Journal of Experimental Social Psychology **37,** 435–442 (2001)
doi:10.1006/jesp.2000.1470, available online at http://www.idealibrary.com on IDEAL®

Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes

Allen R. McConnell

Miami University

and

Jill M. Leibold

Michigan State University

Received June 16, 2000; revised November 17, 2000; accepted November 17, 2000; published online June 8, 2001

Heretofore, no research has shown that meaningful variability on the Implicit Association Test (IAT) relates to intergroup discrimination or to explicit measures of prejudice. In the current study, White undergraduates interacted separately with White and Black experimenters, and their behavior during these social interactions was assessed by trained judges and by the experimenters themselves. The participants also completed explicit measures of racial prejudice and a race IAT. As predicted, those who revealed stronger negative attitudes toward Blacks (vs Whites) on the IAT had more negative social interactions with a Black (vs a White) experimenter and reported relatively more negative Black prejudices on explicit measures. The implications of these results for the IAT and its relations to intergroup discrimination and to explicit measures of attitudes are discussed.

Since LaPiere's (1934) classic demonstration of attitude–behavior inconsistency toward a Chinese couple traveling across the United States, social psychologists have invested a great deal of energy into developing techniques to assess group attitudes in ways that circumvent problems resulting from limited introspective access, experimenter effects, and social desirability concerns. Recently, researchers have employed various social cognition approaches to assess prejudice that minimize the problems involved with explicit reports of attitudes (e.g., Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Wittenbrink, Judd, & Park, 1997). The current work focuses on the most recent of these techniques, the Implicit Association Test (IAT), to examine the extent to which it relates to intergroup behavior and to explicit measures of racial attitudes.

The IAT has become a widely used instrument to measure attitudes in general, and prejudices toward groups in particular. It assesses attitudes by having people quickly categorize stimulus words using two response keys. In racial IAT studies, the stimulus words are names that are racially stereotyped (e.g., Jamal and Sue Ellen) or adjectives that have evaluative connotations (e.g., wonderful and disgusting). In critical trial blocks, participants categorize these words using two keys, each of which has two response options mapped to it. Typically, White participants categorize the words more quickly when "Black or undesirable" is mapped onto one key response and "White or desirable" is mapped onto the other key response than when the opposite set of key mappings (i.e., "Black or desirable" and "White or undesirable") are used (Greenwald et al., 1998). The difference in the average response latency between these two sets of key mappings is known as the IAT effect. Presumably, larger IAT effects reflect stronger associations in memory between the concept pairings (i.e., those re-

This research was supported by NIMH Grants MH58449 and MH60645. The authors gratefully acknowledge the outstanding contributions of Ayana Burroughs, Kristy Dean, and Amy Johnson, who served admirably as experimenters in this study. The diligent efforts of Ben Diaz and Zachary Eichenhorn in coding videotapes are also greatly appreciated. The helpful comments of Frank Bernieri, Chris End, and two anonymous reviewers on a previous draft of this article are greatly appreciated.

Address correspondence and reprint requests to Allen R. McConnell, Department of Psychology, Miami University, Oxford, OH 45056-1601. E-mail: mcconnar@muohio.edu.

0022-1031/01 $35.00

sponses that shared the same response key) that facilitated judgment.

Social psychologists who study group prejudice have been drawn to the IAT because of its large effect size (Greenwald et al., 1998; Nosek, Banaji, & Greenwald, 2000) and because even people who know that the IAT assesses group prejudice still reliably produce the IAT effect, indicating its robustness and apparent imperviousness. As a result, the IAT appears to circumvent many of the problems of traditional, explicit measures of prejudice. Researcher enthusiasm and large effect sizes notwithstanding, the IAT effect has not been demonstrated to be related to behavior toward group members. Although strong between-group differences have revealed favoritism for religious, ethnic, age-related, and racial ingroups (Greenwald et al., 1998; Rudman, Greenwald, Mellott, & Schwartz, 1999), meaningful variability in the strength of the IAT effect has not been shown to be related to one's discriminatory behavior.

Other implicit measures of prejudice, for example, have been shown to relate to intergroup behavior (Dovidio et al., 1997; Fazio et al., 1995). Indeed, the current research married the methodologies of this previous research to examine whether the IAT predicts intergroup discrimination. Thus, the primary goal of the current study was to explore whether this relation exists, which would help substantiate the predictive utility of the IAT.

In addition, the current work also examined whether the IAT relates to explicit measures of prejudice. It has been argued that implicit and explicit measures of attitudes tap into different knowledge and thus should be unrelated (Greenwald & Banaji, 1995; Greenwald et al., 1998), whereas others have found relations between the two (Dovidio et al., 1997, Experiment 2; Wittenbrink et al., 1997; cf., Dovidio et al., 1997, Experiment 3). To the extent that explicit measures of prejudice are reactive and subject to normative pressures, a lack of correspondence between implicit and explicit measures of prejudice is not surprising (Dunton & Fazio, 1997; Fazio et al., 1995). Although other factors, such as desire to avoid discriminatory responses because they are inconsistent with one's values (e.g., Plant & Devine, 1998), can influence behavior toward group members, minimizing self-presentation concerns should, at least, increase the likelihood of observing attitude–behavior consistency (Fazio, 1990). With respect to the IAT, Greenwald et al. (1998) did not find a correlation between the IAT and explicit measures of prejudice (i.e., feeling thermometer and semantic differential scales). The current work examined whether a relation between the IAT and explicit measures would be revealed under conditions designed to minimize self-presentation concerns.

In sum, the current study explored the relations among the IAT, intergroup behavior, and explicit reports of prejudice. Participants met with a White experimenter, and later with a Black experimenter, in structured social interactions. These interactions were videotaped and later assessed by trained judges. Also, the Black and White experimenters independently assessed their interaction during the course of the experiment. Thus, a within-subjects design allowed us to examine how each participant behaved toward a Black experimenter relative to a White experimenter. Before interacting with the Black experimenter, participants privately completed a series of questionnaires to assess their attitudes toward Blacks and Whites in a minimally reactive situation. Afterward, they completed a race IAT before having an unanticipated social interaction with a Black experimenter. It was predicted that those who revealed relatively more negative attitudes toward Blacks on the IAT would behave in a relatively less friendly fashion toward the Black experimenter. This finding would substantiate the predictive validity of the IAT and suggest that it assesses individuals' idiosyncratic attitudes.

Two other empirical questions were also examined. First, would the IAT relate to explicit reports of prejudice? Previous research on implicit measures has yielded mixed results. Second, would explicit reports of prejudice relate to behavior toward the Black experimenter? It was our belief that the likelihood of observing significant relations between explicit measures of prejudice and other outcomes (i.e., IAT, behavior) would be improved under conditions in which participants felt minimal presentational concerns.

METHOD

Participants

At Michigan State University, 42 White undergraduates enrolled in introductory psychology courses participated in exchange for extra credit.

Measures

Explicit measures of prejudice. Participants completed semantic differential scales for Blacks, semantic differential scales for Whites, a feeling thermometer for Blacks, and a feeling thermometer for Whites (in that order). Each measure was completed on a separate page in a questionnaire booklet. Seven-point scales were used for the semantic differential word pairings: beautiful–ugly, good–bad, pleasant–unpleasant, honest–dishonest, and nice–awful. Participants also reported their attitudes toward Blacks and Whites using a feeling thermometer, which ranged from 0° (*extremely unfavorable*) to 100° (*extremely favorable*).

IAT task. Participants completed a word-based IAT task, which presented 96 stimulus words: 24 Black-associated names (e.g., Jamal and Yolanda), 24 White-associated names (e.g., Fred and Mary Ann), 24 desirable words (e.g., wonderful and awesome), and 24 undesirable words (e.g., offensive and disgusting). Names were always presented in

TABLE 1
Trial Blocks Used in the IAT Task

Block(s)	Type of judgment	Left key	Right key
1	Name discrimination	Black	White
2	Adjective discrimination	Undesirable	Desirable
3 and 4	Prejudice consistent combination	Black or Undesirable	White or Desirable
5	Reversed name discrimination	White	Black
6 and 7	Prejudice inconsistent combination	White or Undesirable	Black or Desirable

Note. Left key refers to categories associated with the "D" response, and right key refers to categories associated with the "K" key response.

uppercase letters, and adjectives were always presented in lowercase letters.

The IAT task was based on Greenwald et al. (1998), using a computer program written by the first author. As Table 1 reports, participants encountered five types of trial blocks across seven different blocks, with each block being composed of 48 trials. For half of the participants, Blocks 3 and 4 presented the prejudice-inconsistent combination and Blocks 6 and 7 presented the prejudice-consistent combination (the left key and right key response options for Blocks 1 and 5 were also reversed). This block order manipulation did not produce any effects and thus receives no further discussion. In Blocks 1, 2, and 5, each of the 48 relevant stimulus words was presented once based on a randomly determined order. In each of the combination blocks, the word types were alternated across trials (i.e., name, adjective, name, adjective, and so forth) with individual stimulus words selected at random from their respective lists until each of the 48 relevant items had been presented once across the two blocks.

Participants were told that they would be making a series of category judgments. On each trial, a stimulus word was displayed in the center of a computer window (24-point black serif text on a gray background), and participants used the "D" or "K" key on the keyboard for their responses. Category label reminders were displayed in blue text on the left and right sides of the window. Participants were told, "Make your judgments as rapidly as possible, but don't respond so fast that you make many errors. Occasional errors are okay. If you do make a mistake, a red X will appear on the screen below the target word. Please press the correct category key to continue. You cannot continue until you make the correct response." Participants were told to keep their index fingers on the "D" and "K" keys throughout the experiment to minimize delays in responding. A 250-ms gray-screen intertrial interval was used. In between blocks, participants were given a self-paced break and instructions for the next block.

Procedure

Participants arrived at the laboratory for an experiment on "word perception" and were greeted by a White female experimenter.[1] They were run individually. Unbeknown to the participant, a hidden video camera was positioned to record the participants' and experimenters' full bodies and their entire range of movements during scripted social interactions. A hidden unidirectional microphone recorded their discussions. They were directed to a rolling desk chair initially positioned 120 cm away from the experimenter's chair, allowing participants to establish a preferred distance from the experimenter. The experimenter explained that because the experiment was brief, the participant would complete four unrelated tasks. For the first task, they were told that the Department of Psychology had asked experimenters to interview students about their experiences in psychology. The experimenter asked the participant four innocuous questions (e.g., "What would you change to improve psychology classes?"), pausing for the participant's response between each question and recording the responses on a report form. The experimenter also told a scripted joke following the second question. This interaction took about 3 min.

Next, participants completed a booklet of questionnaires that purportedly were being used to develop future experiments. They were told that it was important for them to answer honestly in order for the future research projects to be successful. The privacy of their responses was stressed by explaining that they would complete the booklet in a private room, place the completed booklet in a sealed envelope, and drop it into a covered box without any experimenter interaction. The booklet contained several questionnaires, only some of which were relevant to the current study. After completing several pages of the booklet, participants completed the semantic differential scales and the feeling thermometer measures. It took participants about 15–20 min to complete the booklet.

While the participant was completing the booklet, the White experimenter assessed her interaction with the participant (details forthcoming). After completing the booklet, participants inserted the sealed survey into a covered box in the laboratory's waiting area. They then found the experimenter, who took them to a private computer workstation to

[1] The sequence of events that participants experienced in the experiment was fixed to minimize suspicion about the overall goals of the study (e.g., initially encountering a Black experimenter may have raised immediate concerns that the study was about racism). Although it is possible that exposure to one's own responses on the IAT or the explicit prejudice measures might affect subsequent behavior toward the Black experimenter, we reasoned that because the interaction with the Black experimenter was unexpected, participants would find it difficult to control their subtle behavioral cues toward her in an extemporaneous social interaction. However, we acknowledge that a fixed-order design may introduce the possibility of unforeseen confounds in the current study.

begin the "word perception" experiment (i.e., the IAT). The experimenter then looked at the clock and mentioned that her shift was almost over and that a new experimenter would assist in completing the fourth task following the word perception experiment. At that point, the White experimenter started the IAT program and excused herself. Participants required about 10 min to complete the IAT.

While participants completed the IAT, a Black female experimenter replaced the White experimenter and greeted participants after they returned from their room after completing the IAT. Once again, the participant was directed to a chair positioned 120 cm from the experimenter's chair, allowing the participant to establish a preferred seating distance. The Black experimenter asked the participant seven questions about the experiment (e.g., "What did you think about the difficulty level of the computer task?" and "Were the instructions clear"?), pausing for the participant's response between each question and recording the responses on an interview form. She also told a scripted joke after the fourth question. Afterward, the experimenter explained that both social interactions had been videotaped, and she asked for the participant's permission to use the videotape for data analyses. One participant refused, and her videotape was erased in her presence, leaving 41 participants for data analyses. Finally, participants were debriefed and thanked for their participation.

Coding of Social Interactions

Trained judges' ratings of participants' behavior. Based on the existing literature documenting behavior cues that convey emotions and attitudes (Crosby, Bromley, & Saxe, 1980; DePaulo, 1992; DePaulo & Friedman, 1998; Duncan, 1969; Eckman & Friesen, 1967; Hendricks & Bootzin, 1976; Kleinke, 1986; Word, Zanna, & Cooper, 1974), 16 behaviors were coded by two trained judges who were unaware of participants' attitudes. Using a scale from 1 (*none*) to 9 (*very much*), judges rated the participant's friendliness during the interaction, the abruptness or curtness of the participant's responses to questions, the participant's general comfort level, how much the participant laughed at the experimenter's joke, and the amount of participant's eye contact with the experimenter. On 5-point scales, they assessed the participant's forward body lean toward the experimenter (vs leaning away), the extent to which the participant's body faced the experimenter (vs facing away), the openness of the participant's arms (vs crossed arms), and the expressiveness of the participant's arms (vs not moving at all). Judges also calculated the distance between the experimenter and the participant's chair at the end of the interaction to gauge social distance. Judges also recorded the participant's speaking time, number of smiles, number of speech errors, number of speech hesitations (e.g., "um"), number of fidgeting body movements (e.g., swinging feet and shifting positions), and number of extemporaneous social comments made by the participant.[2] The judges rated each participant's interaction with the White experimenter and with the Black experimenter separately. The videotape showed both the participant and the experimenter, and the judges were instructed to only attend to the audio for ratings associated with the interaction dialogue (e.g., curtness of responses).

Experimenters' ratings. Each experimenter completed a 5-item inventory after their interaction with the participant. Using a scale ranging from 1 (*not at all*) to 9 (*extremely*), experimenters recorded their assessment of the participant's degree of eye contact, the abruptness or curtness of the participant's responses, the participant's friendliness, the participant's perceived comfort level during the interaction, and the experimenter's own comfort level during the interaction.

RESULTS

Data Reduction

Overview. The data analytic strategy was to transform all measures, implicit and explicit, into difference scores that reflected the relative degree of prejudice against Blacks (i.e., relatively more positive attitudes toward Whites than Blacks and relatively more positive behaviors toward Whites than Blacks). Thus for *all measures,* larger positive scores reflected greater negativity toward Blacks than Whites.

IAT. To reduce the positive skew inherent in response latency data (Greenwald et al., 1998; Ratcliff, 1993), a log transformation was applied to each response latency. IAT effect scores were computed by comparing mean response latency of trials in Blocks 3 and 4 to trials in Blocks 6 and 7. The accuracy of any given trial was ignored, and extreme latencies were recoded such that those less than 300 ms were scored as 300 ms and those greater than 3000 ms were scored as 3000 ms.[3] The mean response latency for the prejudice-consistent block trials was subtracted from the mean response latency for the prejudice-inconsistent block trials. Thus, larger positive IAT effect scores reflected relatively stronger negative Black attitudes and relatively stronger positive White attitudes.

Explicit measures of prejudice. The five semantic differential scales revealed good reliability for Blacks ($\alpha = .91$) and for Whites ($\alpha = .89$). Thus, the mean of each set of scales was calculated, and a difference score was

[2] Readers may contact the authors for details about the behavior coding protocols.

[3] Analyses were also conducted discarding responses in Blocks 3 and 6 (which presumably are more sensitive to task learning effects), as reported by Greenwald et al. (1998). Identical results obtained. Additional analyses using other trimming criteria (e.g., omitting incorrect trials, omitting trials with responses slower than 2 standard deviations from the mean) produced equivalent results.

computed such that larger scores reflected holding more positive attitudes toward Whites than Blacks. A difference score was also computed for the feeling thermometer (subtracting the Black thermometer from the White thermometer). Because both of these difference scores were strongly related, $r = .45$, $p < .01$, each difference score was standardized and the two z scores were added to create the explicit measure of prejudice score, which reflected the overall relative degree to which participants held more positive attitudes toward Whites than Blacks.

Experimenters' ratings of the interaction. The White and Black experimenters' ratings of their social interactions were examined. Difference scores were computed for each assessment, whereby larger scores reflected more positive behavior being perceived by the White experimenter than by the Black experimenter. These five difference scores showed good reliability ($\alpha = .81$), thus an experimenters' rating score was calculated based on the sum of the five (standardized) difference scores. Thus, positive values on this experimenters' rating score represented the extent to which the White experimenter, compared to the Black experimenter, reported a more positive social interaction.

Judges' ratings of the interaction. Two trained judges assessed the videotapes for positive and negative behaviors revealed by the participants, independently assessing each participant's interaction with both the Black and the White experimenter. Difference scores were calculated such that positive values always reflected relatively greater positivity being exhibited toward the White experimenter than toward the Black experimenter. These 16 difference score ratings were divided into two categories: molar judgments that captured overall interaction quality and specific social behaviors.

The molar judgments (interjudge agreement in parentheses) included abruptness or curtness of participant's responses ($r = .48$, $p < .01$), participant friendliness ($r = .43$, $p < .01$), and participant's general comfort level ($r = .53$, $p < .01$). Because of the good interjudge agreement, the mean of the judges' (standardized) differences scores were computed. These three difference scores revealed good reliability ($\alpha = .78$), thus a judges' molar rating was computed by taking the sum of the three difference scores. Therefore, more positive values reflected relatively more positive behaviors being exhibited toward the White experimenter than toward the Black experimenter.

In addition to the molar ratings, the judges also assessed specific participant behaviors for evidence of bias between the experimenters. Each judge's rating was standardized and a difference score was computed whereby larger scores reflected more positive behavior being revealed to the White experimenter than to the Black experimenter. The judges showed good interjudge agreement in their difference scores (see Table 3), thus the mean of their difference scores was computed for each of the 13 specific behaviors.

TABLE 2
Implicit and Explicit Measures of Prejudice Means, Effect Sizes, and Comparisons to Zero (i.e., No Group Preference)

Measure	*M*	Cohen's *d*	*t*(40)
IAT effect	162.81 ms	0.88	11.47**
Explicit prejudice measure difference scores			
Semantic differential	0.22	0.27	2.58*
Feeling thermometer	11.34°	0.63	4.52**

Note. Larger, positive values reflect relatively more positive attitudes toward Whites than Blacks. IAT effect size and inferential statistics were performed on log-transformed values, but the IAT effect mean is reported in a real-time metric. $N = 41$.

* $p < .05$.

** $p < .001$.

Descriptive Analyses

As Table 2 reveals, significant racial bias was exhibited in participants' implicit and explicit measures of prejudice. That is, participants revealed more positive attitudes toward Whites than Blacks on the IAT, semantic differential, and feeling thermometer measures. The effect size was large for the IAT, moderate for the feeling thermometer, and small for the semantic differential (Cohen, 1988). The IAT effect size observed is consistent with previous research (Greenwald et al., 1998). In contrast, Greenwald et al. only found significant racial bias for one of two explicit measures of prejudice (i.e., feeling thermometer), whereas significant prejudice was found in both explicit measures in the current study.

Correlational Analyses

Zero-order correlations between the IAT effect, the explicit measure of prejudice score (i.e., the combination of the feeling thermometer and semantic differential difference scores), the experimenters' ratings, the judges' molar ratings, and the judges' ratings of 13 specific biased social behaviors were calculated. With respect to the primary hypothesis, Table 3 reveals that there were significant correlations between the IAT and the experimenters' rating of social interaction bias and between the IAT and the judges' molar ratings of social interaction bias. Specifically, as participants' IAT scores reflected relatively more positive attitudes toward Whites than Blacks, social interactions were more positive toward the White experimenter than the Black experimenter as assessed both by trained judges and by the experimenters themselves. In addition to finding evidence that the IAT related to the experimenters' and judges' molar assessments, larger IAT effect scores predicted greater speaking time, more smiling, more extemporaneous social comments, fewer speech errors, and fewer speech hesitations in interactions with the White (vs Black) experimenter.

TABLE 3

Correlations between IAT, Explicit Measures of Prejudice, Experimenters' Ratings, and Judges' Molar Ratings and Assessments of Biased Participant Social Behavior

	Prejudice measures		Social interaction bias ratings	
	IAT	Explicit	Experimenters'	Judges' molar
Explicit measure of prejudice	.42**			
Experimenters' ratings	.39*	.33*		
Judges' molar ratings	.34*	.26	.41**	
Biased participant social behaviors				
Forward leaning (.64***)	−.26	.12	.05	−.08
Facing experimenter (.77***)	−.03	−.08	.31*	−.03
Body openness (.47**)	.17	.02	.20	.43**
Expressiveness (.60***)	.09	−.20	.00	.25
Eye contact (.35*)	.25	.20	.20	.55***
Seating distance (.69***)	.26	.14	.31*	.15
Speaking time (.85***)	.51**	.18	.41**	.30
Smiling (.71***)	.39*	.21	.15	.28
Speech errors (.53***)	.42*	.05	.14	−.03
Speech hesitation (.53***)	.35*	.13	−.07	.11
Fidgeting (.42**)	−.06	−.15	.00	.02
Laughter at joke (.56***)	.19	.03	.27	.35*
Social comments (.46**)	.32*	.02	.12	.44**

Note. All measures are coded such that larger, positive values reflect relatively more positive attitudes and behaviors toward Whites in comparison to Blacks. Values in parentheses indicate interjudge correlations. $N = 41$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

In addition to providing strong support for the primary hypothesis, the correlational analyses addressed the two empirical questions raised in the introduction as well. First, a significant correlation between the IAT and explicit reports of prejudice was observed. Specifically, as participants revealed relatively more positive attitudes toward Whites than Blacks on the IAT, they reported more positive evaluations of Whites than Blacks on the explicit measures of prejudice. The second empirical question received mixed support. That is, the explicit measures of prejudice score was positively related to experimenters' ratings of biased interactions, indicating that participants who reported relatively more positive attitudes toward Whites on explicit measures were perceived as more positive in social interactions by the White experimenter than by the Black experimenter. However, the explicit measures of prejudice score was unrelated to any of the judges' ratings (molar or specific social behaviors).

Finally, the remaining correlations addressed what was related to the judges' ratings. For instance, the experimenters' ratings of biased social interaction were positively related to the judges' molar ratings, indicating significant agreement between the experimenters' ratings and the judges' molar ratings of social interaction bias. The experimenters' judgments also corresponded (all in the expected direction) with judges' ratings of bias in terms of facing the experimenter, seating distance, and speaking time. In other words, the experimenters appeared to be especially sensitive to facing the experimenter, social distance cues, and speaking time as factors that related to their perceptions of biased social interaction. Finally, the judges' molar ratings were in correspondence with their specific ratings of body openness, eye contact, laughter at the scripted jokes, and extemporaneous social comments (all in the expected direction).

DISCUSSION

The current work is the first study to demonstrate relations among the IAT, intergroup discrimination, and explicit measures of prejudice. Although the IAT has become popular because of its large effect size and difficulty to inhibit, any psychological tool is only as good as its ability to predict human behavior. Indeed, it was found that the IAT was related to biases in intergroup social interactions. Therefore, researchers can be confident that attitudes assessed by the IAT do relate to intergroup behavior. These findings also suggest that the IAT does assess personal attitudes in that idiosyncratic variability in implicit measures of prejudice was related to behavior. Moreover, the ability of the IAT (unlike explicit measures of prejudice) to predict several specific biased social behaviors as assessed by independent observers is consistent with the claim that implicit measures of attitudes are especially predictive of behavioral leakage (Dovidio et al., 1997).

In addition to establishing a link between the IAT and discriminatory behavior, the current study also found a relation between the IAT and explicit measures of prejudice. Previous work by Greenwald et al. (1998) found no such relation. Further, those researchers did not observe as strong of evidence of racial prejudice in their explicit measures. The current experiment, in contrast, found reliable evidence of both. In the Greenwald et al. study, participants completed explicit measures of racial prejudice *after* completing the IAT. Because of the transparency of the IAT, it is conceivable that their methodology sensitized participants to the overall purpose of the entire study, increasing the likelihood that their subsequent explicit reports were influenced by social desirability concerns more so than by their personal attitudes. The current study, in contrast, attempted to minimize these concerns by having participants complete the IAT after the explicit measures. Accordingly, strong prejudice was found on explicit measures and it was related

to IAT scores. Whether methodological differences between the current study and Greenwald et al. account for this outcome is unclear because task order was not manipulated in this study. However, the current findings suggest that implicit and explicit measures may tap the same attitude representation, though clearly correlational evidence is far from unequivocal. Yet, to the extent that parsimony is desirable, the position that implicit and explicit attitude measures tap similar knowledge has considerable appeal.

Although the current work found that the IAT predicted discrimination and explicit measures of prejudice, some potential limitations should be acknowledged. For example, the sequence of events that participants experienced was fixed in order to minimize suspicion about the purpose of the study. Therefore, it would be desirable to manipulate the order of events in future research to ensure that the fixed order did not produce unintended consequences. Also, the design of the study resulted in participants having completed the explicit measures of prejudice and the IAT just before interacting with the Black experimenter. Because these tasks would result in the conscious activation of racial attitudes, accessibility of these attitudes would be quite high when they encountered the Black experimenter. This greater accessibility makes it more likely that attitude–behavior consistency would be exhibited (e.g., Fazio, Powell, & Williams, 1989; Fazio & Williams, 1986; Snyder & Kendzierski, 1982; Snyder & Swann, 1976). Also, greater attitude accessibility might increase the likelihood that the experimenter would be categorized as "Black" rather than as a member of another applicable social category (Smith, Fazio, & Cejka, 1996), making it more likely that participants' racial attitudes would predict their behavior toward the experimenter. It seems reasonable to assume that such attitude–behavior consistency and categorization effects would naturally occur for those who chronically have highly accessible racial attitudes, but the question remains open as to the implications of attitude expression for individuals whose attitudes are, typically, not highly accessible. Future research should explore whether the expression of group attitudes and its subsequent effects on activation results in different behavior being exhibited from those who vary in attitude accessibility.

The current study also provides insights for researchers considering how to assess intergroup interactions. Our approach was to rely both on experimenters' perceptions (Fazio et al., 1995) and trained judges' assessments of videotapes (Dovidio et al., 1997) to examine behavior. In the current study, many more participant social behaviors were coded than were examined by Dovidio et al., who only reported examining time talking, eye contact with the experimenter, and number of eye blinks. Although both the judges and experimenters assessed the same interactions and showed reliable agreement in their reports, the two groups differed in some respects. For instance, experimenters used naive theories for assessing behavior while engaging in a demanding social interaction, whereas the judges had the benefit of more cognitive resources, the opportunity to replay the interactions, and exposure to the scientific literature on assessing social behavior.

In one sense, the correlation between our experimenters' reports and our judges' molar ratings suggests that the labor-intensive effort required to code specific social behaviors may not be necessary. However, the judges' ratings of specific social behaviors revealed five relations with the IAT but none with the explicit measures of attitudes. This is consistent with Dovidio et al. (1997), who found that only implicit measures of prejudice related to nonverbal behavior. Without collecting the judges' ratings in the current study, this asymmetry between implicit and explicit measures would have gone undetected. Another interesting finding was the discrepancy between the experimenters' ratings and the judges' molar ratings with respect to how each related to the specific behaviors coded by the judges. At the present time, we are far from a complete understanding of what leads to differences between the experimenters' and judges' assessments of social interactions. Future work needs to address this issue, however, because it is clear that each approach to assessing intergroup behavior is capturing something slightly different. Despite this uncertainty, we feel quite confident about the demonstration of the predictive validity of the IAT in the current study because it was reliably related to the independent assessments of social interactions offered by the experimenters and by the trained judges.

Finally, the current study reiterates the importance of making behavior the ultimate criterion for the value of psychological methods. Across the history of social psychology, the value of studying attitudes has been called into question because of concerns that attitudes do not predict behavior (e.g., LaPiere, 1934; Wicker, 1969), are beyond one's introspective capability (e.g., Nisbett & Wilson, 1977; Wilson, Hodges, & LaFleur, 1995), or are often influenced by normative pressures (e.g., Ajzen & Fishbein, 1973, 1980). More recent treatments of attitudes have recognized that cognitive associations, often those beyond our awareness, greatly influence our behavior (Bargh & Chartrand, 1999; Wegner & Bargh, 1998), especially when normative pressure is minimal (Fazio, 1990). The thrust of this emerging perspective is that indirect assessment of attitudes may not only be valuable to circumvent problems such as social desirability, but may be crucial to assess the mechanisms that often direct behavior. It is clear that the IAT holds much promise as a tool to assess attitudes, and the current work demonstrates its predictive utility. However, future work will be required to better understand the mechanisms that underlie the IAT and to predict when it will, and will not, relate to explicit measures of attitudes. At the very least, the current work suggests that such efforts can proceed with the

assurance that the IAT assesses personal attitudes that relate to social behavior in meaningful ways.

REFERENCES

Ajzen, I., & Fishbein, M. (1973). Attitudinal and normative variables as predictors of specific behaviors. *Journal of Personality and Social Psychology,* **27,** 41–57.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice Hall.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist,* **54,** 462–479.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of black and white discrimination and prejudice: A literature review. *Psychological Bulletin,* **87,** 546–563.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin,* **111,** 203–243.

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 3–40). New York: McGraw–Hill.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology,* **56,** 5–18.

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

Duncan, S. (1969). Nonverbal communication. *Psychological Bulletin,* **72,** 118–137.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin,* **23,** 316–326.

Eckman, P., & Friesen, W. V. (1967). Head and body cues in the judgment of emotion: A reformulation. *Perceptual and Motor Skills,* **24,** 711–724.

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

Fazio, R. H. (1995). Attitudes as object–evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.

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

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process? *Journal of Consumer Research,* **16,** 280–288.

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.

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.

Hendricks, M., & Bootzin, R. (1976). Race and sex as stimuli for negative affect and physical avoidance. *Journal of Social Psychology,* **98,** 111–120.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin,* **100,** 78–100.

LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces,* **13,** 230–237.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review,* **84,** 231–259.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2000). Harvesting implicit group attitudes and stereotypes from a demonstration website. Unpublished manuscript, Yale University.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology,* **75,** 811–832.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin,* **114,** 510–532.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition,* **17,** 437–465.

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology,* **71,** 888–898.

Snyder, M., & Kendzierski, D. (1982). Acting on one's attitudes: Procedures for linking attitude and behavior. *Journal of Experimental Social Psychology,* **18,** 165–183.

Snyder, M., & Swann, W. B. (1976). When actions reflect attitudes: The politics of impression management. *Journal of Personality and Social Psychology,* **34,** 1034–1042.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 446–496). New York: McGraw–Hill.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues,* **25,** 41–78.

Wilson, T. D., Hodges, S. D., & LaFleur, S. J. (1995). Effects of introspecting about reasons: Inferring attitudes from accessible thoughts. *Journal of Personality and Social Psychology,* **69,** 16–28.

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

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology,* **10,** 109–120.