group, $F(1, 60) = 6.06$, $p = .02$, and a significant three-way interaction between group, measure, and condition, $F(1, 60) = 11.82$, $p = .001$. As described above, when this 3-way interaction is broken down by measure, implicit and explicit normative regard demonstrated the opposite pattern of results, as Muslim participants had more positive implicit normative regard in the control condition but less positive implicit normative regard when affirmed (and non-Muslim participants demonstrated the opposite pattern of results), whereas on explicit normative regard, Muslim participants had more negative normative regard in the



**Fig. 1.** Mean scores on implicit normative regard towards Muslim for both Muslim and non-Muslim participants in the affirmation vs. control group. Higher scores indicate a stronger association between "most people like" and "Muslim."



**Fig. 2.** Mean scores on explicit normative regard towards Muslims for both Muslim and non-Muslim participants in the affirmation vs. control group. Higher scores indicate more positive norms towards Muslims.

control condition than the affirmation condition (and non-Muslim participants again demonstrated the opposite pattern of results).[11]

Study 1 demonstrated that implicit normative regard among two groups was more positive than explicit normative regard, presumably because the reality of discrimination impedes the expression of positive explicit normative regard but not implicit normative regard. In Study 2 we demonstrated that a group affirmation reduces implicit normative regard presumably by bolstering social identity. Up to this point, however, we have not presented convincing evidence that implicit normative regard and implicit attitudes are different constructs that can differentially predict behavior. The purpose of Study 3 was to provide such evidence by examining whether implicit normative regard and implicit attitudes differentially predict collective action.

## Study 3

Social identity theory (Tajfel & Turner, 1979) argues that members of devalued groups can engage in collective action as one way to alter their group status and subsequently protect their group identity. It seems most likely that members of devalued groups would engage in collective action when they are aware of the discrimination that exists against their group. We reason, however, that implicit, but not necessarily explicit, normative regard should predict collective action.

Specifically, we argue that members of devalued groups are most likely to engage in collective action when they have a sense that their group is viewed negatively (i.e., they have negative normative regard) because when the group is viewed negatively, the need for the group to be viewed positively is strongest. If normative regard is explicit, however, it is likely to arise from conscious reflection on experiences of discrimination. Given that their explicit normative regard is rooted in actual experiences of discrimination, recognition of this discrimination may undermine people's motivation to engage in collective action. In the face of discrimination, they may come to believe that collective action simply cannot promote the groups' interests. This means that explicit normative regard should not drive collective action.

In contrast, when members of devalued groups have negative implicit normative regard we reasoned that these associations will predict collective action. Negative implicit normative regard is likely to arise from relatively unconscious and unelaborated associations between one's group and negative evaluations of others (Gawronski & Strack, 2004; Gawronski, Strack, & Bodenhausen, 2009) rather than explicit recognition of discrimination. Because these associations are not well elaborated, they are unlikely to undermine people's motivations to engage in collective action. This means that implicit normative regard will motivate collective action without being undermined by the explicit recognition of the reality of discrimination. Therefore, we expect that implicit normative regard will be a better predictor of collective action than explicit normative regard.

As collective action is directed at changing the status of one's group, we expect that normative regard (and specifically implicit

normative regard) will be a stronger predictor of collective action than attitudes (both implicit and explicit). Although attitudes (i.e., personal evaluations of one's group) may have some relation to collective action, collective action as a behavior should be most closely tied to normative regard because it directly assesses how the group is viewed.

We chose feminists as the devalued group to study because feminists tend to be viewed negatively by others (Haddock & Zanna, 1994) and thus they should be motivated to protect their social identity through collective action. We compared feminists to women who did not identify with feminism because these women should not be motivated to protect their social identity (as women in general tend to be rated quite positively, Eagly & Mladinic, 1989). Thus, we assessed women's identification with feminism and over-selected women who highly identified with feminism to test whether feminists engage in collective action to improve the status of their group but non-feminists do not. We expected that positive implicit normative regard would predict less volunteering behavior among feminists when asked to volunteer for a relevant cause (i.e., a cause that benefits feminists) but not for an unrelated cause (i.e., a cause that benefits students).

We assessed attitudes and normative regard (both implicitly and explicitly) online, and then randomly assigned participants to volunteer for a cause, framed as advancing the interests of feminists or students. We then assessed their intention to volunteer for the cause. We predicted that only women who identified with feminism, who had relatively negative implicit normative regard about feminists, and who were asked to volunteer for a feminist cause would show greater collective action.

### Method

#### Participants

Forty-nine female participants (age range from 18 to 57, $M = 20.04$) from the introductory psychology participant pool participated in this study in exchange for course credit.

#### Materials

*Identification with feminism.* At the beginning of the semester, participants completed a measure of identification with feminism (Reid & Percell, 2004). These questions were "I am a feminist," "I am strongly identified with being a feminist," and "being a feminist is central to who I am." We also asked participants how favorable their attitude was towards feminists, and combined this item with the other three (Cronbach's $\alpha = .89$).

*Activism Orientation Scale (Corning & Myers, 2002).* At the beginning of the semester, participants responded to 28 questions ranging from 0 (*extremely unlikely*) to 3 (*extremely likely*) assessing the likelihood that they would engage in various politically-oriented activist behaviors in the future. An example item is "display a poster or bumper sticker with a political message?" (Cronbach's $\alpha = .96$).

*Implicit attitudes and normative regard.* Participants completed two IATs before coming into the lab; a measure of implicit attitudes (with the category labels "I like/I don't like"), and a measure of implicit normative regard (with the category labels most "people like/most people don't like"). The second category was "feminist/housewife." Example items are "activist" and "liberated" for the feminist category, and "domestic" and "traditional" for the housewife category.

*Explicit normative regard.* We assessed explicit normative regard using a modified evaluative thermometer. Participants were asked to ignore their own attitudes towards feminists and filler groups (union members, gay men and lesbians, African Canadians, and English

---

[11] It is possible that Muslim participants affirmed a social identity other than their Muslim identity and that this may explain the results in this study. By affirming another identity, their Muslim identity may have been made less important to them, allowing them to acknowledge that there are negative norms towards Muslims. Consistent with this possibility, in this study only 4 of the 17 participants in the affirmation condition chose to affirm their Muslim identity. Moreover, although we could not test the significance of this difference due to the low N, Muslims who affirmed their Muslim identity did have more positive implicit normative regard than Muslims who did not affirm their Muslim identity ($M$(affirmed Muslim ID) = 56.76, $SD = 142.76$; $M$(affirmed other ID) = −2.69, $SD = 195.17$). This finding does not explain the results of this study because the Muslims in the affirmation group *explicitly* thought that norms towards their group were *less* negative than those who had not been affirmed. This differential pattern of results for implicit and explicit normative regard helps to rule out the possibility that these results are solely due to affirming a non-devalued identity.

Canadians) and instead asked to report how most undergraduates at their university would rate the groups on a scale from 0 (*extremely unfavorable*) to 100 (*extremely favorable*).

*Intention to volunteer.* We measured participants' intention to engage in several volunteer behaviors (sign a petition, wear a button, join the group, go to a rally on campus, distribute flyers on campus, organize the flyer distribution and/or rally, protest outside the National Parliament building, and stand for election to the executive of the group), each one requiring more effort. A separate group of participants rated how much effort each behavior required (Bass, Cascio, & O'Connor, 1974). Each behavior chosen was then multiplied by these effort ratings and summed.

### Procedure

Participants completed the identification with feminism and the activism orientation scales at the beginning of the term in a mass-testing session. They then completed the two IATs (implicit attitudes and implicit normative regard) in two counter-balanced sessions. The measure of explicit normative regard came after the measure of implicit normative regard. The online sessions were spaced 4 days to 1 week apart.

Participants arrived to the lab in groups of three to six, ostensibly to engage in group work. They were randomly assigned to either the feminist condition or the control condition. In the feminist condition participants were met by a female experimenter wearing a button that had "Feminists for Child Care" and a pink female symbol on it. In the lab these same buttons were present, as were pamphlets with the symbol on them. In the control condition, the button had "Waterloo Students for Child Care" and cartoon children on it. Participants were told that they would be completing a short experiment and, because the study was short, the experimenter was allowed to recruit them to participate in various activities for a group to which she belonged on campus. Participants were asked to indicate an interest in participating in the group only if they were committed to doing so, because volunteering and then backing down would be problematic for the researcher, allowing us to assess committed behavioral intentions.

In order to increase participants' collective identity and sense of interdependence, participants then completed the supposed experiment by reading a paragraph and circling the number of times that the word "we" was used (Brewer & Gardner, 1996). Then, participants completed the dependent variable of interest, in which participants were given the choice to indicate whether they intended to volunteer for Feminists for Child Care in the feminist condition or Waterloo students for Child Care in the control condition.

After completing this questionnaire, participants were given a funnel debriefing in which we probed for suspicion and fully explained the nature of the experiment.

### Results and discussion

We replicated our previous group differences. Identification with feminism was marginally correlated with positive implicit attitudes about feminists, $r(64) = .24$, $p = .06$, and with implicit normative regard about feminists, $r(65) = .23$, $p = .06$.

We regressed intentions to volunteer on implicit normative regard, condition, and participants' identification with feminism, their two-way interactions, and the three-way interaction.[12] There were no significant main effects. There was, however, a significant interaction between identification with feminism and condition, $\beta = .60$, $t(41) = 2.48$, $p = .02$, such that identification with feminism predicted intentions to volunteer in the feminist condition, but not in the control condition. This two-way interaction was qualified by the

predicted three-way interaction, $\beta = .59$, $t(41) = 2.52$, $p = .02$, as illustrated in Fig. 3. In the control condition, neither implicit normative regard, $t(41) = 1.00$, *ns*, feminism, nor their interaction predicted intentions to volunteer, *ts* < 1. In the feminist condition, identification with feminism predicted intentions to volunteer, $\beta = .57$, $t(41) = 2.77$, $p = .008$, as did the interaction between identification with feminism and implicit normative regard, $\beta = .62$, $t(41) = 3.40$, $p < .001$, such that more positive implicit normative regard predicted lower intentions to volunteer for women who identified with feminism, $\beta = .57$, $t(41) = 5.67$, $p < .0001$, but did not predict for women who did not identify with feminism, $\beta = -.15$, $t < 1$. This predicted effect was not explained by participants' implicit attitudes, explicit normative regard, or their general tendency to engage in activism (as assessed by the activism orientation scale) because when we controlled for these variables this three-way interaction remained significant, $\beta = .59$, $t(39) = 2.46$, $p = .02$. Thus, women who identified with feminism and had positive implicit normative regard about feminists showed less intention to volunteer for a feminist cause. Ironically then, those who protect their social identity by forming positive implicit normative regard may be less likely to attempt to improve the status of their group.

Recall that we predicted that implicit normative regard would predict collective action, but that in this study implicit attitudes and explicit normative regard would not. Was this indeed the case? Our results supported these predictions. Implicit attitudes, explicit normative regard, and the activism orientation scale, were not related to intentions to volunteer in either condition, and did not interact with condition or participants' level of feminism to predict intentions to volunteer, all *ts* < 1.20, *ns*.

In addition, in this study, as in Study 1, implicit attitudes and implicit normative regard were correlated, $r(47) = .32$, $p = .03$, suggesting that although implicit attitudes and implicit normative regard are related, they are not redundant. As we predicted, feminists' willingness to engage in collective action was related to their implicit normative regard towards their group but not their implicit attitude towards feminists, supporting our assertion that these two measures are distinct. Thus, we have found evidence that, in some situations at least, implicit normative regard can predict a behavioral intention that implicit attitudes cannot.

We also found, as we predicted, that implicit but not explicit normative regard predicted collective action. Past research has found that explicit normative regard predicts collective action (Louis & Taylor, 1999), but this was not the case in our study. It is also possible
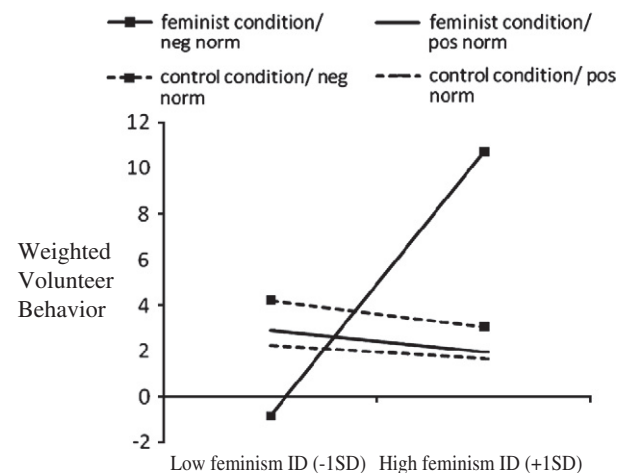


**Fig. 3.** Three-way interaction between condition, feminism, and implicit normative regard, predicting weighted number of volunteer behaviors. "Feminist condition/neg norm" refers to participants randomly assigned to the feminist condition with negative implicit normative regard towards feminists. Normative regard is plotted at one standard deviation above and below the mean.

---

[12] Three participants were excluded from these analyses because their implicit normative regard scores were more than three standard deviations from the mean.

that explicit normative regard did not predict behavior because it was assessed using only one item. A more thorough assessment of whether feminists believed most people liked feminists might have been related to their propensity to engage in collective action to benefit their group, but as we have argued above we would not expect such a finding.

As we predicted, only women who identified with feminism engaged in collective action in this study, supporting our assertion that despite overall favorable ratings of women, some groups of women (such as feminists, Haddock & Zanna, 1994), or women in some domains (such as traditionally masculine domains like engineering) are evaluated negatively and discriminated against, and it is these groups of women who may be most likely to attempt to protect their social identity. Our results support this conclusion, but future research will be needed to examine whether some sub-groups of devalued groups are more likely to attempt to protect their social identity than others. More generally we theorize that discrimination will only make identity protection at an explicit level more difficult for the sub-group that is discriminated against. For example, discrimination against feminists may have little effect on women's ability to believe that women in general are liked, but would make it quite difficult for feminists to believe that feminists are well liked.

## General discussion

Throughout this paper, we have provided evidence that members of devalued groups (e.g., African-Canadians, feminists, Asian-Canadians, and Muslims) have more positive implicit normative regard about their group than do members of the majority. We have argued that this positive normative regard arises from a motivation to protect one's social identity and found support for our argument in Study 2, which demonstrated that when this motive is met through other means (in this case, an affirmation), this group difference in implicit normative regard was no longer evident. Study 3 demonstrated that this new measure of implicit normative regard can also predict behavioral intentions (i.e., collective action), over and above participants' explicit beliefs of what the majority thinks of their group, their implicit attitudes towards their group, and even their general tendencies towards activism, all of which did not predict behavioral intentions in this situation.

We have argued that implicit and explicit measures of normative regard differentially reflect the experiences of devalued groups. Supporting this assertion, we found that African-Canadians and Muslims both say that most people do not like their group on explicit measures, but on implicit measures, members of both groups believed that most people actually do like their group (as compared to members of non-devalued groups). Further supporting the notion that implicit and explicit measures of normative regard assess different constructs, we found that, when affirmed, implicit and explicit normative regard demonstrated opposite patterns of results. When affirmed, Muslims implicitly believed that most people like their group less, but they explicitly believed that most people like their group more, than when Muslims were not affirmed. We argue that explicit measures are affected by the reality of inequality that members of devalued groups face whereas implicit measures are not, and these differing patterns of results support this claim.

Our findings not only suggest that explicit normative regard is distinct from implicit normative regard, but that implicit attitudes are distinct from implicit normative regard. First, we found that members of devalued groups demonstrated the same pattern of results on both implicit and explicit measures of attitudes, whereas they had different patterns on implicit and explicit normative regard. Second, we found that implicit normative regard predicted behavioral intentions over and above measures of implicit attitudes, which in this research did not predict intentions to engage in collective action.

In other research (Yoshida et al., 2010) we have found other results that further discriminate between these two measures. First, we have found that implicit personalized attitudes (Olson & Fazio, 2004) and implicit normative regard independently predict scores on a more traditional IAT. Second, we have found that the amount of time Asians have spent in North America influenced their implicit normative regard but not their implicit attitudes towards the elderly. Third, implicit normative regard was able to predict whether female engineers intended to drop out of engineering, over and above their implicit attitudes. Fourth, hearing an audience laugh at a racist joke about people from the Middle East (vs. hearing the same joke with no laughter) made implicit normative regard towards people from the Middle East more negative but did not influence implicit attitudes, and this implicit normative regard (but not implicit attitudes) predicted cutting funds for a Muslim organization on campus. Fifth, we have found that implicit normative regard predicted response time on the shooter bias task (Correll, Park, Judd, & Wittenbrink, 2002) whereas implicit attitudes, explicit attitudes, and explicit normative regard did not. Together with findings from this paper, these differential patterns of results suggest that these measures are not the same, and that implicit normative regard can add to our understanding of the experiences of members of devalued groups.

In this paper we have suggested that members of devalued groups develop positive implicit normative regard in order to protect their social identity. If members of devalued groups do not strongly identify with a devalued group to which they belong, however, we suggest that they will not protect that devalued group membership. Consistent with this logic, in Study 3 we found that female participants who did not identify with feminism did not have positive implicit normative regard towards feminists. We also suggest that individuals will only be motivated to protect a social identity when they are strongly identified with that social group and may not protect a social identity that is new. Also consistent with this logic, in the paper cited above we found that female engineering students' developed more negative implicit normative regard towards female engineers over their first year of university, presumably because this identity was new and not fully formed. Future research will need to examine whether protection of a devalued social identity changes over time.

We began by asking how members of groups that experience discrimination could possibly reconcile their desire to be a member of a group that is liked with evidence that their group is in reality not liked. Throughout this paper, we have suggested that, by believing that their group is liked by others at a spontaneous level, members of devalued groups can meet this need without directly confronting reality. Unfortunately, if members of devalued groups do not have to confront reality, then they do not have to attempt to change the reality that their group faces. Thus, as demonstrated in Study 3, it seems that positive effects at the individual level may have more insidious consequences at the group level.

## References

Bargh, J. A. (1990). Auto-motives: preconscious determinants of social interaction. In E. T. Higgins, & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior, vol. 2.* (pp. 93−103)New York: Guilford Press.

Bargh, J. A. (2007). *Social psychology and the unconscious: The automaticity of higher mental processes.* Frontiers of social psychology. New York: Psychology Press.

Bass, B. M., Cascio, W. F., & O'Connor, E. J. (1974). Magnitude estimations of expressions of frequency and amount. *The Journal of Applied Psychology, 59*, 313−320.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42*, 192−212.

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among Africa-Americans: implications for group identification and well-being. *Journal of Personality and Social Psychology, 77*, 135−149.

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71*, 83−93.

Chartrand, T. L., & Bargh, J. A. (2002). Nonconscious motivations: Their activation, operation, and consequences. In A. Tesser, D. A. Stapel, & J. V. Wood (Eds.), *Self and*

motivation: Emerging psychological perspectives (pp. 13—41). Washington: American Psychological Association.

Corning, A. F., & Myers, D. J. (2002). Individual orientation toward engagement in social action. Political Psychology, 23, 703—729.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: using ethnicity to disambiguate potentially threatening individuals. Journal of Personality and Social Psychology, 83, 1314—1329.

Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. Personality and Social Psychology Bulletin, 20, 503—513.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: the self-protective properties of stigma. Psychological Review, 96, 608—630.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. Personality and Social Psychology Bulletin, 15, 543—558.

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., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: an investigation of the 1984 presidential election. Journal of Personality and Social Psychology, 51, 505—514.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: affirming the self through derogating others. Journal of Personality and Social Psychology, 73, 31—44.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: dissonance changes explicit, but not implicit attitudes. Journal of Experimental Social Psychology, 40, 535—542.

Gawronski, B., Strack, F., & Bodenhausen, G. V. (2009). Attitudes and cognitive consistency: the role of associative and propositional processes. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), Attitudes: Insights from the new implicit measures (pp. 85—117). New York: Psychology Press.

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., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: an improved scoring algorithm. Journal of Personality and Social Psychology, 85, 197—216.

Gunz, A. J. (2004). Compunction's beta: The power of the motivation to avoid prejudice. Unpublished doctoral dissertation. University of Waterloo

Haddock, G., & Zanna, M. P. (1994). Preferring "housewives" to "feminists": categorization and the favorability of attitudes towards women. Psychology of Women Quarterly, 18, 25—52.

Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2005). Types of high self-esteem and prejudice: how implicit self-esteem relates to ethnic discrimination among high

explicit self-esteem individuals. Personality and Social Psychology Bulletin, 31, 693—702.

Louis, W. R., & Taylor, D. M. (1999). From passive acceptance to social disruption: towards and understanding of behavioral responses to discrimination. Canadian Journal of Behavioral Science, 31, 19—28.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: self-evaluation of one's social identity. Personality and Social Psychology Bulletin, 18, 302—318.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). Origins of achievement motivation. In D. C. McClelland, J. W. Atkinson, R. A. Clark, & E. L. Lowell (Eds.), The Achievement Motive: Century Psychology Series (pp. 275—318). East Norwalk, CT: Appleton-Century-Crofts.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of evaluative associations on the implicit association test: personalizing the IAT. Journal of Personality and Social Psychology, 86, 653—667.

Reid, A., & Percell, N. (2004). Pathways to feminist identification. Sex Roles, 50, 759—769.

Sherman, D. K., Kinias, Z., Major, B., Kim, H. S., & Prenovost, M. (2007). The group as a resource: reducing biased attributions for group success and failure via group affirmation. Personality and Social Psychology Bulletin, 33, 1100—1112.

Son Hing, L. S., Bobocel, D. R., & Zanna, M. P. (2002). Meritocracy and opposition to affirmative action: making concessions in the face of discrimination. Journal of Personality and Social Psychology, 83, 493—509.

Spencer, S. J., Peach, J. M., Yoshida, E., & Zanna, M. P. (2010). Learning what most people like: how implicit attitudes and normative evaluations shape prejudice and stereotype threat, and are shaped by social identity protection and culture. In J. Forgas, J. Cooper, & W. Crano (Eds.), The Psychology of Attitude and Attitude Change: The 12th Sydney Symposium of Social Psychology (pp. 155—169)..

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), Advances in experimental social psychology, Vol. 21. (pp. 261—302)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. 187—204). Monterry, CA, USA: Brooks/Cole.

Taylor, D. M. (1997). The quest for collective identity: the plight of disadvantaged ethnic minorities. Canadian Psychology, 38, 174—190.

Thrash, T. M., Elliot, A. J., & Schultheiss, O. C. (2007). Methodological and dispositional predictors of congruence between implicit and explicit need for achievement. Personality and Social Psychology Bulletin, 33, 961—974.

Yoshida, E., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2010). Not all automatic associations are created equal: How implicit normative evaluations are distinct from implicit attitudes and uniquely predict meaningful behavior. Unpublished manuscript, University of Waterloo, Waterloo, Ontario, Canada.