

## Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination

Denise Sekaquaptewa,<sup>a,\*</sup> Penelope Espinoza,<sup>a</sup> Mischa Thompson,<sup>a</sup>  
Patrick Vargas,<sup>b</sup> and William von Hippel<sup>c</sup>

<sup>a</sup> Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1109, USA

<sup>b</sup> Department of Advertising, University of Illinois, IL, USA

<sup>c</sup> School of Psychology, University of New South Wales, Sydney, Australia

Received 7 March 2001; revised 11 February 2002

### Abstract

Two experiments examined whether a measure of implicit stereotyping based on the tendency to explain Black stereotype-incongruent events more often than Black stereotype-congruent events (Stereotypic Explanatory Bias or SEB) is predictive of behavior toward a partner in an interracial interaction. In Experiment 1 SEB predicted White males' choice to ask stereotypic questions of a Black female (but not a White male or White female) in an interview. In Experiment 2 the type of explanation (internal or external attribution) made for stereotype-inconsistency was examined. Results showed that White participants who made internal attributions for Black stereotype-incongruent behavior were rated more positively and those who made external attributions were rated more negatively by a Black male confederate. These results point to the potential of implicit stereotyping as an important predictor of behavior in an interracial interaction.

© 2002 Elsevier Science (USA). All rights reserved.

**Keywords:** Stereotyping; Intergroup behavior; Prejudice; Interracial interaction

Stereotyping and prejudice are difficult to measure because people are often unwilling to admit negative attitudes and beliefs about social groups (Fazio, Jackson, Dunton, & Williams, 1995). Additionally, people may sometimes be unable to accurately report on these topics because how they think and feel about social groups may not be consciously accessible to them (Greenwald & Banaji, 1995). Thus, researchers are faced with a substantial "willing and able" problem when attempting to measure prejudice and stereotyping.

In response to this "willing and able" problem, researchers turned to measures of implicit prejudice and stereotyping. Such measures are thought to tap consciously inaccessible group-based attitudes and beliefs (Greenwald & Banaji, 1995). Most research attention has focused on implicit prejudice measures, which are intended to assess the degree of positivity or negativity an individual implicitly associates with social groups

(e.g., Greenwald, McGhee, & Schwartz, 1998). Somewhat less research attention has focused on implicit stereotype measures (e.g., Wittenbrink, Judd, & Park, 1997), and implicit stereotyping, which we define as the unintended influence of stereotypes on information processing (cf. Brewer, 1996). In part, this focus on prejudice rather than stereotypes/stereotyping probably emerged because prejudice has traditionally been thought to be more consequential than stereotyping for behavioral outcomes such as discrimination (Brigham, 1971; Stangor, Sullivan, & Ford, 1991).

To the extent that measures of implicit prejudice and stereotyping assess important processes relevant to intergroup attitudes and perceptions (von Hippel, Sekaquaptewa, & Vargas, 1995, 1997), it seems reasonable to expect them to relate to intergroup behavior. Yet such demonstrations are rare. In one study, White participants who implicitly favored Whites over African-Americans were rated by observers as having more positive interactions with a White than a Black experimenter (McConnell & Leibold, 2001; see also

\* Corresponding author. Fax: 1-734-647-9440.

E-mail address: [dsekaqua@umich.edu](mailto:dsekaqua@umich.edu) (D. Sekaquaptewa).

Fazio et al., 1995). Similarly, Whites high in implicit prejudice showed greater indications of anxiety (e.g., eyeblinking) when interacting with a Black partner (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). The dearth of studies on the predictive utility of such measures (particularly on implicit stereotyping, concerning which no studies could be located) indicates the need for more experimental investigations in this area.

If stereotyping is defined as the use of stereotypes in information processing (Brewer, 1996), then there are almost as many ways to operationalize implicit stereotyping as there are known stereotypic biases. Of these biases, we focused on an explanatory bias adapted from Hastie (1984) and introduced in our earlier work (von Hippel et al., 1997, Experiment 3). This explanatory bias emerges when one is more likely to provide explanations for behaviors that are inconsistent with expectancies than for behaviors that are consistent with expectancies. For example, if one expects an individual “James” to be unintelligent, learning that “James received an A on the test” may instigate explanatory processing, in an attempt to make sense of the incongruity (“...because it was an easy test”). Learning that “James received a D on the test,” on the other hand, is unlikely to instigate explanatory processing. Because such expectancies can be based on stereotypes, an explanatory bias can emerge in response to stereotype-inconsistency as well. To the extent that this stereotypic explanatory bias (SEB) reflects the unintended influence of stereotypes on processing, it is well-suited to the goal of measuring implicit stereotyping.<sup>1</sup>

If implicit stereotyping as indicated by SEB reflects differences in the way perceivers process stereotype-relevant information, then it seems likely that people who vary in SEB would react differently during interactions with stereotyped individuals. People who show SEB should tend to selectively discount counter-stereotypic behaviors from Blacks, and thereby behave in a more negative manner when interacting with a Black person. According to this logic, we predicted that respondents who show SEB would display discriminatory behavior towards a Black but not a White individual. In the lab, discriminatory behaviors are likely to be relatively subtle, involving nonverbal behaviors (Dovidio et al., 1997) or behavioral choices that are not clearly associated with discrimination. For example, Rudman and Borgida (1995) identified sexist behavior in men who chose to ask subtly sexist/stereotypic questions in an interview context. Experiment 1 was conducted using a similar procedure adapted for interracial interaction.

White male participants engaged in a mock job interview with either a White or Black interviewee (actually a research assistant). Participants were given a list of Black stereotypic and neutral questions to select for use in their interview. It was predicted that White participants who showed SEB would tend to choose stereotypic questions to ask of a Black but not a White applicant.

Because prejudicial attitudes have been shown to be better predictors of behavior than endorsement of stereotypes (e.g., Brigham, 1971), the Modern Racism Scale (MRS; McConahay, Hardee, & Batts, 1981) was also administered. We predicted that SEB would account for variance in behavioral choice beyond that explained by the MRS. Because implicit measures are frequently unrelated to more traditional, direct measures of racial attitudes (Dovidio et al., 1997; Kawakami & Dovidio, 2001; von Hippel et al., 1997), we expected that the SEB measure would not be correlated with the MRS.

## Experiment 1

### Method

#### Participants

Fifty-five White males participated in partial fulfillment of psychology course requirements.

#### Materials

Stereotypic explanatory bias was assessed by presenting participants with a series of 25 sentence beginnings, 16 of which were designed to measure responses to Black stereotype-consistent behaviors (e.g., easily made the team) and Black stereotype-inconsistent behaviors (e.g., got a job at Microsoft). Behaviors were paired with 50% male and 50% female African-American (e.g., Marcellus, Lakisha) and White names (e.g., Adam, Deborah). Nine race-neutral behaviors (e.g., Linda ate a sandwich) were also included as filler items. It was necessary to include White targets in the measure to ensure that participants were responding to the combination of the target's race and the race stereotypicality of the behavior, as opposed to only the behavior itself. Additionally, the SEB items included both positive and negative Black stereotypic behaviors (i.e., easily made the team; blasted loud music in his car) and positive and negative counter-stereotypic behaviors (i.e., enrolled at Princeton; refused to dance). Participants were asked to add words to the end of the sentence stem in any manner that created a grammatically correct sentence (see Hastie, 1984). SEB is evidenced by providing more explanations for Black targets engaging in Black stereotype-inconsistent than consistent behaviors.

<sup>1</sup> In a pretest SEB was demonstrated in students enrolled in a course on stereotyping and was uncorrelated with explicit stereotype endorsement,  $r(59) = -.005$ ,  $p = .97$ , supporting the idea that the bias is implicit and distinct from conscious stereotype endorsement.

The Modern Racism Scale (a seven-item self-report measure) was also administered.

### *Procedure*

Under the guise of a study on interview skills, participants were asked to engage in a mock job interview with a research assistant (RA), said to be in training to make a video on interviewing skills. It was explained that it was the participants' role to help with this training by providing the opportunity for the RA to interact with someone unfamiliar to him or her. It was also explained that the interview skills session would not take long, so that after it was over the participant would take part in an unrelated survey on social attitudes to complete the hour. Participants were randomly assigned to interact with one of three RAs.

The RAs were a White male, a White female, and a Black female. Both a White and Black female RA were used to assess whether the SEB measure predicts behaviors specifically toward members of the targeted stereotyped group (African-Americans) as opposed to other stereotyped outgroup members (women).

After being introduced, the participant was separated from the RA. The participant was told he would serve in the role of interviewer and the RA would be the interviewee, and that the job the RA was interviewing for was a restaurant supervisor position. Participants were given a series of interview questions from which they were to choose one option from each of 14 pairs. Interleaved among eight filler question pairs were six question pairs designed so that the participant would be forced to choose between mildly stereotypic or nonstereotypic question wording (see Rudman & Borgida, 1995).<sup>2</sup> For example, participants were asked to choose between the questions "...some people think they can get away with stealing food, silverware, even cash. Have you ever had any trouble like this?" (stereotypic), and "... some people think they can get away with taking work supplies home. Have you ever experienced it, and what did you do about it?" (nonstereotypic). Participants were allowed to look over the question pairs, and then the interview began.

As the participant chose a question and asked it of the RA, he made a checkmark on a sheet indicating which question in each pair he selected. The RA responded by giving answers that were scripted to be pleasant but neutral in tone, sufficient to answer either the stereotypic or nonstereotypic questions, and were identical across the race and gender of the RA. After this interview, the RA left the room while the participant

rated the RA on interview skills (skilled interviewer, capable, and confident).

Next, the experimenter explained that the RA would now serve as the interviewer and the participant as the interviewee, in order for the RA to practice asking as well as answering questions. It was explained to the participant that he should "role play" and answer the questions as if it were an actual job interview, and that his responses were not important to the exercise. The RA then asked the subject a series of open-ended questions related to a restaurant job. The RA then left the room again while the participant completed a second questionnaire asking about the RA's skills at asking questions. During this time, the RA completed a questionnaire rating his/her overall impression of the participant (liking of the participant, friendliness of the participant), in order to test whether the RA's impression of the participant would be predicted by SEB (similar to Fazio et al., 1995).

After the participant finished rating the RA, the experimenter explained that the interview skills study was over, and the experimenter and the RA thanked the participant. The experimenter then escorted the participant to another room to take part in an "intercollegiate survey on attitudes about social groups." Within this survey was the SEB measure followed by the MRS. Finally, participants were probed for suspicion and fully debriefed.

### *Results*

#### *Scoring SEB*

The sentence completions that participants provided on the SEB measure were coded by two independent judges as to whether they explained the behavior in the sentence beginning or simply continued the sentence without explaining. Because the judgments of the two coders were highly correlated ( $r = .91$ ,  $p < .001$ ), these ratings were collapsed across judges.

An SEB score was derived by subtracting the number of explanations provided for Black stereotype-consistent events (Black actors engaging in stereotypically Black behaviors) from the number of explanations provided for Black stereotype-inconsistent events (Black actors engaging in counter-stereotypic or stereotypically White behaviors; i.e., BW–BB, where the first letter indicates the race of the target and the second letter indicates whether the behavior is Black or White stereotypic). Higher positive scores indicated greater Black SEB (B-SEB). B-SEB scores ranged from  $-2.50$  to  $2.50$ ,  $M = .77$ ,  $SD = 1.06$ , and differed significantly from zero,  $t(54) = 5.40$ ,  $p < .001$ .

A second SEB score was calculated in order to test whether Whites' behaviors toward a Black interaction partner is predicted by the processing of Black as opposed to White targets engaged in stereotype-incongruity. This second SEB score was derived by subtracting

<sup>2</sup> An independent sample of 20 White undergraduates rated each question on stereotypicality, on a scale from 1 (not at all stereotypic) to 10 (very stereotypic). Stereotypic questions ( $M = 4.44$ ) were rated significantly higher than neutral questions ( $M = 2.15$ ),  $t(19) = 6.24$ ,  $p < .001$ .

the number of explanations provided for White stereotype-consistent events (Whites engaging in stereotypically White behaviors) from the number of explanations provided for White stereotype-inconsistent events (Whites engaging in counter-stereotypic behaviors; i.e., WB–WW). Higher positive scores indicated greater White SEB (W-SEB). W-SEB scores ranged from  $-2.50$  to  $1.50$ ,  $M = -.67$ ,  $SD = .98$ , and differed significantly from zero,  $t(54) = -5.08$ ,  $p < .001$ .

#### Scoring MRS

Response choices on the MRS ranged from 1 (strongly disagree) to 7 (strongly agree), and were summed such that higher scores indicated higher modern racism. Scores on the MRS ranged from  $7.00$  to  $36.00$ ,  $M = 18.84$ ,  $SD = 5.92$ .

#### Question choice

The number of stereotypic questions asked ranged from 0 to 6. In order to test the primary hypothesis that B-SEB would predict question choice with a Black but not a White interviewee, question choice was analyzed in a simultaneous multiple regression, using B-SEB, MRS and type of RA (White male, White female, and Black female) and interactions as predictor variables. Results showed that the B-SEB by RA type interaction term accounted for significant variance in question choice,  $\beta = .34$ ,  $p < .03$ , whereas the MRS by RA type interaction term did not,  $\beta = .01$ ,  $p = .96$ . Within cell regression analyses showed that B-SEB significantly predicted question choice for the Black female RA, but not for the White male or White female RA conditions (see Table 1).

In order to compare B-SEB to W-SEB, an additional multiple regression analysis was conducted including W-SEB in the model. Results showed that the B-SEB by RA type interaction term remained a significant predictor of question choice,  $\beta = .32$ ,  $p < .05$ . The W-SEB by RA type interaction term was not significant,  $\beta = .17$ ,  $p = .24$ .<sup>3</sup>

No significant effects of B-SEB, W-SEB, or MRS emerged in participants' ratings of the RA or the RA's impression of participants and none of these variables were reliably correlated with each other.

#### Discussion

Consistent with predictions, Experiment 1 revealed that the tendency to engage in SEB (i.e., to provide more explanations for Black stereotype-inconsistency than Black stereotype-consistency) predicted the number of stereotypic questions White male participants chose to ask while interviewing a Black female RA. SEB was not

Table 1

Question choice predicted by B-SEB and MRS for Black female, White male, and White female interviewees (Experiment 1)

	$\beta$	$p$	$r$	$p$
(a) Black female RA				
SEB	.521	.05	.53	.02
MRS	.216	.42	.21	.23
(b) White male RA				
SEB	-.257	.36	-.20	.22
MRS	.167	.54	.12	.32
(c) White female RA				
SEB	-.309	.20	-.38	.04
MRS	-.094	.69	-.22	.16

predictive of this question choice for interviews with either a White male or White female RA. Scores on the Modern Racism Scale did not predict this question choice with any of the RAs, nor did it predict the competence ratings of the RA by the participant, and the RA's final impression of the participant. SEB also did not predict these evaluations.

In light of previous research showing that White participants high in implicit prejudice were disliked by a Black female experimenter (Fazio et al., 1995; see also McConnell & Leibold, 2001), it was somewhat surprising that SEB failed to predict the RA's evaluations of the participant. It is possible, however, that the interaction in the current experiment may have been sufficiently constrained by forcing participants to play the relatively scripted roles of interviewer and interviewee that there was insufficient latitude in behavior for the predicted effect to emerge. Impression ratings might be predicted by SEB in a more spontaneous, unscripted social interaction.

#### Experiment 2

Although Experiment 1 provided initial evidence that SEB can predict behavioral outcomes, there are several questions left unanswered. Because SEB did not predict impressions of the participants, perhaps because the interaction was too scripted, the first goal of Experiment 2 was to provide a less constrained interaction setting. Second, in Experiment 1 SEB was found to predict behavior beyond that predicted by the MRS, but the MRS was not itself an effective predictor. It may be that SEB predicted unique variance in behavior because it was paired with an explicit rather than an implicit prejudice measure. Therefore, Experiment 2 included an implicit prejudice measure (the Implicit Associations Test, or IAT; Greenwald et al., 1998).

Finally, and perhaps most importantly, in Experiment 1 responses on the SEB measure were coded simply as whether they provided an explanation for the

<sup>3</sup> An analysis without B-SEB also showed that W-SEB was not a significant predictor.

behaviors. Not all explanations are equally potent dismissals of the implications of a behavior, however, as some types of explanations are likely to be more stereotype-maintaining than others. In particular, attributions to the situation rather than the actor may be particularly stereotype-maintaining. For example, responding to, “Shaniqua scored high on the SAT...” by adding “because she took preparation courses” maintains the stereotype that Blacks are unintelligent more so than the explanation “because she is smart.” When SEB involves primarily internal attributions for stereotype-inconsistency, SEB may actually promote stereotype change or reduction, whereas when SEB involves primarily external attributions, it is likely to lead to stereotype maintenance. If SEB predicts discriminatory behavior due to its stereotype-maintaining properties, a form of SEB that instead promotes stereotype reduction may actually predict positive interracial interactions. Therefore, in Experiment 2, the coding of the SEB measure included categorization of internal and external explanations, with the prediction that external SEB may be more related to negative behavioral outcomes than internal SEB.

### *Method*

#### *Participants*

Seventy-nine White participants (27 male and 52 female) participated in partial fulfillment of psychology course requirements.

#### *Procedure*

Upon reporting to the lab, the participant was shown to a cubicle and led to believe other students occupied the remaining cubicles. It was explained that participants were to be paired with another participant for an experiment on game-playing, but that the task would not take long, so they would also take part in an unrelated survey on social attitudes to complete the hour. Participants were randomly assigned to interact with either a Black male or a White male confederate.

The participant and confederate were escorted to a larger room for the game-playing task. The participant and confederate played three rounds of a paper-and-pencil game similar to an extended version of tic-tac-toe. Each round was terminated after 3 min regardless of whether one person had won. The confederate was instructed to not try to win, nor initiate conversation, but to respond cordially if the participant initiated conversation. The experimenter recorded the outcome of each round of the game. Between the second and the third round, the experimenter left the room for 2 min, presumably to attend to other students. This provided an opportunity for the participant to have an unstructured interaction and possibly initiate conversation with the confederate.

After the final round of the game, the participant returned to the cubicle to complete a filler questionnaire regarding strategies used during the game. During this time, the confederate completed the primary dependent measures. The confederate rated his impression of the participant, using the following items: I liked the participant; the participant was nice; was friendly; was selfish; was cold towards me; didn't want to play the game with me (the final three were reverse scored). The confederate also rated the participant on the positive behaviors he saw the participant display during the interaction, using the following items: the participant looked me in the eye; spoke to me before or during the game; maintained a closed posture by crossing his/her arms (reverse scored).

After completing the filler questionnaire, participants were instructed to begin the second study as in Experiment 1. The survey packet instructed the participant to complete the SEB measure, then use the computer located in the cubicle for the IAT (the race version of the IAT was given as in Greenwald et al., 1998), and when finished to complete the rest of the survey packet, which contained the MRS.<sup>4</sup> Participants were then probed for suspicion, debriefed, and thanked.

### *Results*

The IAT computer program recorded reaction times in a categorization task wherein participants responded to Black and White names and pleasant and unpleasant words. Outliers in the IAT data were trimmed, and the data log transformed prior to analysis, according to the procedure described by Greenwald et al. (1998). Averaged reaction times to unpleasant words paired on the same response key as Black names, and pleasant words paired on the same response key as White names, were subtracted from averaged reaction times to unpleasant words paired on the same response key as White names, and pleasant words paired on the same response key as Black names. Higher positive difference scores indicated more negative associations to Blacks and/or more positive associations to Whites (see Greenwald et al., 1998). IAT difference scores ranged from  $-174.00$  to  $879.00$ ,  $M = 315.14$ ,  $SD = 184.37$ , and differed significantly from zero,  $t(78) = 15.29$ ,  $p < .001$ .

The SEB measure was identical to that used in Experiment 1 and responses were again scored as to whether they explained the behavior in the sentence stem. (As in Experiment 1, W-SEB was computed and found

<sup>4</sup> Because responding on the SEB measure may be altered when participants report racial attitudes first, whereas responding on the IAT seems largely uncontrollable (Kim & Greenwald, 1998), the SEB was given first, followed by the IAT, and the MRS was given last, as it is apparent to participants that it measures racial prejudice (Fazio et al., 1995).

not to relate to any dependent measures, so results are reported for B-SEB only.) In addition, responses judged to be explanations were further scored as internal (i.e., dispositional) or external (i.e., situational) explanations. As such, two separate SEB scores were derived from the measure. An external SEB score (E-SEB) was derived by subtracting the number of external explanations provided for Black stereotype-consistent events from the number of external explanations provided for Black stereotype-inconsistent events. An internal SEB score (I-SEB) was derived by subtracting the number of internal explanations provided for Black stereotype-consistent events from the number of internal explanations provided for Black stereotype-inconsistent events.

E-SEB scores ranged from  $-2.00$  to  $1.00$ ,  $M = -.11$ ,  $SD = .49$ , and differed significantly from zero,  $t(78) = -2.05$ ,  $p < .05$ . I-SEB scores ranged from  $-3.00$  to  $1.50$ ,  $M = -.25$ ,  $SD = .79$ , and differed significantly from zero,  $t(78) = -2.78$ ,  $p < .01$ .<sup>5</sup> Higher scores on E-SEB and I-SEB indicate that more explanations (external or internal) were generated for Black stereotype-inconsistency than consistency. Additionally, for some analyses, a difference score was created by subtracting E-SEB from I-SEB. Positive values on this SEB difference score indicated that the participant showed more I-SEB than E-SEB.

The MRS was scored as in Experiment 1. MRS scores ranged from  $7.00$  to  $34.00$ ,  $M = 13.22$ ,  $SD = 5.00$ .

#### *Ratings of the interaction by the Black and White interaction partners*

The confederates' responses to the six items indicating his feelings toward the participant (I liked the participant, the participant was friendly, etc.) and the three items indicating his observation of the nonverbal behaviors of the participant (eye contact, initiating conversation, and closed posture) were combined to form a social interaction score (Cronbach's  $\alpha = .86$ ). Confederates' ratings on this measure did not differ by participant gender,  $F's < 1$ , so data were collapsed across gender.

The social interaction scores were analyzed in a multiple regression analysis, using confederate race, SEB difference score, IAT, MRS, and interactions as predictor variables. Results showed that the SEB difference score by confederate race interaction term was a significant predictor, suggesting that among participants who interacted with a Black confederate, those who showed more E-SEB received lower social interaction scores, whereas those who showed more I-SEB received higher social interaction scores. The IAT by confederate

race and MRS by confederate race interaction terms were not significant (see Table 2).

In order to further examine how I-SEB and E-SEB predicted the nature of the interaction differentially for Black and White confederates, two separate regression analyses were conducted including the I-SEB and E-SEB scores separately. These analyses showed that the E-SEB by confederate race interaction term was significant,  $\beta = -.23$ ,  $p < .05$ , and the I-SEB by confederate race interaction was also significant, but in the opposite direction,  $\beta = .28$ ,  $p < .05$ . Within cell regression analyses showed that in the Black confederate condition, E-SEB predicted more negative social interaction scores,  $\beta = -.67$ ,  $p < .001$ , and I-SEB predicted more positive social interaction scores,  $\beta = .33$ ,  $p < .05$ . These scores were not significant predictors of social interaction scores in the White confederate condition,  $\beta = -.13$ ,  $p = .43$ , and  $\beta = -.24$ ,  $p = .17$ , for E-SEB and I-SEB, respectively.

#### *Correlations between SEB, IAT, and MRS*

I-SEB and E-SEB were not significantly correlated with the IAT,  $r = .01$ ,  $p = .96$ , and  $r = .15$ ,  $p = .19$ , respectively. E-SEB was significantly correlated with the MRS,  $r = .22$ ,  $p < .05$ , and I-SEB was not,  $r = .19$ ,  $p = .10$ . The MRS was not significantly correlated with the IAT,  $r = .16$ ,  $p = .15$ . I-SEB and E-SEB were significantly negatively correlated,  $r = -.23$ ,  $p < .05$ .<sup>6</sup>

#### *Discussion*

The results of Experiment 2 support the idea that stereotypic explanatory bias can predict the nature of an interracial interaction. Regression analyses showed that E-SEB and I-SEB were significant predictors of the nature of an interaction between White participants and their Black partner, in opposite directions. Black confederates rated their social interaction with White participants more negatively when the participants provided external attributions for Black stereotype-inconsistency, whereas they rated their social interaction with White participants more positively when the participants provided internal attributions for Black stereotype-inconsistency. White confederates did not differ in their social interaction ratings depending on I-SEB or E-SEB. IAT and MRS scores did not predict social interaction ratings.

The results of Experiment 2 extend the results of Experiment 1, by showing that the type of explanations generated for Black stereotype-inconsistency is a critical

<sup>5</sup> In contrast to Experiment 1, average SEB scores were negative in Experiment 2, indicating that participants showed lower SEB overall in this study. This unexpected result may be limited to this particular sample; however, the possible diminishing effect of a spontaneous interracial interaction on SEB should be addressed in future research.

<sup>6</sup> Because IAT and MRS were administered after the social interaction took place, it is possible that they failed to predict social interaction scores because prior interaction with a Black individual altered responding on these measures. While this is more likely regarding the MRS than the IAT (see<sup>4</sup>), this is still a possibility.

Table 2

Social interaction score predicted by SEB difference score, IAT, and MRS for White and Black confederate (Experiment 2)

	$\beta$	$p$	$r$	$p$
SEB difference score $\times$ confederate race	.32	.005	.32	.002
IAT $\times$ confederate race	.05	.65	.03	.41
MRS $\times$ confederate race	-.005	.96	.01	.48

*Note.* Positive values on the SEB difference scores indicate that the respondent showed more internal than external SEB. Therefore, participants who showed E-SEB had more negative and those who showed I-SEB had more positive interactions with a Black confederate. Either form of SEB was not related to social interactions with a White confederate.

factor in predicting the nature of an interracial interaction. Attributing Black counter-stereotypic behavior to external forces (e.g., “Marcellus got a job at Micro-soft... because he knew someone there”) predicts negative interactions, whereas attributing such behavior to internal traits or abilities (e.g., “... because he’s good with computers”) predicts positive interactions. These results suggest that providing external attributions for stereotype-inconsistency has greater stereotype-maintaining properties than providing internal attributions (cf. Crocker, Hannah, & Weber, 1983; Deaux, 1976).

## General discussion

These experiments explored the idea that implicit stereotyping can predict behaviors in interracial interactions. A paper-and-pencil measure of implicit stereotyping was developed, based on the tendency to provide explanations for Black stereotype-inconsistency (the stereotypic explanatory bias, or SEB). In Experiment 1, SEB predicted whether White male participants asked stereotypic questions of a Black (but not a White) interviewee in a mock job interview. In Experiment 2, SEB derived from external attributions predicted negative social interactions between White participants and Black confederates, as indicated by the ratings of the interaction by the confederate. SEB derived from internal attributions, on the other hand, predicted positive social interactions between White participants and Black confederates. The IAT and MRS did not predict these outcomes. These experiments suggest that the tendency to spontaneously explain stereotype-inconsistency can have a significant relationship with how an individual behaves during interactions with members of the stereotyped group, and that this relationship may be independent of level of racial prejudice.

An interesting feature of SEB is its ability to predict not only negative behavioral outcomes, but positive ones as well. Examining the locus of the attribution generated in response to stereotype-inconsistency (internal vs. external) allowed the directional prediction of behavior in relatively unstructured social interactions in Experiment 2. However, in Experiment 1, stereotypic question choice was predicted by an assessment of SEB that did not differentiate between internal and external

explanations. Does locus of explanation differentially predict outcomes such as stereotypic question choice?

To address this question, we conducted supplementary analyses of the SEB data from Experiment 1. Explanations were re-coded as internal or external and I-SEB and E-SEB were tested as predictors of stereotypic question choice. Results showed that neither the I-SEB nor the E-SEB by confederate type interaction terms were significant, although the direction of the betas indicated that I-SEB and E-SEB were both related to asking more stereotypic questions of a Black than a White confederate (I-SEB  $\beta = .21$ ,  $p = .24$  and E-SEB  $\beta = .25$ ,  $p = .14$ ).

Although our explanation of this is entirely post hoc, it seems possible that I-SEB and E-SEB predict in the same direction in Experiment 1 but in different directions in Experiment 2, because of the difference in the dependent variable in the two studies. I-SEB and E-SEB seem to predict behavioral outcomes differently when the outcome is based on impressions that are presumably heavily dependent on interpretations of another’s behavior (for example, liking, friendliness of partner; Experiment 2). But for outcomes that are not, or are more “cold” (e.g., choosing mildly stereotypic questions in Experiment 1), I-SEB and E-SEB may work in the same manner. This hypothesis should be further explored in future research.

In sum, the current findings provide a demonstration that measures of stereotyping can predict intergroup behavior. To our knowledge, this is the only report documenting the relationship between implicit stereotyping and behavior. Thus, these results may help establish the predictive validity of implicit stereotyping measures. In addition, these findings add to the literature on the relationship between stereotypes and discrimination, by providing evidence that measures of stereotyping can predict discriminatory behavior when the measures tap implicit stereotypic processing.

## Acknowledgments

The authors wish to thank Oscar Ybarra, James Hilton, and James Jackson for helpful comments on earlier drafts of this manuscript. Experiment 2 was

conducted as part of a master's thesis written by the second author.

## References

- Brewer, M. B. (1996). When stereotypes lead to stereotyping: The use of stereotypes in person perception. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping*. New York: Guilford Press.
- Brigham, J. C. (1971). Racial stereotypes, attitudes, and evaluations of and behavioral intentions towards Negroes and Whites. *Sociometry*, 34, 360–380.
- Crocker, J., Hannah, D. B., & Weber, R. (1983). Person memory and causal attributions. *Journal of Personality and Social Psychology*, 44, 55–66.
- Deaux, K. (1976). Sex: A perspective on the attribution process. In J. H. Harvey, W. I. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Dovidio, J., 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. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (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.
- 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.
- Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, 46, 44–56.
- Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin*, 27, 212–225.
- Kim, D., & Greenwald, A. G. (1998). Voluntary controllability of implicit cognition: Can implicit attitudes be faked? Paper presented at the 70th annual meeting of the Midwestern Psychological Association, Chicago, IL.
- McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined? It depends on who's asking and what is asked. *Journal of Conflict Resolution*, 25, 563–579.
- McConnell, A. R., & Leibold, 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.
- Rudman, L., & Borgida, E. N. (1995). The afterglow of construct accessibility: The behavioral consequence of priming men to view women as sexual objects. *Journal of Experimental Social Psychology*, 6, 493–517.
- Stangor, C., Sullivan, L. A., & Ford, T. E. (1991). Affective and cognitive determinants of prejudice. *Social Cognition*, 9, 359–380.
- von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology*, 27, 177–254.
- von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1997). The Linguistic Intergroup Bias as an implicit indicator of prejudice. *Journal of Experimental Social Psychology*, 33, 490–509.
- 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.