Social Cognition, Vol. 19, No. 2, 2001, pp. 97-144

IMPLICIT ATTITUDES AND RACISM: EFFECTS OF WORD FAMILIARITY AND FREQUENCY ON THE IMPLICIT ASSOCIATION TEST

Scott A. Ottaway, Davis C. Hayden, and Mark A. Oakes
Western Washington University

Greenwald, McGhee and Schwartz (1998) described a new method-the Implicit Association Test (IAT)-for unobtrusively measuring racial attitudes. This article assesses the validity of the IAT by investigating whether Greenwald et al.'s implicit racism findings resulted from two confounds present in their studies: differential familiarity and frequency of the words that comprised their target concepts. Experiment 1 produced large IAT effects when both low and high familiarity words comprised nonsocial target categories (insects and flowers) and demonstrated that the IAT is more sensitive when high familiarity exemplars form the target concepts. In Experiment 2, we obtained large implicit racism effects for both African American and Hispanic racial groups even when the familiarity and frequency of the names that comprised the racial categories were controlled and even though participants described themselves as unprejudiced. Additionally, explicit self-reports of racial attitudes were only weakly related to the IAT measures. These experiments indicate that (a) although familiarity clearly exerts an important influence on the IAT, the use of low familiarity stimuli does not eliminate the sensitivity of the IAT, (b) stimulus familiarity and frequency can not account for the implicit racism effect and (c) stimulus familiarity is an important moderating variable that can influence the sensitivity of

This research was supported by a Bureau of Faculty Research award from Western Washington University to Scott Ottaway. A portion of the manuscript was completed while the first author was a visiting researcher at Indiana University in Bloomington with John Kruschke.

We wish to thank Tony Greenwald, Russ Fazio, John Skowronski and Bill von Hippel for their valuable comments. Additionally, Richard J. Harris, Theresa Treat, George Cvetkovich, Saera Khan and David Baggenstos are to be thanked for their helpful suggestions. Cristy Silverman and Nick Bissell assisted with data collection. We also thank Sean Draine for making his experiment development (www.microsecond.com) freely available and Bernd Wittenbrink and Charles Judd for the Discrimination and Diversity scales.

Please address correspondence to Scott Ottaway, Department of Psychology, Western Washington University, Bellingham, WA. 98225, (360) 650-3073, or via electronic mail to ottaway@wwu.edu. The author's World-Wide-Web page is at www.wwu.edu/~ottaway. Mark Oakes is now at the University of Washington.

implicit attitude measures. We discuss the results in relation to the validity of the Implicit Association Test and theories of implicit social cognition.

Since the mid 1980s, numerous researchers have focused their effort on examining a variety of types of implicit cognition, including implicit memory (Roediger, 1990a; Schacter, 1987), implicit learning (Stadler & Frensch, 1997), implicit reasoning (Sloman, 1996), and implicit social cognition (Fazio, 1990; Greenwald & Banaji, 1995; Lewicki, 1986). One assumption underlying these investigations is that, under some conditions, participants are not aware of knowledge (memories, attitudes, beliefs) that influences their behavior (recall, reaction times, evaluations). Recently, a great deal of research and theorizing has focused on understanding the implicit nature of attitudes, stereotypes and bias (Banaji & Greenwald, 1994; Devine, 1989; Dovidio, Evans, & Tyler, 1986; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton & Williams, 1995; Greenwald & Banaji, 1995; Jacoby, Kelley, Brown, & Jasechko, 1989; Smith & Branscombe, 1988; Srull & Wyer, 1979; Wittenbrink, Judd, & Park, 1997). Although both implicit and explicit attitudes derive from an individual's previous experiences, some researchers argue that implicit and explicit attitudes are conceptually distinct cognitive processes (Devine, 1989; Greenwald & Banaji, 1995). Specifically, explicit attitudes are deliberate evaluative dispositions that are amenable to introspection and under conscious control. In contrast, implicit attitudes are automatic evaluative dispositions that typically occur without conscious reflection (Fazio, 1990) and are not necessarily available for introspection or conscious control (Banaji & Bhaskar, 2000; Greenwald & Banaji, 1995).

The purpose of this paper is to evaluate a measure of attitudes—the Implicit Association Test—developed by Greenwald, McGhee, and Schwartz (1998). The Implicit Association Test (IAT) is an indirect measure[1] of attitudes because it does not directly ask participants to

1. We refer to The Implicit Association Test as an "indirect measure" similar to the terminology used by Greenwald, McGhee, and Schwartz (1998). The term "implicit measure" is considered by many to be equivalent because both terms indicate that one assesses attitudes without directly asking participants to reveal their attitudes. The implicit and explicit memory literature has contained a debate about the appropriateness of the term implicit measure (cf., Richardson-Klavehn & Bjork, 1988; Roediger, 1990b). For the current purposes, both terms are synonymous in that they indicate unobtrusive measures that (a) do not rely upon self-report and (b) attempt to minimize the participants' awareness that their attitudes are being assessed. Similarly, we treat the terms "implicit attitude" and automatic attitude as synonymous because both terms suggest that a subject may at some time be unaware of their attitudes.

verbalize their attitudes; rather, one infers attitudes based on response latencies to evaluatively congruent and incongruent concepts (e.g., flower + pleasant; insect + pleasant). One important feature of the IAT task is that it allows unobtrusive examination of both universally held as well as socially charged attitudes (e.g., racial prejudices). We report two experiments that examine whether stimulus characteristics (subjective familiarity and frequency of occurrence) affect IAT performance, independent of an individual's implicit or explicit attitude. Of particular interest is whether Greenwald et al.'s findings suggesting the wide-spread prevalence of implicit racial prejudices can be explained by their confounding of name frequency/familiarity across racial groups. Our experiments also examine the relationship between the IAT and explicit racial attitude measures, the external validity of the IAT by examining attitudes toward Hispanics and the sensitivity of the IAT to gender differences in attitudes.

ATTITUDES, SELF-PRESENTATION FORCES AND INDIRECT MEASURES

Attitudes are favorable or unfavorable responses and/or biases to an object. The measurement of attitudes traditionally has involved the use of various self-report evaluation methods including adjective checklists and rating scales (e.g., Katz & Braly, 1933; McConahey, Hardee, & Batts, 1981). However, these measures are susceptible to selfpresentation effects: participants can conceal attitudes that they do not wish to express. For example, an individual may be reluctant to express a negative bias toward Hispanics in order to avoid being labeled as prejudiced. Whether due to internal (self-presentation to self, Greenwald & Breckler, 1985) or external (impression formation of others) concerns, self-presentation forces can lead to an inaccurate report of one's attitudes. Moreover, some attitudes may not be accessible to an individual and therefore can not be measured by verbal report measures (Greenwald & Banaji, 1995). A number of researchers have developed measures of attitudes intended to minimize the conscious activation of self-presentation defenses (Banaji & Greenwald, 1994; Bargh, Chaiken, Govender, & Pratto, 1992; Devine, 1989; Dovidio et al., 1986; Fazio, Sanbonmatsu, Powell & Kades, 1986; Fazio et al., 1995; Gaertner & McLaughlin, 1983; Greenwald et al., 1998). These measures are indirect in that one infers attitudes based upon examination of response latencies rather than self-report measures. The assumption underlying these measures is that paired concepts that are responded to more quickly share an association different from that of paired concepts responded to more slowly. This assumption is consistent with numerous associative network models of memory (e.g., Anderson & Bower, 1973), as well as models of human category learning (Kruschke, 1992; Rosch, 1975).

The Implicit Association Test operates in a fashion similar to other indirect measures and shares the same goal: to measure the strength of association between target concepts. In the IAT, participants complete a series of discriminations that involve a target concept and an evaluative dimension. In the single discrimination tasks, members of one concept (e.g., pleasant words) are assigned to one response key while members of a contrasting concept (e.g., unpleasant words) are assigned to another response key. For the combined discrimination tasks, one concept and attribute value are assigned to a single response key (e.g., African American names and pleasant words) while the contrasting concept and attribute are assigned to the other response key (e.g., Caucasian names and unpleasant words). Instructions during both the single and combined discrimination tasks emphasize that participants should categorize concept members as rapidly as possible.

For example, consider Greenwald et al.'s (1998) Experiment 3, which examined attitudes to a racial out-group. In this IAT study the target-concept was racial identity and the evaluative dimension consisted of pleasantness (see Table 1 for an overview). In Task A, participants classified names as African American with one hand and Caucasian with the other hand. In Task B, participants discriminated between pleasant and unpleasant words. Task C consisted of a combined discrimination task. Caucasian names and pleasant words (Caucasian + pleasant) shared a response key assigned to the left hand while African American names and unpleasant words (African American + unpleasant) shared a response key assigned to the right hand. In Task D, participants repeated the first discrimination task (African American vs. Caucasian) except for a reversal of response key assignments. In Task E, participants completed a reversed combined-discrimination task in which the racial groups were paired with the opposite evaluative attribute (i.e., Caucasian + unpleasant; African American + pleasant).

TABLE 1. Overview of the Implicit Association Task (upper panel) and Examples of Stimuli (lower panel)

	Task Sequence				
	A	B	C	D	E
Response	Target concept discrimination	Attribute discrimination	First combined task	Reversed target concept discrimination	Reversed combined task
Left hand	Caucasian names	pleasant	Pleasant + Caucasian	African American names	Pleasant + African American
Right hand	African American names	unpleasant	Unpleasant + African American	Caucasian names	Unpleasant + Caucasian
	Sample Stimuli				
	A	B	C	D	E
Left hand	Chip/Harry/Josh	caress/freedom/health	Paul/miracle/Brandon	Theo/Alphonse/Terryl	Lionel/honest/Alonzo
Right hand	Lavan/Marcellus/Rashaun	abuse/crash/filth	Lamont/disaster/Jamel	Brad/Greg/Jonathan	Hank/poverty/Roger
Measures	Rts Errors	RTs Errors	RTs Errors	RTs Errors	RTs Errors

Note. The Implicit Association Task consists of a series of two-choice discrimination tasks (see text for a complete description). Three tasks involve discriminations between exemplars from two target concepts (Tasks A and D) or discriminations of opposite poles of an attribute dimension (Task B). In Task C, the target and attribute tasks are combined. In this example, Caucasian names and pleasant words are assigned to the left hand and African American names and unpleasant words are assigned to the right. In Task E, the target and attribute combinations are reversed: African American names and pleasant words are assigned to the left hand and Caucasian names and unpleasant words are assigned to the right hand. Attitudes towards target concepts are assessed by examining the difference between the two combined tasks (e.g., Task E - Task C). In the full IAT, a variety of task-order combinations are examined. The IAT allows assessment of attitudes towards a variety of target concepts (e.g., flowers and insects, Caucasian and Hispanic names). The first names, target concepts and attribute dimension in this figure are taken from Greenwald, McGhee, and Schwartz (1998, Experiment 3). A similar IAT method is used in Experiment 2 with the target concepts of African American, Hispanic and Caucasian names. Rts = response times.

The difference between response latencies on the two combined tasks (Tasks C and E) serves as the IAT measure of attitudes. If participants differentially associate the target concepts with the attribute dimension, then discriminations that involve pairing of evaluatively equivalent concepts should be faster than discriminations involving evaluatively less similar concepts. In short, Greenwald et al. argued that the difference in response latencies on the two combined tasks-the IAT effect-provides an indirect measure of differences in attitudes toward the target categories. It is important to highlight that the IAT task consists of a number of built-in methodological controls for variables that could influence response latencies. These controls include the use of the same stimuli in all combined tasks and counterbalancing of order of tasks, response key assignments, and pairings of target and attribute dimensions.

In Greenwald et al.'s (1998) third experiment, participants responded more slowly during blocks that associated African American names and pleasant words than during blocks when African American names and unpleasant words shared a response key. The average difference was 179 ms with an effect size of $d = 1.16$ (by convention, $d \geq .80$ is considered a large effect). Greenwald et al. interpreted this finding as an implicit attitudinal preference for Caucasian over African American, an effect they labeled implicit racism. Moreover, they reported only a weak relationship between their IAT measure and explicit measures of attitudes and racist beliefs (average $r < .20$). This dissociation calls into question the validity of explicit measures, or, at the very least raises the question of which measure, implicit or explicit, accurately reflects an individual's prejudice. Based on these and other results, Greenwald et al. (1998) concluded that their IAT findings "are discouraging in indicating the pervasiveness of unconscious forms of prejudice" (p. 15).

Prior to the acceptance of the IAT measure as an index of automatic attitudes, validity issues and alternative interpretations of the IAT must be addressed. As evidence of the validity of the IAT task, Greenwald et al. reported another IAT experiment that used nonsocial target concepts (flowers vs. insects; musical instruments vs. weapons) and the attribute dimension of pleasantness. The intention was to examine the sensitivity of the IAT to near-universal attitudes (flower + good, insect + bad; musical instrument + good, weapon + bad). As predicted, the IAT method indicated a preference for flowers over insects and musical instruments over weapons. Importantly, explicit attitude measures of the target concepts correlated among themselves but not with the implicit attitude measures. Moreover, they also found larger correlations among the implicit measures in comparison to correlations among the explicit measures. Greenwald et al. argued that the small correlations between implicit and explicit attitude measures are consistent with similar findings by other researchers (e.g., Fazio et al., 1995) and a theoretical orientation that advocates for the distinction between implicit and explicit attitude constructs (Devine, 1989; Fazio, 1990; Greenwald & Banaji, 1995).

FAMILIARITY AND FREQUENCY EXPLANATIONS OF THE IMPLICIT RACISM EFFECT

One plausible alternative explanation of the IAT effect, suggested by Greenwald et al. (1998), is based on the possibility that subjects may have had differential prior exposure to the target concepts (e.g., African American and Caucasian first names). It is reasonable to assume that members of these categories may differ in their amount of prior exposure (i.e., familiarity and/or frequency of occurrence). Based on the assumption that preferred stimuli are responded to more quickly, Greenwald et al.'s implicit racism effect could be due to a positive bias toward more familiar and/or frequently occurring stimuli (e.g., Zajonc, 1968) This prior-exposure interpretation seriously questions the usefulness and discriminant validity of the IAT as an indirect measure of racial prejudice.

Although Greenwald et al. argued on rational grounds that this was an unlikely explanation of their IAT effect, they presented no empirical data to support their conclusion. Inspection of Greenwald et al.'s names suggests that their African American and Caucasian names were confounded with both the familiarity and frequency such that Caucasian names were more familiar and frequently occurring. To examine the magnitude of this confounding, we developed a list of over 300 male and female names commonly given to African American, Hispanic and Caucasian Americans. Participants then rated the names on a 5-point familiarity scale, with "1" corresponding to very unfamiliar and "5" corresponding to very familiar (details of these ratings are provided in Appendix A). Table 2 displays the mean familiarity ratings of Greenwald et al.'s names by racial cate-

TABLE 2. Mean Familiarity Ratings and Cumulative Frequency of Occurrence of First Names Used by Greenwald, McGhee, and Schwartz (1998) and in Experiment 2

Names from Greenwald, McGhee, and Schwartz's Experiment 3

	Familiarity		Cumulative Frequency	
	Females	Males	Females	Males
Race of Name	M (SD)	M (SD)	M (SD)	M (SD)
African American	1.77 (.51)	2.12 (.61)	7 (9)	8 (18)
Caucasian	4.18 (.69)	4.28 (.66)	40 (20)	32 (18)

Names from Experiment 2

	Familiarity		Cumulative Frequency	
	Females	Males	Females	Males
Race of Name	M (SD)	M (SD)	M (SD)	M (SD)
African American	2.77 (.50)	3.02 (.35)	19 (6)	13 (9)
Hispanic	2.80 (.78)	2.96 (.20)	29 (14)	10 (7)
Caucasian	2.91 (.71)	3.00 (.38)	31 (29)	9 (9)

Note. n = 156. Ratings of Familiarity are based on a 5 point Likert-like scale with 1 = "Very Unfamiliar" and 5 = "Very Familiar" and were completed by Caucasian American undergraduates. Cumulative Frequency statistics indicate the average percentage of names from a U.S. Census database, regardless of racial identity, that occur less frequently. The database contained the names of 6.1 million people. Cumulative Frequency values range from 0 to 90 with higher values indicating more frequent names. Statistics are based on n = 25 and n = 10 names for Greenwald et al. and Experiment 2, respectively. Appendix A describes the Familiarity and Frequency measures in detail.

gory and gender. Among Caucasian American undergraduates, Greenwald et al.'s African American names were much less familiar (average M = 1.94) than the Caucasian names (average M = 4.23). Thus, the confounding of familiarity with race of name could account for the implicit racism effect.

We also assessed the frequency of occurrence of Greenwald et al.'s names because familiarity and frequency are not necessarily equivalent constructs. Whereas high frequency names are often very familiar (e.g., Frank), low frequency names may or may not be familiar. For instance, some first names (e.g., Juanita) might be common to members of one racial group but, because this group contains relatively fewer members, these names will occur infrequently relative to all names. Thus, some infrequently occurring names may be highly familiar to some ethnic groups but not others. We assessed name frequency by determining each name's cumulative frequency in a U.S. Census database consisting of over 6 million names. This analysis, also presented in Table 2, indicated that Greenwald et al.'s

African American names (average $M = 7$) occurred less frequently than their Caucasian names (average $M = 36$). In summary, Greenwald et al.'s implicit racism effect could have resulted from a positive bias toward names that are either more familiar or that occur more frequently. Thus, it is a real possibility that the IAT method may not reflect implicit prejudicial attitudes to a racial out-group.

OVERVIEW

It is important to demonstrate that the results from the IAT are not due to a positive bias to more familiar and/or frequently occurring stimuli. In our experiments, we use the IAT to measure attitudes towards nonsocial and social target categories that vary in their familiarity and/or frequency of occurrence. Except for the word and name lists, both experiments closely follow Greenwald et al.'s methodology in order to provide a stringent test of the IAT's validity, and to make direct comparisons with their findings. Experiment 1 compared attitudes to the non-social target concepts of insects and flowers when these concepts were comprised of either all low or all high familiarity words. Experiment 2 directly examined the implicit racism effect in Caucasian American females by comparing attitudes to the contrasting racial categories of African American versus Caucasian and Hispanic versus Caucasian. We equated the familiarity and frequency of the first names across the three racial categories (see the bottom half of Table 2). Inclusion of Hispanic names allowed investigation of the generality of the IAT's sensitivity to racial bias.

EXPERIMENT 1

Participants in Experiment 1 completed two target concept discriminations. In one phase, they discriminated between high familiarity insects (e.g., flea) and flowers (e.g., carnation). In the other phase, the discriminations consisted of low familiarity insects (e.g., bedbug) and flowers (e.g., nightshade). In both phases, the evaluative dimension consisted of pleasant and unpleasant words of moderate familiarity. We calculated an IAT effect for each phase by comparing reaction times (RTs) on the evaluatively compatible task (flower + pleasant versus insect + unpleasant) with RTs on the evaluatively incompatible task (flower + unpleasant versus insect + pleasant). Comparison of IAT effects for the low and high familiarity conditions allowed us to assess the role of familiarity. If the IAT effect results from a positive bias toward more familiar stimuli, the IAT effect should diminish or disappear when low familiarity insects and flowers serve as the target categories. Additionally, the design allowed examination of gender differences in the IAT, although we had no specific expectation of a difference.

METHOD

PARTICIPANTS AND DESIGN

Fifty-six (28 male and 28 female) primarily Caucasian undergraduate volunteers from Western Washington University received course credit for their participation and were randomly assigned to one of the experimental conditions. No participants were eliminated based on Greenwald et al.'s criterion of 75% accuracy on each block. The experiment's design allowed examination of the effect of four variables: Word Familiarity (low and high), Order-of-Word-Familiarity (whether low familiarity words were presented in the first phase or second), Order-of-Combined tasks (whether the compatible task was first or second) and Gender-of-Participant. Word Familiarity was the sole within-subjects factor.

MATERIALS

Sixty words consisting of 10 pleasant, 10 unpleasant, 20 insects and 20 flowers formed the target concepts. The insect category was equally divided into high (e.g., aphid, flea, roach) and low (e.g., hornet, maggot, bedbug) familiarity words. Similarly, we divided the flower category into high (e.g., carnation, daisy, rose) and low (e.g., forsythia, lupin, nightshade) familiarity words. Appendix A describes the method for obtaining the ratings of familiarity and pleasantness. Appendix B contains the complete word lists.

The mean familiarity ratings (with larger values indicating more familiarity) for high and low familiarity insects were 4.88 (SD = .27) and 2.35 (SD = .74), respectively. These two insect categories did not differ in their average ratings of pleasantness (larger values indicate *more* pleasantness): low familiarity insects had a M = 2.58 (SD = .55) and high familiarity insects had a M = 2.56 (SD = .55). The mean familiarity ratings for high and low familiarity flowers were 4.83 (*SD* = .20) and 2.21 (*SD* = .59), respectively. Although generally comparable, there was a trend for high familiarity flowers to be rated as more pleasant than low familiarity flowers (*M* = 4.16, *SD* = .21, and *M* = 3.88, *SD* = .22). As can be seen, all flower categories were rated as more pleasant than all insect categories. We also attempted to match word length across the target categories.

As expected, pleasant words were more pleasant than the unpleasant words (*M* = 4.50, *SD* = .29, and *M* = 1.62, *SD* = .17, respectively). Additionally, pleasant and unpleasant words had comparable average ratings of familiarity (*M* = 2.58, *SD* = .82, and *M* = 2.90, *SD* = .48). Eighty percent of the pleasant and unpleasant words and 35% of the flower and insect words were identical to Greenwald et al.'s words.

IAT DESCRIPTION

The IAT method involved nine consecutive tasks. Each participant completed two IAT phases with each phase consisting of flowers and insects that were either all high or all low in familiarity. The first IAT phase involved five consecutive tasks: (a) initial target-concept discrimination, (b) evaluative dimension discrimination, (c) first combined task, (d) reversed target-concept discrimination and (e) a second combined task. The IAT method uses the same evaluative dimension (e.g., pleasant-unpleasant) in both phases, eliminating the need to repeat the evaluative discrimination task in the second phase. Consequently, the second phase consisted of only four tasks: (f) initial target-concept discrimination, (g) first combined task, (h) reversed concept-target discrimination and (i) second combined task. The IAT measures are obtained for the first phase by comparing RTs in tasks (c) and (e) and for the second phase by comparing RTs in (g) and (i).

PROCEDURE

All participants were tested individually on 200 MHz Pentium PC-compatibles running Inquisit Experiment Software, developed by Sean C. Draine. Upon arrival, we informed participants about the nature of the categorization tasks and encouraged them to respond as quickly as possible without sacrificing accuracy. Participants used the left and right forefingers to respond with the "A" and "5" (on numeric keypad) keys. Response keys were randomly assigned and counterbalanced across subjects and tasks. Response keys for the attribute dimension remained constant within each IAT phase.

Each task commenced with instructions about the category discrimination(s) and assignment of response keys (left or right) to categories. Throughout all trials, labels corresponding to the category assignment of response keys reminded participants of the relationship between categories and response keys. Target stimuli consisted of black upper-case letters presented in the center of a light gray background. A 250 ms intertrial delay occurred after each correct response and before the start of the next trial. When participants made an incorrect response, the word "ERROR" appeared in red letters for 300 ms followed by the intertrial delay. In order to encourage accuracy and rapid responding, each task concluded with a summary of the participant's mean percent correct and latency to respond.

Response blocks consisted of 40 trials and order of words was randomly determined for each participant in each block. For the single category discrimination tasks (Tasks a, b, d, f and h), participants completed two blocks for a total of 80 trials. All words occurred once before reappearing in these blocks. The combined tasks (c, e, g and i) consisted of three blocks (120 trials), rather than two, and the 10 stimuli from each category (insects, flowers, pleasant, unpleasant) occurred only once.

Due to an oversight, there were three methodological differences between these experiments and Greenwald et al.'s Experiment 1. First, on the combined tasks, words for the target-concept discrimination and the attribute discrimination did not appear on alternating trials. Instead, they were randomly ordered. Second, participants received feedback on their accuracy and their average response latency only at the completion of each task rather than at the end of each block. Finally, we did not collect explicit attitude measures in Experiment 1, as the primary purpose was to examine whether word familiarity moderates the size of IAT effects.

RESULTS

Overview of Data Analyses. Response times (RTs) and proportion of errors were calculated for each trial block. Examination of the distributions of response latencies indicated the presence of small proportions of extremely fast and slow responses. Because of the limited theoretical interest of these outliers and their potential to distort the means and inflate variances, we recoded RTs less than 300 ms to 300 ms, and RTs greater than 3000 ms were recoded to 3000 ms. Moreover, all ANOVAs and effect sizes were based on log transformed RTs to obtain a more acceptable stability of variance. Following Greenwald et al., we omitted the practice trials (the first 40 trials for each task) in the RT analyses and figures. However, in order to examine the response strategies of participants, the error analyses included these trials. These data reduction conditions were similar to Greenwald et al. (1998).

In contrast to Greenwald et al.'s analysis of RTs from all trials, including trials in which participants responded incorrectly, we examined RTs for only error-free trials because of the significantly greater error rates for low-frequency words (see the discussion of errors below). Effect sizes (d = mean/SD) were based on *unpooled* standard deviations because of the presence of interactions among factors. The use of various pooled standard deviations increased smaller effect sizes and reduced larger ones. Importantly, we obtained the same pattern of findings and statistical conclusions across both experiments, regardless of whether we (a) included or excluded RTs from error trials, (b) analyzed transformed or untransformed RTs, and (c) used pooled or unpooled standard deviations as the effect size unit. Finally, we employed the multivariate approach to within-subjects designs and an alpha level of .05 for all analyses.

The IAT Effect Measures. Figures 1A and 1B display the nine consecutive tasks completed in Experiment 1 by the factor of Order-of-Word-Familiarity (either low or high familiarity targets first). Within each figure, the levels of the factor of Order-of-Compatible conditions (flower + pleasant first or flower + unpleasant first) are displayed in the upper and lower panels and the Gender-of-Participant is displayed in the columns. White bars represent evaluatively compatible conditions (flowers + pleasant) and black bars represent incompatible conditions (flower + unpleasant). The difference in latency between the incompatible and compatible conditions results in the IAT effect measure. Back-transformed IAT effects were obtained by log transforming RTs, computing block means and then back transforming these means to milliseconds. It is important to highlight that we based the IAT disruption effects displayed in each of the figures on untransformed RTs, while all measures of effect size are based upon the back-transformed RTs. Table 3 contains the mean

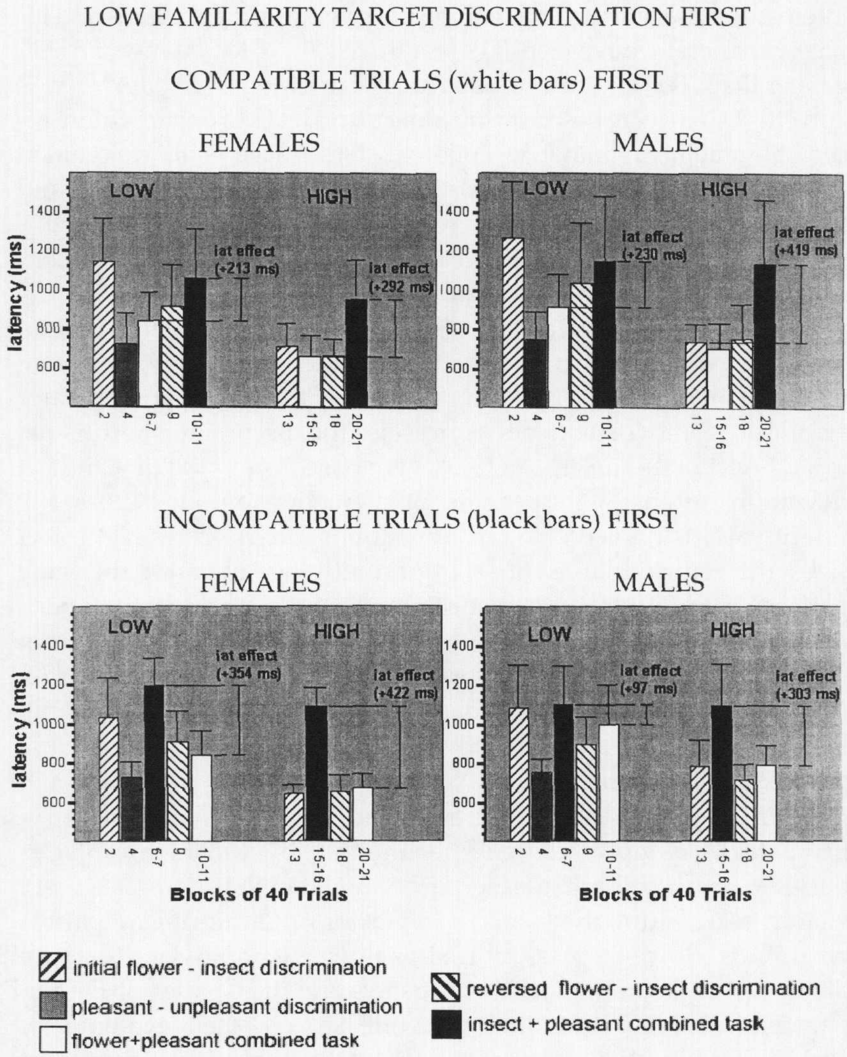

FIGURE 1A. Mean latency results of Experiment 1 for conditions in which low familiarity words occurred first. Results are presented by Gender-of-Participant (columns) and Order-of-Compatible-Combined Tasks (upper and lower panel).

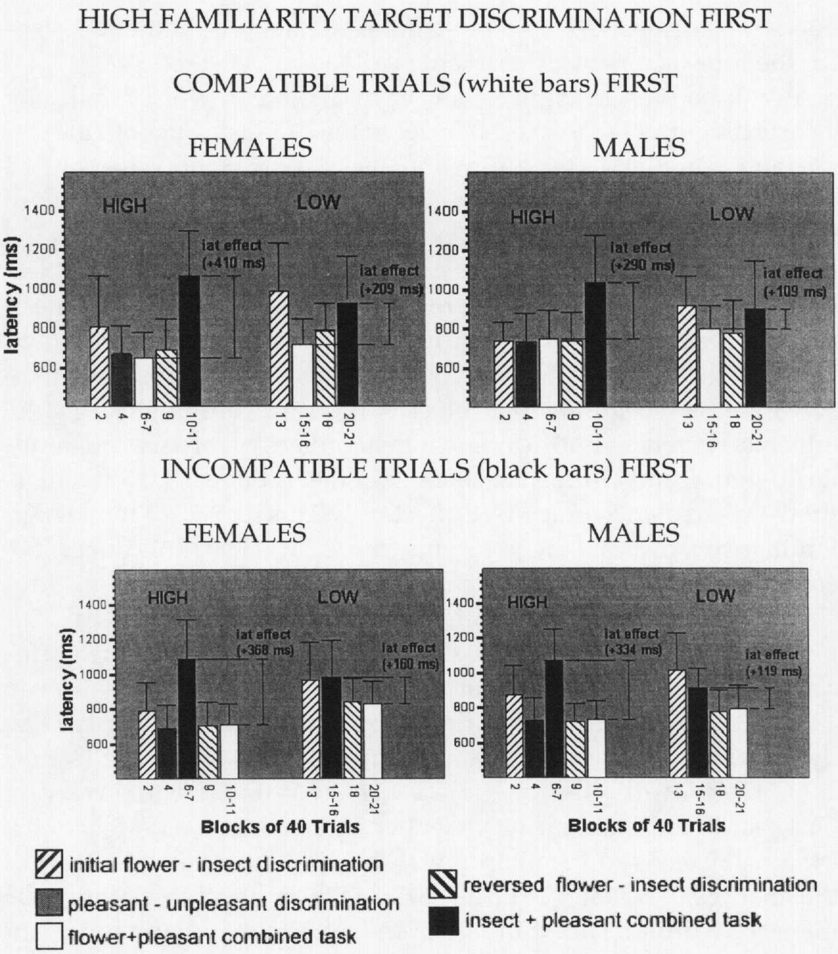

FIGURE 1B. Mean latency results of Experiment 1 for conditions in which high familiarity words occurred first. Results are presented by Gender-of-Participant (columns) and Order-of-Compatible-Combined Tasks (upper and lower panel).

Note. N = 56. The notable feature of this figure is that large IAT effects occurred to both low and high familiarity target concepts. All IAT effects are untransformed latencies. Error bars are standard deviations based on the 14 participants contributing to each mean. The first block of trials in each condition served as training trials and are excluded from the figure. Response latencies are untransformed. The upper panels contain conditions in which low familiarity words were presented first and the lower panels display conditions in which high familiarity words occurred first.

back-transformed IAT effects, standard deviations, and effect sizes for the four factors of Experiment 1.

As can be seen in Figure 1 (A, B) and Table 3, word familiarity clearly moderates response latencies in the IAT task. For both the single and combined discrimination tasks, participants responded more slowly to low familiarity target exemplars in comparison to high familiarity exemplars. These slower responses indicate that the categorization of low familiarity exemplars was more difficult. Second, large IAT effects occurred regardless of word familiarity (mean d = 1.75). As displayed in Figure 1 (A, B), participants performed faster on the flower + pleasant combinations (white bars) than on insect + pleasant combinations (black bars). Third, although large IAT effects occurred for both low and high familiarity words, word familiarity significantly influenced the size of the IAT effect. Low familiarity words had an average IAT effect of 160 ms (d = 1.26) while high familiarity words averaged 296 ms (d = 2.24). These IAT effects are comparable to Greenwald et al.'s average IAT effect of 176 ms (d = 1.54). Therefore, consistent with the findings of Greenwald et al., Experiment 1 indicated a more positive attitude toward flowers than insects, regardless of the familiarity of the flowers and insects.

A 2 x 2 x 2 x 2 (Word-Familiarity x Order-of-Word-Familiarity x Order-of-Compatible-Tasks x Gender) mixed ANOVA on the IAT effect indicated the presence of a Word-Familiarity x Order-of-Word-Familiarity x Gender interaction, F (1, 48) = 4.03, MSE = .0056. Word Familiarity was the only other reliable effect in the analysis, F (1, 48) = 161.81, MSE = .0056, indicating that high familiarity words resulted in larger IAT effects than low familiarity words (296 vs. 160 ms, respectively). This finding indicates that word familiarity moderates IAT effect. All other effects had a $p > .20$. In order to explain the 3-way interaction and examine both order and gender effects, we examined the effects of Gender and Order-of-Word-Familiarity and their interaction at each level of word familiarity (low and high) (Maxwell & Delaney, 1990, pp. 330).

Gender Effects. For high familiarity insects and flowers, a 2 x 2 (Order-of-Word-Familiarity x Gender) between-subjects ANOVA indicated no effect of order, $F < 1$, and no interaction, F (1, 48) = 2.34, MSE = .0056, $p > .10$. Only the effect of gender on the IAT measure was reliable, F(1, 48) = 11.88, MSE = .0056. Although both had large IAT effects (see top of Table 3), males (282 ms) had a slightly smaller IAT

TABLE 3. Mean IAT Latencies (ms) and Effect sizes (d) for Experiment 1 displayed by the factors of Word Familiarity, Order of Word Familiarity, Order of Combined Tasks, and Gender of Participant

High Familiarity Words

		Gender of Participant					
Order of	Order of	Females			Males		
Word Familiarity	Combined Tasks	*M*	*SD*	d	*M*	*SD*	d
Low Familiarity First	Compatible First	240	99	2.41	360	222	1.62
	Incompatible First	349	106	3.30	258	112	2.29
High Familiarity First	Compatible First	348	137	2.54	236	208	1.14
	Incompatible First	308	142	2.17	272	112	2.43
	Average	311		2.60	282		1.87

Low Familiarity Words

		Gender of Participant					
Order of	Order of	Females			Males		
Word Familiarity	Combined Tasks	*M*	*SD*	d	*M*	*SD*	d
Low Familiarity First	Compatible First	179	140	1.28	216	238	.91
	Incompatible First	309	137	2.26	74	106	.70
High Familiarity First	Compatible First	180	102	1.76	89	150	.59
	Incompatible First	133	91	1.46	99	87	1.14
	Average	200		1.69	120		.84

Summary of IAT effects

		M	d
Word Familiarity	High	296	2.24
	Low	160	1.26
Gender	Females	255	2.15
	Males	201	1.35
Order of Combined Tasks	Compatible First	231	1.53
	Compatible Second	225	1.97

Note. This table indicates that large IAT effects were obtained for all factors examined in Experiment 1. Mean IATs are back-transformed log RTs. The effect size measure d = [(re-transformed mean log RT)/SD]. The SD is not pooled across any other factors and is therefore based on n = 7. Small, medium and large effect sizes correspond to d values of .2, .5 and .8, respectively. All effect sizes are statistically significant at $p < .001$ with 55 d.f. The SEM for these analyses was pooled across all factors except Word Familiarity (Low or High).

effect than females (311 ms). We investigated this gender difference by examining the average speed of responding on the compatible and incompatible combined tasks (white and black bars in Figure 1, respectively). Males (747 ms) performed more slowly than females

(675 ms) on the compatible combined tasks, F (1,48) = 7.32, MSE = .002. However, no gender difference was present on the incompatible combined tasks (1083 and 1048 ms for males and females, respectively). Based on the assumption that faster responding indicates a stronger association, Experiment 1 suggests that males have less favorable attitudes toward flowers and/or more favorable attitudes toward insects than females.

For low familiarity words, a second 2 x 2 (Order-of-Word-Familiarity x Gender) ANOVA of the IAT effect revealed no interaction, $F < 1$. Males once again had a smaller IAT effect (120 ms) than females (200 ms), F (1, 48) = 8.65, MSE = .0056. Inspection of response latencies revealed the same pattern that was present with high familiarity words: Males performed more slowly than females on the compatible combined tasks (880 and 806 ms, respectively), F (1,48) = 3.88, MSE = .002, but they had comparable response latencies on the incompatible combined tasks (1016 and 1040 ms, respectively).

Order Effects. The three-way interaction (Word-Familiarity x Order-of-Word-Familiarity x Gender) is explained by the influence of Order-of-Word-Familiarity on the size of the IAT effect for low familiarity words, F (1, 48) = 25.30, MSE = .0056, but not high familiarity words. While the order of the familiarity of the target concepts (high familiarity discrimination tasks first or second) did not influence the IAT effects for high familiarity words, smaller IAT effects resulted when low familiarity words occurred in the second phase (125 ms, see Figure 1B) than when they occurred first in the first phase (194 ms). This finding most likely reflects a practice effect and suggests that IAT effects can be reduced with extended practice, especially when concept discriminations consist of low familiarity stimuli. It is important to emphasize, however, that the more favorable attitudes toward flowers did not disappear with practice.

Finally, there was a single difference between the current results and the findings of Greenwald et al. Experiment 1 revealed no effect of Order-of-Combined tasks (compatible condition first vs. second). Similar IAT effects occurred, regardless of the order in which compatible and noncompatible combined tasks (231 vs. 225 ms, respectively). In contrast, Greenwald et al. found larger IAT effects when compatible combined discriminations occurred first.

Error Analyses. One can also compute the difference in average error rates between the incompatible and compatible conditions as an index of automatic associations. We refer to this measure as the IAT error effect. Figure 2 contains the IAT error effects for Experiment 1. Figure 2 displays the average error rate for all blocks of Experiment 1. For reasons of economical presentation, we collapsed the results across gender. The average error rate for all blocks that involved low familiarity words was nearly twice (M = 10.8%) the error rate for blocks that contained high familiarity words (M = 5.6%). Thus, Experiment 1 also indicates that word familiarity affects IAT performance by increasing error rates.

Analysis of the difference in errors between compatible and incompatible blocks—the IAT error effect—revealed that errors were more frequent on incompatible trials. Almost all IAT error effects (M = 3.4%) differed from the expected value of zero. The only exceptions were when low familiarity words occurred in the second phase (right panels of Figure 2, M = 0.4% and M = –0.3%, $p > .20$). Average IAT error effect sizes for high and low familiarity words were d = 1.12 and d = 1.01, respectively. This similarity between high and low familiarity words provides further support for the conclusion from the RT analyses: Participants had more positive attitudes toward flowers than insects regardless of word familiarity.

A $2 \times 2 \times 2 \times 2$ mixed ANOVA on the IAT errors revealed a Word-Familiarity × Order-of-Word-Familiarity interaction, $F(1,48) = 13.82$, $MSE = .0015$. The IAT error effect for low familiarity words was smaller than the effect for high familiarity words only when low familiarity words occurred in the second phase (see right-hand column of Figure 2). This finding is also consistent with the RT finding that extended practice reduces the IAT effect for low familiarity words. An interaction between Word Familiarity and Gender on the IAT error effect approached statistical significance, $F(1,48) = 3.73$, $MSE = .0015$, $p = .059$. Simple effects tests indicated that females had a larger IAT error effect for blocks with high versus low familiarity words (5.1% vs. 2.1%, respectively, $F(1, 24) = 6.54$, $MSE = .0015$). In contrast, males had comparable error rates on high and low familiarity words, (3.4% and 3.1%, respectively). A second set of simple effects indicated that females had a greater IAT error effect than males for high familiarity words (5.1% vs. 3.4%, respectively, F (1,24) = 2.99, $MSE = .0015$). This finding is also consistent with the assumption that males have less favorable attitudes toward flowers and less unfavorable attitudes toward insects than fe-

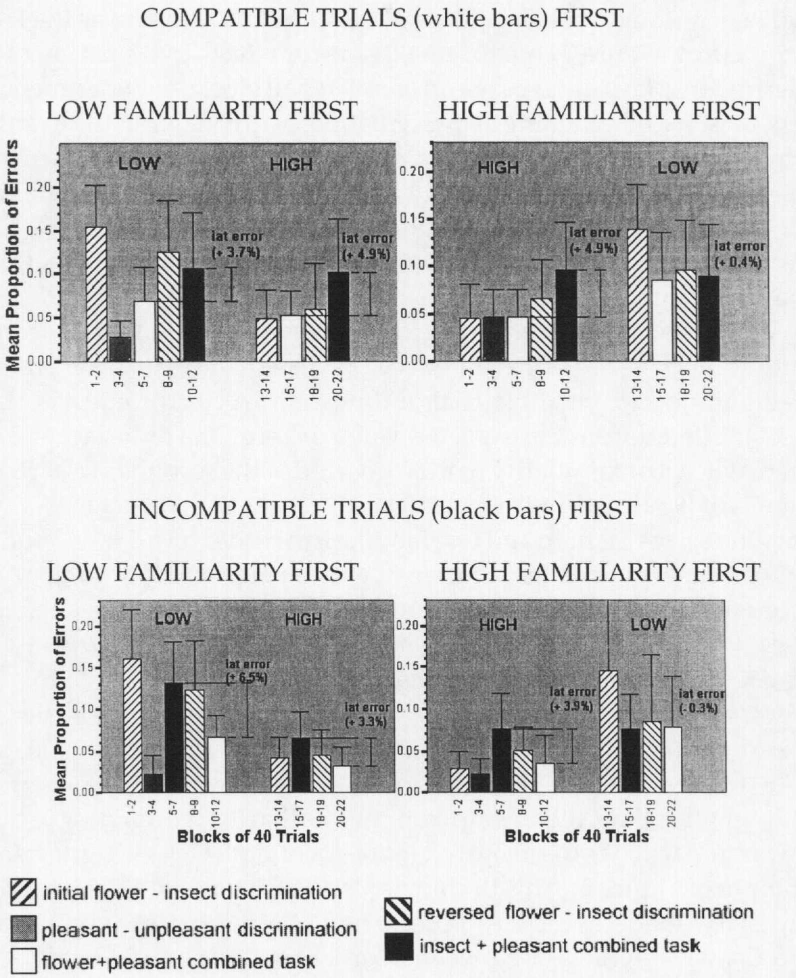

FIGURE 2. Mean proportion of errors in Experiment 1 presented by Order-of-Compatible-Combined Tasks (upper and lower panels), Order-of-Word-Familiarity (columns), and Word Familiarity.

Note. N = 56. The main finding in this figure is that the IAT effect can also be obtained when errors are examined. Error bars are standard deviations based on the 14 participants contributing to each mean. All trials (except the first two in each task) were used to calculate the proportion of errors.

males. A final important result is that we observed no reliable difference between females and males for the IAT error effect for low familiarity words (2.09% and 3.12%). The gender difference in error rates for high but not low familiarity words is important because it indicates that stimulus familiarity also influences the sensitivity of the IAT to group differences.

DISCUSSION

In addition to replicating the results of Greenwald et al. (1998), Experiment 1 clearly demonstrated that word familiarity moderates performance on the Implicit Association Test. First, across all blocks, participants responded more slowly and made more errors to low familiarity words. Second, IAT effects were smaller when low familiarity exemplars formed target concepts (mean d = 1.26), in comparison to when high familiarity stimuli formed the target concepts (mean d = 2.24). Additionally, word familiarity interacted with amount of practice: When low familiarity words occurred in the second phase, smaller IAT effects and more errors occurred. These results indicate that the use of low familiarity stimuli reduces, but does not eliminate, the sensitivity of the IAT. However, the use of low familiarity words potentially introduces error variance that may (a) minimize the magnitude of group differences and (b) affect the rank ordering of individuals. Thus, low familiarity stimuli may reduce the IATs value as an individual differences measure. Based on these results, we strongly recommend that experimenters who use the IAT and other indirect attitude measures attempt to control word familiarity when attempting to assess automatic attitudes.

Experiment 1 also demonstrated the IATs sensitivity to gender differences. Males had smaller IAT effects than females (mean ds of 1.35 and 2.15, respectively), although it is important to recognize that both genders had large IAT effects. Examination of response latencies indicated that males responded more slowly than females on evaluatively compatible tasks (e.g., flower + pleasant). In contrast, males and females had comparable response latencies on the evaluatively incompatible tasks (e.g., flower + unpleasant). This IAT finding suggests that males are less likely to associate flowers with pleasantness and insects with unpleasantness and is consistent with gender differences in evaluations of pleasantness reported by Bellezza et al. (1986). We found a similar trend in pilot ratings of the pleasantness/unpleasantness of insects and flowers (see Hayden & Ottaway, in preparation, for a further discussion). These IAT gender differences are also consistent with Fujita, Diener, and Sandvik (1991), who argue that women have greater positive and negative emotional intensity than men. However, the gender difference obtained with the IAT is ambiguous because it could also indicate that males are slower at combined discrimination tasks, and that a ceiling effect for RTs in the incompatible conditions could have resulted in the gender differences. We believe that convergent evidence from either indirect or explicit attitudes measures is necessary to confirm the sensitivity of the IAT to gender differences.

Lastly, comparable IAT effects occurred in Experiment 1 regardless of the order of the combined tasks (compatible condition first vs. second). In contrast, across three experiments, Greenwald et al. reported a trend for larger IAT effects when compatible conditions occurred first (although this effect was statistically significant only in their first experiment). This discrepancy may be due to the use of fewer trials in Experiment 1 (120 vs. Greenwald et al.'s 200 trials on combined tasks) and/or the exclusion of error trials in our analyses.

FAILURE TO SUPPORT PRIOR-EXPOSURE AND SPEED-ACCURACY TRADEOFF EXPLANATIONS OF THE IAT

Greenwald and associates have proposed that the IAT measures implicit attitudes. The findings of the current experiment are inconsistent with this familiarity/frequency explanation, which proposes that the IAT reflects differential positive bias to high familiarity words. If the IAT effect was solely due to a positive bias toward more familiar stimuli, then the inclusion of low familiarity insects and flowers should have minimized or eliminated the IAT effect. However, both low and high familiarity words resulted in large IAT effects ($d = 1.26$ and $d = 2.24$, respectively), comparable to those obtained by Greenwald et al. ($d = 1.54$). Although familiarity clearly exerts an important influence on the IAT, the use of low familiarity stimuli does not eliminate the sensitivity of the IAT. Therefore, experiment 1 supports the idea that the Implicit Association Test assesses participant's automatic attitudes toward target categories.

Experiment 1 also failed to support a speed-accuracy tradeoff explanation[2] of the IAT effect. It is possible that participants performed more slowly on incompatible blocks (e.g., flower + unpleasant) - and thus had an IAT effect - because they consciously attempted to minimize errors and therefore ensure accuracy at the cost of response speed. This argument presupposes that IAT effects result from an explicit "cautious responding" strategy, rather than an automatic attitude resulting from past experience. Our pattern of errors and response latencies are, however, inconsistent with this account's prediction of fewer errors and slower response latencies on incompatible blocks. Participants were more likely to make errors on incompatible blocks, suggesting that they did not adopt a different criterion of accuracy. Moreover, for low familiarity words, participants had comparable average-response latencies on the incompatible combined tasks and the two single discrimination tasks (1028 vs. 1053 ms, respectively). Participants clearly made an effort to respond rapidly during both types of discrimination tasks. Taken in conjunction, the pattern of response latencies and errors are inconsistent with a speed-accuracy tradeoff explanation of the IAT effect.

EXPERIMENT 2

Experiment 1 is remarkable for the large IAT effects resulting from both low and high familiarity words and the substantial overlap with Greenwald, McGhee, and Schwartz's (1998) findings. Experiment 2 investigated whether stimulus familiarity and frequency of occurrence also influence the implicit racism effect. Greenwald et al. argued that the IAT task is an important potential indirect measure of race-related attitudes which, due to their socially sensitive nature, may not be accurately assessed by explicit self-report measures (e.g., Fazio et al., 1995; Wittenbrink, Judd, & Park, 1997). The disturbing implication of Greenwald et al.'s results is that racism is still prevalent, even though this prejudice is not reflected by self-report measures. As detailed in the Introduction, analysis of the first names used by Greenwald et al. confirmed that either name familiarity or frequency of occurrence could be responsible for their implicit racism findings. In the present experiment, Caucasian

2. We thank Richard J. Harris for suggesting this possible interpretation.

American undergraduates categorized Caucasian, African American and Hispanic names matched for familiarity and frequency of occurrence. The primary objective of Experiment 2 was to examine the implicit racism effect when name familiarity and frequency were equated across racial categories. Because Experiment 1 demonstrated that large IAT effects occur even with low familiarity stimuli, we expected to replicate Greenwald et al.'s implicit racism effect with our moderate familiarity and frequency names.

A second objective of Experiment 2 involved the assessment of the IAT's construct validity. Greenwald et al.'s results indicated prejudice of Caucasians towards African Americans, as well as mutual prejudice between Japanese and Koreans. If the IAT assesses attitudes to socially sensitive domains, we also expected an implicit racism effect toward Hispanics. This prediction was based on the out-group status of Hispanics relative to our Caucasian participants as well as reports of similarities between African Americans and Hispanics in types of discrimination experienced (Ong, 1991) and levels of poverty and housing segregation (Massey & Eggers, 1990). A failure to obtain implicit racism towards Hispanics would seriously question the generality of the IAT method as a measure of racial prejudice. Finally, we included male and female names to examine the generality of the IAT effect. If the IAT is sensitive to anti-African American and anti-Hispanic attitudes, than we should expect to obtain large IAT effects for both male and female names. However, we expected larger IAT effects for male names because males names always occurred in the first IAT phase while female names always occurred in the second phase.

METHOD

PARTICIPANTS AND DESIGN

Thirty-three female undergraduates from the same subject pool of Experiment 1 received course credit for their voluntary participation. All participants self-described themselves as Caucasian-American. Upon meeting the participants, we informed them that the study could potentially reveal attitudes toward another race that they may not prefer to reveal. The experimenter then assured participants that their responses would remain anonymous and reminded them of their right to terminate their participation at any time without loss of course credit. No participants exercised this option. Furthermore, no participants were eliminated because of excessive errors. However, one subject was replaced because their data were accidentally erased from the computer.

The design consisted of two between-subject factors: Race-of-Names (African American vs. Caucasian or Hispanic vs. Caucasian) and Order-of-Combined tasks (Non-Caucasian + pleasant combinations in the first phase or second). The sole within-subjects factor was gender-of-name (male names in Phase 1, female names in Phase 2). Because African American males are more likely to be the victim of discrimination than African American females (Klonoff & Landrine, 1999), we always presented male names first to maximize the opportunity of obtaining correlations between the IAT and explicit racial measures. Experiment 2 was identical to Experiment 1, except that (a) target categories consisted of first names commonly given to members of different racial groups, (b) we included explicit measures of attitudes toward racial groups, discrimination and mulitculturalism, (c) we used only names of moderate familiarity, and (d) all participants were female.

MATERIALS AND PROCEDURE

The critical manipulation of Experiment 2 was the selection of first names readily identified as being common names of African-Americans, Hispanic-Americans, or Caucasian-Americans. We chose the 60 first names (30 female, 30 male) based on four criteria: First, names had to be commonly given to African-Americans, Caucasian-Americans, or Hispanic-Americans. Second, names had to have an average familiarity rating of approximately 3 ("somewhat familiar") on a 5 point scale. This use of moderate familiarity names was necessary because few African American and Hispanic first names are highly familiar to our Caucasian American undergraduates. Similarly, it was difficult to find any low familiarity Caucasian names because most Caucasian names are highly familiar to the Caucasian Americans who completed the ratings. Third, names from each racial group were required to have similar average cumulative frequencies in a U.S. Census database. Thus, name frequency was also equated across the three racial categories. Lastly, from the names that satisfied the first three criteria we selected names used by Greenwald et al. (there was a 65% and 30% overlap in African American and Caucasian names, respectively). Appendix A describes in more detail the name selection sources and name ratings.

Table 2 displays the mean ratings of familiarity and mean cumulative frequencies for the names used in Experiment 2. In comparison to Experiment 1s use of low and high familiarity stimuli, the names used in Experiment 2 were of moderate familiarity. Similarly, a comparison to the names used by Greenwald et al. indicated that our Caucasian names were less familiar and our nonCaucasian names were more familiar. The following are examples of names from each racial category and gender: African American males (Darnell, Terrence, Tyrone), African American females (Latisha, Rochelle, Tasha), Hispanic males (Luis, Pedro, Ricardo), Hispanic females (Juanita, Monica, Rosa), Caucasian males (Alfred, Barry, Marty) and Caucasian females (Dorothy, Kandace, Peggy). Appendix B contains the entire set of 60 names. The pleasant and unpleasant words were identical to those used in Experiment 1.

PROCEDURE

IAT Measures. Each participant completed one of two target discriminations of nonCaucasian versus Caucasian names - either African American versus Caucasian or Hispanic versus Caucasian. For each non Caucasian + Caucasian combination of names, participants completed two IAT phases: one with male names followed by one with female names. The presentation of female names in the second phase follows the method used by Greenwald et al. All other procedural, random assignment, counterbalancing, and timing details were identical to Experiment 1.

Racial Attitude and Demographic Questionnaires. After the computer tasks, participants completed four questionnaires that assessed their attitudes to Caucasians and either African Americans or Hispanics. These questionnaires were identical to those used by Greenwald et al. except for modifications that allowed us to assess attitudes toward Hispanics as well as African Americans. Written instructions reminded participants that the experiment could disclose attitudes that they might not wish to reveal, emphasized that the experimenter who assisted them would never see their data, that their names would never be associated with their responses, and that they would personally place their data in a sealed envelope before leaving. Participants then opened an unsealed envelope that contained the questionnaires.

The Feeling Thermometer questionnaire consisted of an illustration of a thermometer numerically labeled at increments of 10 ranging from 0 to 99. The thermometer was anchored at three points (0, 50, & 99) with the labels "cold or unfavorable," "neutral," and "warm or favorable." Participants marked the appropriate position corresponding to their general feeling of "warmth or coolness" toward the target racial group. Participants completed two Feeling Thermometers (one for Caucasians and one for non-Caucasians). The order of the ratings (Caucasians first or second) was counterbalanced across subjects. Scoring of the Feeling Thermometer consisted of calculating the difference between non-Caucasian and Caucasian ratings (e.g., African American-Caucasian). Possible scores ranged from -99 to +99 with higher scores reflecting pro non Caucasian attitudes.

Next, we administered the Semantic Differential Questionnaire which uses polar-opposite adjectives to assess attitudes toward a target category. This questionnaire contains five 7-point scales, consisting of adjective pairs (beautiful-ugly, good-bad, pleasant-unpleasant, honest-dishonest, and nice-awful). Participants indicated their attitude towards the target racial categories (non Caucasians and Caucasian) by circling the number corresponding to their choice.[3] Scoring consisted of averaging the five responses resulting in a score that could range from -3 to +3 with higher scores reflecting more positive attitudes. A difference score (non Caucasian-Caucasian) was then calculated for each participant. This score ranged from -6 to +6 with higher scores reflecting more positive attitudes to the non Caucasian racial group.

Finally, participants completed both the Discrimination and Diversity Scales (Wittenbrink et al., 1997). The Discrimination Scale contains 11 questions that assess beliefs about the causes and prevalence of discrimination in the United States. A sample statement

3. Due to an oversight and contrary to Greenwald et al., we did not provide participants specific instructions about how to respond (i.e., indicate zero) on the Semantic Differential Task if they felt that the adjectives were irrelevant to the target category. It appears that this difference did not have a major impact due to the highly similar results between Greenwald et al. and the Semantic Differential Task of Experiment 2.

from this scale is "Discrimination against Blacks is no longer a problem in the United States." On a 5-point scale with the anchors of *strongly disagree* and *strongly agree*, participants indicated to what extent they agreed with each statement. Participants circled "0" if they neither agreed or disagreed with the statement. The Diversity Scale assesses beliefs about multiculturalism and uses the same 5-point response scale. A sample from the four items is: "Whites will need to learn about *Black* culture if positive interethnic relations are to be achieved." Higher scores on these scales usually indicate greater endorsement of antidiscrimination and multiculturalism. Following Greenwald et al., we reversed the scoring on these scales so that higher scores reflect less endorsement of racist beliefs.

As illustrated in the examples, Wittenbrink et al.'s measures originally used African Americans as a racial referent. To assess beliefs about discrimination and diversity as they relate to Hispanic groups, references to African Americans were replaced with references to Hispanics (e.g. "Discrimination against *Hispanics* is no longer a problem in the United States"). Because of this modification, we assessed the internal consistency of both scales. For the present sample, the internal consistency for the Discrimination Scale was high and comparable for the African American and Hispanic versions (Cronbach's α = .84 and .86, respectively). However, the internal consistency of the Diversity Scale was much lower for the original version than the Hispanic version (α = .46 and .92, respectively). For the final requirement of the experiment participants completed a demographic questionnaire in which they indicated their gender and ethnicity.

RESULTS AND DISCUSSION

IAT Analyses. Figure 3 presents the nine IAT tasks completed in Experiment 2 displayed by Race-of-Name (African American + Caucasian, Hispanic + Caucasian) and Order-of-Compatible Combined tasks.[4] Consistent with the findings of Greenwald et al., Caucasian+pleasant combinations (white bars) produced faster

4. This labeling presupposes that the Caucasian + pleasant combination is evaluatively compatible relative to the African American + pleasant and Hispanic + pleasant combinations, at least for the Caucasian-American participants of Experiment 2 and is consistent with that used by Greenwald et al.

responding than the non Caucasian + pleasant combinations (black bars). The IAT effect (based on retransformed RTs) averaged 108 ms (d = 1.00) for the African American + Caucasian combinations and 94 ms (d = .82) for the Hispanic+Caucasian combinations. All IAT effects in Figure 3 were reliable, except for one IAT effect of only 10 ms. Thus, even when name familiarity and frequency were equated across the non Caucasian and Caucasian racial categories, large implicit racism effects occurred to African American and Hispanic target concepts.

A 2×2×2 (Gender-of-Name × Race-of-Name × Order-of-Compatibility-Conditions) mixed ANOVA was conducted on the log transformed IAT effects. Important, there was no effect of Race-of-Name, $F < 1$, indicating that the Hispanic + Caucasian and African American + Caucasian combinations resulted in similar IAT effects. Only the factors of Gender-of-Name, $F(1,28) = 4.70$, $MSE = .011$, and Order-of-Compatibility conditions, $F(1,28) = 5.70$, $MSE = .022$, were reliable. Male names resulted in a larger IAT effect ($M = 128$ ms, d = 1.14) than female names ($M = 75$, d = .68). This result may reflect either a greater negative attitude toward non Caucasian males over non Caucasian females, or it may reflect an order effect because evaluation of male names always occurred before evaluation of female names. In support of the former explanation, Klonoff and Landrine (1999) reported that male African-Americans were more likely to be the targets of racism than female African-Americans. Future research will need to address whether the IAT is sensitive to racial attitudes that depend on the gender of the target.

Finally, non Caucasian-of-Compatibility conditions indicated that larger IAT effects occurred when participants completed the non Caucasian + pleasant combined tasks first in comparison to completing it second. It is worth noting that this result is also inconsistent with Greenwald et al.'s finding of larger IAT effects when incompatible tasks precede compatible tasks.

Error Analyses. The IAT error effect—the difference in error rates for compatible versus incompatible blocks—averaged only 1.2% and was not reliable for any factors. The small number of errors obtained with the moderate familiarity stimuli used in this experiment is consistent with finding from Experiment 1 of a smaller number of errors for high relative to low familiarity stimuli. As noted by a reviewer, the error rates for the name discrimination tasks (~13%) indicated

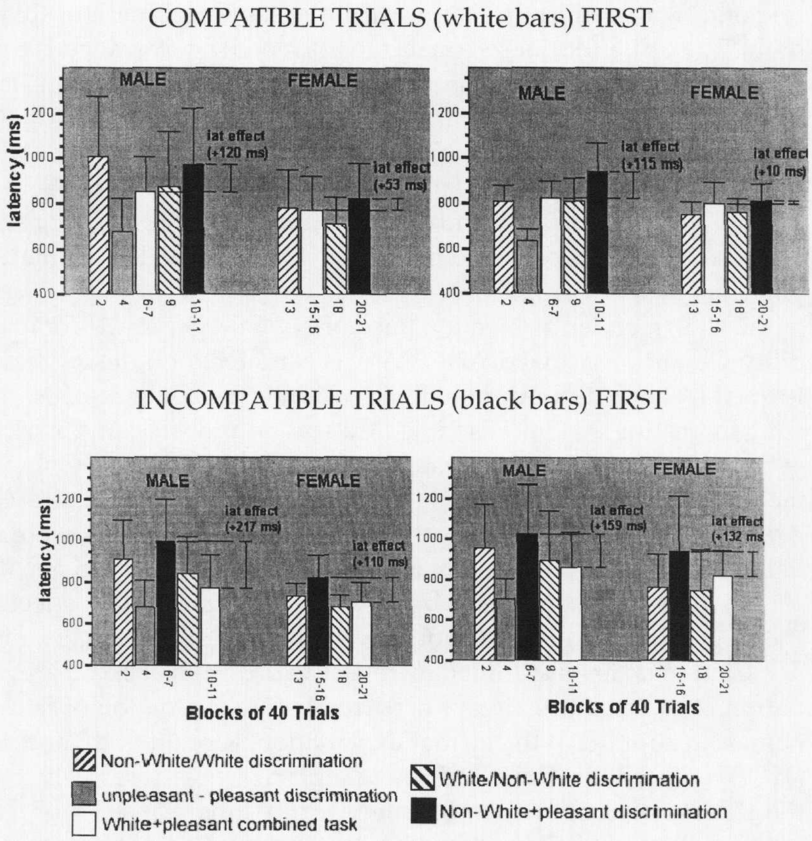

FIGURE 3. Mean response latencies of Experiment 2 presented by Gender-of-First-Names (columns) and Order-of-Compatible-Combined Tasks (upper and lower panel).
Note. N = 32. IAT effects are based on untransformed RTs. Error bars are standard deviations based on the 8 participants contributing to each mean. The first block of training trials in each condition are excluded from the figure. Response latencies are untransformed. Male names were always presented in the first phase followed by female names in the second phase.

that participants may have had difficulty with the initial name discriminations between Caucasian and non Caucasian names. For instance, names such as Monica or Sylvia may have been mistakenly categorized as Caucasian rather than Hispanic females names. Importantly, we excluded the first 40 trials of each task and all error trials from the reaction time analyses in order to minimize the effect of individual differences in ease of identifying the racial identity of names. Furthermore, we found an identical pattern of findings when the error trials were included and excluded from the analyses. Given that (a) the African American names are less susceptible to this criticism, (b) the theoretically predicted IAT effects were obtained for both African American and Hispanic names and (c) our results replicated Greenwald et al.'s findings, we believe that any difficulty of classifying names did not significantly influence our implicit racism findings.

Of additional importance for understanding IAT performance is the finding that errors on incompatible trials were not less than errors on compatible trials. This finding contradicts a "cautious responding" interpretation of IAT effects which predicts both slower RTs and fewer errors on incompatible trials. This prediction was not supported by either experiment. Consequently, our experiments clearly demonstrate that the IAT effect is not a function of a speed-accuracy tradeoff and thus are inconsistent with an explicit 'cautious responding' interpretation of the IAT effect.

RELATIONSHIP BETWEEN IMPLICIT AND EXPLICIT MEASURES

Experiment 2 also allowed us to address whether Greenwald et al.'s finding of only weak correlations between the IAT and explicit attitude measures resulted from their confounding of name familiarity and frequency. Table 4 presents the summary statistics for three sets of measures: explicit attitudes (Feeling Thermometer and Semantic Differential), racist beliefs (Diversity and Discrimination Scales) and implicit attitudes (IAT effects). The upper and lower tables separately present the results for participants that completed African American versus Caucasian and Hispanic versus Caucasian discrimination tasks.

As can be seen in Table 4, effect sizes for the IAT, the Feeling Thermometer, and the Discrimination Scale revealed pro-Caucasian attitudes relative to non Caucasians. In contrast, the Semantic Differential questionnaires and the Diversity Scale indicated equal preference for non Caucasians and Caucasians. These result replicates Greenwald et al., and is striking given that our first experiment demonstrated that low familiarity stimuli reduce the size of IAT ef-

TABLE 4. Summary Statistics for Experiment 2's Explicit and Implicit Attitude Indexes

	African Americans and Caucasians				
Index[a]	*M*	*SD*	d[b]	t(15)	*p*
Feeling Thermometer	–9.88	11.44	–.86	–3.45	< .005
Semantic Differential	–.087	.44	–.19	–0.80	> .10
Discrimination	–.659	.64	–1.03	–1.03	< .002
Diversity	–.156	.70	–.22	–.89	> .10
IAT (male names)	–140	122	–1.15	–4.60	< .001
IAT (female names)	–77	91	–.85	–3.38	< .005
	Hispanics and Caucasians				
Index	*M*	*SD*	d[b]	t(15)	*p*
Feeling Thermometer	–14.40	16.68	–.86	–3.34	< .005
Semantic Differential	–.333	.75	–.44	–1.72	> .10
Discrimination	–.642	.67	–.96	–3.74	< .003
Diversity	–.183	.43	–.43	–1.66	> .10
IAT (male names)	–116	104	–1.12	–4.46	< .002
IAT (female names)	–73	113	–.51	–2.58	< .04

Note. N = 32. Participants were Caucasian-American females.
[a]On the Feeling Thermometer, Semantic Differential and IAT measures, negative scores indicate a preference for Caucasians relative to African Americans/Hispanics. Scores of 0 indicate equal preference for non Caucasian and Caucasian. On the Discrimination and Diversity scales, positive scores indicate endorsement of anti-African American/anti-Hispanic beliefs. The potential range of the Feeling Thermometer was -99 to +99 and the potential range of the Semantic Differential was -6 to +6. The IAT measures are latencies that have been transformed to natural logarithms and then back-transformed into milliseconds. The Discrimination and Diversity scores range from -2 to +2 and are reversed relative to their usual scoring.
[b]The effect size measure d= [(mean)/SD]. The SD is not pooled across any other factors and is therefore based on n = 16. Small, medium and large effect sizes correspond to d values of .2, .5, and .8, respectively.

fects. An important finding contained in Table 4 is that the IAT effects in Phase 1 (male names) were larger (mean d = –1.14) than the largest explicit attitude measure (mean d = –.86 for the Feeling Thermometer) and the Discrimination Scale (mean d = –1.03). (As described in Table 4, all measures have been coded so that negative values indicate a pro-Caucasian preference. As an example, a negative effect size indicates a pro-Caucasian orientation while a positive effect size indicates a pro-African American orientation on the IAT.).

Because there were no reliable differences in IAT effects to African Americans and Hispanics, we collapsed across this factor for all implicit and explicit attitude measures to increase statistical power. Table 5 contains the correlations among these measures. Correlations between the IAT measures and the measures of racism and explicit

TABLE 5. Correlations among Experiment 2's Indirect and Explicit Measures of Racial Attitudes and Racist Beliefs

	Explicit Attitudes		Racist Beliefs		Implicit Attitude
Indexes	1	2	3	4	5
1. Feeling Thermometer					
2. Semantic Differential	.67				
3. Diversity	.45	.36			
4. Discrimination	.55	.58	.60		
5. IAT (male names)	.21	.21	–.03	.29	
6. IAT (female names)	.14	.15	.06	.20	.42

Note. Correlations are collapsed across the non-Caucasian name categories (African American, Hispanic). IAT effects are log transformed for these analyses. N = 32 for all correlations; correlations that are statistically significant at the 2-tailed p-value of .05 are printed in bold, rcrit,.05(31) = .35, rcrit,.01(31) = .40.

attitudes were small and not statistically reliable. The average of these correlations was only r(31) = .15. In contrast, correlations between the explicit attitude and racist belief measures were reliable and moderately large, average r(31) = .53. These findings parallel Greenwald et al.'s results, and therefore demonstrate that their dissociation between the IAT and explicit measures was not due to a confounding of racial category of names with name familiarity and frequency. Moreover, the finding of only a weak relationship between the IAT and four explicit measures is consistent with other experiments demonstrating the independence of indirect and self-report measures of race-based prejudices (Devine, 1989; Dovidio et al., 1997; Fazio et al., 1995; von Hippel, Sekaquaptewa & Vargas et al., 1997).

GENERAL DISCUSSION

We have organized the General Discussion into four sections. In the first section, we review results that establish the validity of the IAT as an indirect measure of automatic attitudes. Next, we discuss the implications of these results for alternative explanations of performance on the IAT. The penultimate section discusses current results in relation to theoretical accounts of implicit social cognition, and the final section addresses the discriminant and external validity of the IAT.

CONSTRUCT VALIDITY OF THE IMPLICIT ASSOCIATION TEST

This article produced six important results that establish the construct validity of the IAT as a measure of implicit attitudes. First, we replicated and extended Greenwald et al.'s finding that the IAT is sensitive to theoretically predicted attitudes toward both nonsocial and social attitudes. Second, and more important, large IAT effects were observed when low, moderate, and high familiarity words formed the target categories. However, stimulus familiarity plays a moderating role in the IAT: The IAT is more sensitive to evaluative associations when high familiarity words form the target concepts. Third, we obtained implicit racism toward two racial out-groups when name familiarity and frequency were equated. Fourth, Experiment 2 extended the implicit racism effect to Hispanics. Fifth, Experiment 1 demonstrated that the IAT is sensitive to gender differences. Finally, as discussed more fully below, we failed to find a relationship between the implicit and explicit measures of attitudes in Experiment 2. Taken in sum, both experiments strongly support Greenwald, McGhee, and Schwartz's (1998) conclusion that the IAT measures automatic evaluative attitudes.

FAMILIARITY, FREQUENCY, AND SELF-PRESENTATION EXPLANATIONS OF IMPLICIT RACISM

The current experiments provide considerable insight into the role of stimulus familiarity on the Implicit Association Task. Our experiments indicate that the combination of differential familiarity with stimuli and the ubiquitous finding of greater liking for more familiar stimuli can not account for the performance in the IAT task. Consistent with a prior-exposure explanation of the IAT effects, the use of more familiar target concepts resulted in larger IAT effects (mean $d = 2.24$). However, contrary to a prior-exposure account, low familiarity words also produced large IAT effects (mean $d = 1.26$). The consistent finding of large IAT effects in Experiment 1, regardless of word familiarity, demonstrates that the IAT does not simply reflect a positive bias to high familiarity stimuli.

Experiment 2 also failed to support a prior-exposure explanation of IAT effects. We found large implicit racism effects for African American-Caucasian evaluative contrasts (mean $d = 1.00$). and Hispanic-Caucasian evaluative contrasts (mean $d = .83$) when all names were of moderate frequency and familiarity. While the results fail to support a prior-exposure explanation of the IAT effect, stimulus familiarity can affect the size of IAT effects. Most obviously, confounding stimulus familiarity or frequency allows subjects to use this alternative dimension to complete the IAT's discrimination tasks. The presence of differential familiarity can therefore make the IAT tasks easier and thus minimize IAT effects. As already discussed, we recommend that researchers who employ the IAT and other indirect attitude measures (e.g., Devine, 1989; Dovidio et al., 1997; Fazio et al., 1995) use high familiarity stimuli in order to minimize task difficulty and error variance. Support for this recommendation is provided by the findings of Dasgupta, McGhee, Greenwald, and Banaji (2000) who examined frequency-discrepant names (low vs. high), as well as frequency-discrepant evaluative dimensions and compared them to matched names and evaluative dimensions in an implicit racism IAT. Contrary to a frequency explanation, larger IATs were obtained when names were matched in frequency than when infrequent black names and frequent Caucasian names were contrasted. Taken in conjunction, our findings and Dasgupta et al.'s results fail to support a prior exposure explanation of results from the IAT.

Our experiments also addressed the possibility that an explicit self-presentation strategy is responsible for IAT effects. Specifically, because participants knew we were interested in their racial attitudes, they may have attempted to avoid the impression of prejudice by adopting a "cautious responding" strategy on the incompatible trials (e.g., Hispanic + pleasant). By attempting to avoid incorrectly categorizing the non Caucasian first names in the unpleasant category, participants may have inadvertently responded more slowly in their attempt to be accurate. However, analysis of error rates was inconsistent with this interpretation. Participants had larger (Exp. 1) or similar (Exp. 2) error rates on the incompatible in comparison to the compatible combined tasks. Additional evidence inconsistent with explicit strategy interpretations of the IAT effect are Banaji and colleagues' (Banaji & Bhaskar, 2000) findings that participants are often unaware of their inability to perform as rapidly on the incompatible relative to the compatible conditions, and are unable to eliminate the IAT effect when instructed to do so. A cautious responding strategy can not explain implicit racism effects obtained with the IAT.

RELATIONSHIP BETWEEN THE IAT AND EXPLICIT MEASURES

Experiment 2 also allowed examination of the relationship between implicit and explicit measures of socially sensitive attitudes. We observed only weak correlations (mean $r = .15$) between the IAT effects and two direct measures of racial attitudes and two direct measures of racist beliefs. In contrast, correlations within direct and indirect measures were much higher (mean $r = .53$ for direct measures, mean $r = .42$ for indirect measures[5]). Moreover, we found the theoretically predicted patterns between the IAT and explicit measures, regardless of whether African Americans or Hispanics comprised the target categories. These results indicate that Greenwald et al.'s failure to find a relationship between the IAT and explicit measures can not be explained by their confounding of racial category with both name familiarity and frequency. Moreover, our replication of the weak relationship between the IAT and explicit measures of racial prejudice is consistent with similar dissociations obtained with a heterogeneous group of indirect and explicit prejudice measures (Banaji & Hardin, 1996; Dovidio et al., 1997; Fazio et al., 1995; von Hippel et al., 1997). In summary, Experiment 2 supports Greenwald et al.'s conclusion that, in addition to reflecting differences in the task requirements, indirect and explicit attitude measures assess different constructs: automatic evaluative associations versus conscious, controlled attitudes susceptible to self-presentation forces.

IMPLICATIONS FOR THEORIES OF IMPLICIT SOCIAL COGNITION

Implicit aspects of social cognition recently received considerable theoretical attention because of their potential to increase the predictive and construct validity of the attitude construct (Greenwald & Banaji, 1995). Similar to theoretical accounts of learning (Reber, 1989) and memory (Jacoby, 1991), recent theories of social cognition are based on the assumption that past experience moderates behavior through the elicitation of implicit and explicit cognitions. Explicit cognitions reflect the operation of conscious controlled activity,

5. The magnitude of this correlation was probably reduced by the practice effect present in Experiment 2. The correlation between IAT effects in Experiment 1 was much larger, $r(55) = .77$, $p < .001$, two-tailed.

while implicit cognitions reflect nonconscious automatic processes (cf., Greenwald & Banaji, 1995; Neasdale & Durkin, 1998 for overviews). Devine (1989) has proposed that both automatic and controlled components play a mediating role in determining racial prejudice (see Bargh et al., 1992 for a related approach). In particular, she argues that if one has acquired socially shared stereotypes about a group, then exposure to a group member results in inescapable, automatic elicitation of the cultural stereotype. However, Devine proposed that automatic attitudes can be overcome by controlled processing in which one makes a conscious decision to act in a nonprejudiced manner.

Fazio has proposed a related model (MODE; Fazio, 1990, Fazio et al., 1995) that focuses on individual differences in automatic and controlled processes. The MODE model differs in two ways from Devine's approach. First, the model assumes that automatic activation of a stereotype does not occur for all individuals. Automatic bias therefore depends on an individual's specific prior experiences rather than being universal across individuals. Second, MODE assumes that the controlled processing also differs among individuals and will be especially influenced by either internal motivations (e.g., an acute displeasure for the negative reaction) or external motivational determinants (e.g., impression management). In support of this model, Fazio et al. (1995) reported a large amount of variability in implicit racial attitudes with a semantic priming task.

The results from our experiments support the MODE model and are inconsistent with Devine's proposal that automatic racial attitudes are shared by all members of a culture. Inspection of each participant's IAT effect supports the assumption that there are individual differences in automatic attitudes: 94% of the Experiment 1 participants had a negative bias to insects. In contrast, only 71% of Experiment 2 participants had IAT effects greater than zero. The smaller proportion of IAT effects in Experiment 2 is inconsistent with Devine's theory that automatic negative evaluations about members of other racial groups are universal. Even when one takes into account that the IAT contains measurement error, Devine's theory would predict a greater uniformity in the implicit racism effect, such as that found in Experiment 1.

Lastly, it is important to note that our results are also consistent with aversive racism theory which proposes that many Caucasians can simultaneously hold conscious egalitarian attitudes and unconscious negative feelings and beliefs about African Americans (Dovidio & Gaertner, 1997; Dovidio et al., 1997). In our second study, participants reported pro-African American or neutral attitudes on self-report measures, but their IAT measures revealed consistent pro-Caucasian beliefs. Taken as a whole, the present experiments support recent theoretical accounts of implicit social attitudes and provide converging evidence for Greenwald et al.'s conclusion that the Implicit Association Task measures automatic expressions of both universal and socially sensitive attitudes and stereotypes.

DISCRIMINANT AND EXTERNAL VALIDITY OF THE IAT

The IAT is a laboratory-based measure of implicit attitudes, and as such has limitations. Important, the discriminant and external validity of the IAT need to be addressed in future research. Just as there is an ecological validity problem when one attempts to generalize from self-report based measures of attitudes (which are prone to self-presentation forces) to actual behavior, the IAT suffers from the similar problem that attitudes elicited by only general semantic instances (e.g., Hispanic names) may not have any predictive relevance for actions and prejudices in real-life settings (e.g., towards a specific Hispanic person). In short, interactions with actual persons may elicit qualitatively different attitudes than would result from exposure to words (e.g., Gilbert & Hixon, 1991), or than would occur if a participant encountered a specific member of a racial group (e.g., Fazio et al., 1995; McConahay, Hardee, & Batts, 1981). Consequently, it is important to establish that the implicit attitudes revealed by the IAT also occur to actual persons (e.g., Chen & Bargh, 1997; Dasgupta, McGhee, Greenwals, & Banaji, 2000; Fazio et al., 1995).

Similarly, it is essential to demonstrate the external validity of the IAT by establishing that implicit attitudes are predictive of behavior outside the laboratory setting (e.g., voting behavior, Carpenter & Banaji, 1997; also see Fazio et al., 1995). The demonstration of a direct relationship between performance on the IAT would provide converging evidence for the validity of the concept of implicit attitudes. Moreover, the external validity of the IAT needs to be established for both averaged group data as well as at the level of the individual subject. While the current experiments support the conclusion that the

Implicit Association Test is sensitive to the average attitudes of a group, the usefulness of the IAT depends upon the demonstration that the IAT can also predict an individual subject's behavior. Unless results from the IAT are predictive of an individual subject's attitudes and/or behavior in other situations (e.g., evaluations of job applicants, the trustworthiness of an eyewitness, likelihood to cease smoking), the theoretical significance and practical utility of the IAT will be limited. The experiments reported in this article indicate that control of word familiarity has the potential to increase the external validity and the IAT's practical applications. Our lab is currently pursuing these issues in order to further the understanding of the relationship between implicit and explicit attitudes.

APPENDICES

APPENDIX A

FAMILIARITY AND PLEASANTNESS RATINGS

For all words and first names used in Experiments 1 and 2, we collected ratings of familiarity and determined the frequency of occurrence of each word in the English language. We also collected pleasantness ratings for the insects, flowers, pleasant and unpleasant word categories. These ratings allowed us to construct word lists matched on both word familiarity and frequency. The following provides a brief overview of participants, criteria for word selection, and rating methods. A more complete description including ratings for individual words can be found in Hayden and O ttaway (in preparation). The methods for rating the African American, Hispanic and Caucasian first names closely followed those employed by Bellezza, Greenwald, and Banaji (1986). The ratings for the flowers, insects, pleasant and unpleasant words are based on a smaller number of participants than the ratings for the first names.

METHOD

PARTICIPANTS

Participants were recruited from the same pool as the other experiments and were tested in groups. For the familiarity ratings of first names, 160 undergraduates (80 females, 80 males) rated either male or female names. For the other word categories, distinct subjects completed the ratings of pleasantness and familiarity. For flowers, 17 females completed the familiarity ratings and 17 females completed the pleasantness ratings. For insects, 17 (13 female, 4 male) completed pleasantness ratings and 17 females completed the familiarity ratings. For the pleasant and unpleasant words, 17 (16 female and 1 male) participants completed the pleasantness ratings and 19 (18 female, 1 male) completed the familiarity ratings.

SOURCES AND SELECTION CRITERIA FOR WORDS USED IN EXPERIMENT 1

We consulted a variety of sources in order to generate the word lists. In addition to the words used by Greenwald et al., Sutherland (1978) provided insect names and Stearn (1992) provided flower names. Pleasant and unpleasant words were selected from Greenwald et al.'s (1998) stimuli, from Bellezza et al. (1986) and from a dictionary. We attempted to create word lists that contained a roughly equal number of low, medium and high familiarity words. The final lists consisted of 67 flower names, 59 insect names and 63 pleasant and unpleasant words. After compiling these lists, we determined each word's frequency of occurrence in the English language. However, relative to all words in the English language, most selected words (e.g., insects) occurred very infrequently or were not listed in the corpuses that we examined (Kucera & Francis, 1967; Thorndike & Lorge, 1944). As a result, accurate estimates of word frequency are unavailable for the words used in Experiment 1. This finding provides further evidence that familiarity is not isomorphic with frequency of occurrence.

Two separate orders of each list were constructed to control for order effects. All words were rated on the dimensions of familiarity and pleasantness. Only instructions for rating the flowers on familiarity and pleasantness are presented. Similar instructions were provided for the ratings of insects and the pleasant and unpleasant words. For purposes of analysis, a response of "very familiar" was coded as a "5" and a response of "very unfamiliar" was coded as a "1." Thus, higher values reflect greater familiarity and pleasantness.

FAMILIARITY OF FLOWERS

The purpose of this experiment is to find out how familiar certain flowers are to college students as flowers differ considerably in this domain. In this experiment you will be rating a list of 67 flowers with regard to how familiar-unfamiliar they are to you. You should read each flower listed very carefully. Then, after you read it, on the response line with the same number as the flower listed, fill in one of the lettered circles (A to E). Make sure you fill in each circle completely. Use the five point scale on the top of each page to rate how familiar-unfamiliar each flower is. If the flower is one that you consider to be very familiar, then give this flower a rating of "A." If the flower is one that is fairly familiar, then give it a rating of "B." If the flower is somewhat familiar, then give it a "C." If the flower is fairly unfamiliar, then give it a rating of "D." Finally, if the flower is one that you consider to be very unfamiliar, then give it a rating of "E." Across all the flowers, try to use all five points on the rating scale. Remember to read each flower listed carefully before you rate it. It will take you about fifteen minutes to complete these ratings. Try to take at least 5 seconds to rate each flower.

PLEASANTNESS–UNPLEASANTNESS OF WORDS

Use the 5 point scale on the top of each page to rate the pleasantness-unpleasantness of each word. If the word is one that you consider to be *very pleasant*, then give this word a rating of "A." If the word is one that is *fairly pleasant*, then give it a rating of "B." If the word is *somewhat pleasant*, then give it a "C." If the word is *fairly unpleasant*, then give it a rating of "D." Finally, if the word is one that you consider to be *very unpleasant*, then give it a rating of "E." Across all the words, try to use all 5 points on the rating scale. Remember to read each word carefully before you rate it.

SELECTION CRITERIA FOR AFRICAN AMERICAN, HISPANIC AND CAUCASIAN NAMES USED IN EXPERIMENT 2

We had considerably more difficulty in compiling a list of first names. The goal was to develop a list of nonoverlapping African American, Hispanic and Caucasian male and female first names that varied in familiarity and frequency of occurrence. We assessed name frequency by examining the 1990 U.S. Department of Justice Census (1998) database of first names, which consists of over 6.1 million names-including 3,184,399 female names (4275 unique names) and 3,003,954 male names (1219 unique names). Although we considered a variety of frequency statistics, we selected cumulative frequency rank as the measure of a name's frequency of occurrence. (All indexes resulted in similar estimates of name frequency). Because the Census Database did not contain the bottom 10% of the least frequently occurring names, the cumulative frequency statistic ranged from 0 to 90 (rather than the usual 0 to 100), with higher scores reflecting less frequently occurring names. We recoded the frequency scale so that high scores reflect more frequently occurring names. This change was made to produce an intuitively consistent measure and improve scale comprehension.

The critical manipulation of Experiment 2 was the selection of names readily identifiable as common names of either African Americans, Hispanic Americans or Caucasian Americans. We compiled a long list of African American, Hispanic and Caucasian names from a variety of books and web sources that offered general baby name information (Babycenter, 1997; Dunkling & Gosling, 1983; Evans, 1992; Shackleford, 1997) as well as name information directed at African Americans (Dillard, 1976; Faulkner, 1994; Monk, 1997) and Hispanic Americans (Arce & Junco, 1995; U.S. Department of Justice, 1973; Woods, 1984).

The frequency of each name in the Census database was then determined. Our initial goal was to identify groups of "high", "moderate", "low", and "very-low" frequency first names for each racial category. However, we were unable to find enough matching names for the high- and very-low frequency categories (e.g., a nonoverlapping list of 10 high-frequency African American names could not be generated). As a result, a majority of the generated list consisted of moderate and low frequency names. All names were combined into a list of 185 male and 158 female names. Once again, we constructed two randomly ordered lists for each gender to control for order effects. Participants read the following instructions for rating the male names. We presented similar instructions for rating the female names.

FAMILIARITY OF MALE NAMES

The purpose of this experiment is to find out how familiar certain first names are to college students as first names differ considerably in familiarity within the United States. Some names you may have encountered frequently and others only rarely, if at all. In this experiment you will be rating a list of 185 male first names with regard to how familiar you are with them.

You should read each name very carefully. Then, after you read it, on the response line with the same number as the name, fill in one of the lettered circles (A to E). Make sure you fill in each circle completely.

Use the 5 point scale on the top of each page to rate the familiarity of each name. If the name is one that is *very familiar* to you, that is, one you often read or hear, then give this name a rating of "A." If the name is one that you are *fairly familiar* with then give it a rating of "B." If the name is *somewhat familiar*, then give it a "C." If the name is *fairly unfamiliar*, then give it a rating of "D." Finally, if the name is one that you have encountered only rarely or have never seen before, then rate the name as *very unfamiliar* by giving it a rating of "E." Across all the names, try to use all 5 points on the rating scale. Remember to read each name carefully before you rate it. We expect you to spend 15 minutes completing these ratings. Try to take at least 5 seconds to rate each name. Any questions?

APPENDIX B

WORDS LISTS

INSECTS

Low Familiarity
bedbug, hornet, lacewing, leaf roller, maggot, mantid, scale, silverfish, spittle bug, and weevil
High Familiarity
ant, aphid, fly, dragonfly, flea, millipede, mite, roach, scorpion, slug, and snail

FLOWERS

Low Familiarity

forsythia, gloxinia, hydrangea, lupin, mimosa, nightshade, peony, salsify, valerian, and wisteria

High Familiarity

carnation, clover, daisy, dandelion, geranium, orchid, rose, sunflower, tulip, and violet

PLEASANT WORDS

diamond, joy, glory, laughter, lucky, miracle, paradise, rainbow, sunrise, and blossom

UNPLEASANT WORDS

accident, bomb, disaster, fail, grief, poison, pollute, poverty, sickness, and thief

FEMALE FIRST NAMES

African American names

Jasmine, Kendra, Latisha, Latoya, Monique, Rochelle, Tasha, Tia, Yolanda, and Yvette

Hispanic Names

Dolores, Felipa, Josefina, Juanita, Margarita, Marta, Monica, Olivia, Rosa, and Sylvia

Caucasian Names

Crystal, Dorothy, Kandace, Mable, Mallory, Margery, Peggy, Ruth, Sherrie, and Shirleen

MALE FIRST NAMES

African American names

Alvin, Damon, Darnell, Deion, Jerome, Leon, Leroy, Terrence, Theo, and Tyrone

Hispanic names

Jorge, Leonard, Luis, Manuel, Miguel, Pablo, Pedro, Ramon, Ricardo, and Rudy

Caucasian names

Alfred, Barry, Chip, Edmund, Hank, Jed, Loren, Marty, Preston, and Wilbur.

Note. N = 156. Ratings of Familiarity are based on a 5 point Likert-like scale with 1 = "Very Unfamiliar" and 5 = "Very Familiar" and were completed by Caucasian American undergraduates. Cumulative Frequency statistics indicate the average percentage of names from a US Census database, regardless of racial identity, that occur less frequently. The database contained the names of 6.1 million people. Cumulative Frequency values range from 0 to 90 with higher values indicating more frequent names. Statistics are based on N = 25 and N = 10 names for Greenwald et al. and Experiment 2, respectively. Appendix A describes the Familiarity and Frequency measures in detail.

REFERENCES

Arce, R. M., & Junco, M. (1995). *Bebes preciosos: 5,001 Hispanic baby names.* New York: Avon Books.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Winston.

Babycenter. (1997). First names. [online] Available at http://www.babycenter.com/babyname/names70.htm [1997, December 5].

Banaji, M. R., & Bhaskar, R. (2000). Implicit stereotypes and memory: The bounded rationality of social beliefs. In D. L. Schacter, E. Scarry et al. (Eds.), *Memory, brain, and belief.* Cambridge, MA: Harvard University Press, 139 - 175.

Banaji, M. R., & Greenwald, A. G. (1994). Implicit Stereotyping and Prejudice. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario symposium* (pp. 55-76). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7,* 136-141.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic activation effect. *Journal of Personality and Social Psychology, 62,* 893-912.

Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research, Methods and Instruments, 18,* 299-303.

Carpenter, C., & Banaji, M. R. (1997). Paper presented at the annual meetings of the Midwestern Psychological Association. Chicago, IL. Cited in Banaji, M. R., & Bhaskar, R. (in press). Implicit stereotypes and memory: The bounded rationality of social beliefs. In D. L. Schacter and E. Scarry (Eds.), *Belief and memory.* Cambridge, MA: Harvard University Press.

Chen, M., & Bargh, J. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541-560.

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

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

Dillard, J. L. (1976). *Black names.* The Hague, the Netherlands: Mouton & Co.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22,* 22-37.

Dovidio, J. F., & Gaertner, S. L. (1997). On the nature of contemporary prejudice: The causes, consequences and challenges of aversive racism. In J. Eberhardt & S. T. Fiske (Eds.), *Racism: The problem and the response.* Newbury Park, CA: Sage.

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.

Dunkling, L., & Gosling, W. (1983). *The facts on file dictionary of first names.* New York: Facts on File Publications.

Evans, C. K. (1992). *Unusual and most popular baby names:* Signet Reference.

Faulkner, B. (1994). *What to name your African-American baby:* St. Martin's Paperback.

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.* (pp. 75-109). New York: Academic Press.

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.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. L., & Kades, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229-239.

Fujita, F., Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: The case for emotional intensity. *Journal of Personality and Social Psychology, 61,* 427-434.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46,* 23-30.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509-517.

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

Greenwald, A. G., & Breckler, S. J. (1985). To whom is the self presented? In B. R. Schlenter (Ed.), *The self and social life,* (pp. 126-146). New York: McGraw Hill.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998) Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1469-1480.

Hayden, D., & Ottaway, S. A. (2001). When subjective familiarity and word frequency are not the same: Ratings of first names, flowers, insects, pleasant and unpleasant words. Manuscript in preparation.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Memory and Language, 30,* 513-451.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989) Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326-338.

Katz, D., & Braly, K. (1933). Racial stereotypes of one-hundred college students. *Journal of Abnormal and Social Psychology, 28,* 280-290.

Klonoff, E. A., & Landrine, H. (1999). Cross-validation of the Schedule of Racist Events. *Journal of Black Psychology, 25,* 231-254.

Kruschke, J. K., (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99, 22-44.*

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Lewicki, P. (1986). *Nonconscious social information processing.* New York: Academic Press.

Massey, D. S., & Eggers, M. L. (1990). The ecology of inequality: Minorities and the concentration of poverty, 1970-1980. *American Journal of Sociology, 95,* 1153-1188.

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data.* Belmont, CA: Wadsworth Publishing.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends upon who is asking and what is asked. *Journal of Conflict Resolution, 25,* 563-579.

Monk, K. (1997). *First names.* [online] Available at http://student-www.uchicago.edu/users/ smhawkin/names/hq.html. [1997, December 5]

Neasdale, D., & Durkin, K. (1998). Stereotypes and attitudes: Implicit and explicit processes. In K. Kirsner et al. (Eds.). *Implicit and explicit mental processes.* (pp. 219-232). New Jersey: Lawrence Erlbaum Associates.

Ong, P. M. (1991). Race and post-displacement earnings among high-tech workers. *Industrial Relations, 30,* 456-468.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118,* 219-235.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 1043-1056.

Roediger, H. L. (1990a). Implicit memory. *American Psychologist, 45,* 1043-1056.

Roediger, H. L. (1990b). Implicit memory: A commentary. *Bulletin of the Psychonomics Society, 28,* 373-380.

Rosch, E. E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192-233.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 501-518.

Shackleford, M. (1997). *First names.* [online]. *Available at http://www.charm.net/~shack/name/babynm.html. [1997, December 5].*

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3-22.

Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology,* 24, 490-504.

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660-1672.

Stadler, M., & Frensch, P. (Eds.). (1997). *Handbook of implicit learning.* Thousand Oaks, CA: Sage Publications.

Stearn, W. T. (1992). *Stearn's dictionary of plant names for gardeners.* London: Cassell.

Sutherland, D. W. S. (1978). *Common names of insects and related organisms.* College Park, MD: Entomological Society of America.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* Teacher's College, NY: Columbia University Press.

U. S. Department of Justice. (1998). [online] Database of first names from the 1990 US census. Available at http://www.census.gov/genealogy/names/. [1998, January 7].

U. S. Department of Justice. (1973). *Spanish name book.* (Vol. M-156): United States Department of Justice: Immigration and Naturalization Service.

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.

Woods, R. D. (1984). *Hispanic first names.* Westport, CT: Greenwood Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, Monograph Supp. No. 2, part 2.