Abstract

Implicit Gender Attitudes

Siri J. Carpenter
2000

The present research examined the favorability of implicit attitudes toward women and men, with a focus on four questions: What role does group membership play in determining such attitudes? Are effects of group membership moderated by cultural construals of the group? What is the relation between implicit and explicit gender attitudes? Are implicit attitudes susceptible to intervention? Four preliminary experiments measured male and female participants' implicit attitudes toward varying construals of women and men, using the Implicit Association Test (IAT). Overall, participants showed more favorable implicit attitudes toward women than toward men. However, this effect was far more pronounced in female participants than in male participants, a finding that Eagly and her colleagues did not reliably obtain in their research using explicit measures of attitude. Additionally, attitudes were affected by the specific construal of the gender groups being evaluated (i.e., mothers vs. fathers, women vs. men, female leaders vs. male leaders). Across all construals, female participants showed strong favorability toward the category female (relative to male), regardless of the particular construal of the category. Male participants, however, showed weaker preferences overall and less valence-consistency in their gender attitudes. These findings drive the conclusion that group membership (i.e., one's own sex) and cultural construal (i.e., the culture's assessment that female=good) both play an important role in defining implicit gender attitudes. Experiment 1 examined the consequences of the valence-

consistency of attitudes toward women and men. This experiment replicated the participant sex effect observed in the four preliminary experiments. More importantly, there was a stronger link between implicit attitudes and explicit candidate preferences among women (the group showing greater valence-consistency) than among men. In Experiment 2, participant sex differences in implicit attitudes toward women were again replicated. Importantly, a mild attempt to influence the strength of implicit association between the concepts *weak* and *female* and between *strong* and *male* (by asking participants to spend five minutes writing an essay about strong women leaders) was successful. Imagining strong women leaders led both male and female participants equally to show a reduction in the implicit stereotype that female=weak and male=strong. The influence of the intervention was, however, restricted to a change in implicit stereotype; it did not influence women's or men's implicit gender attitudes, suggesting that the two processes, implicit attitude and stereotype, may function independently.

Implicit Gender Attitudes

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Siri J. Carpenter

Dissertation Director: Mahzarin R. Banaji

December 2000

Table of Contents

Table of Contents (continued)

List of Figures and Tables

Acknowledgements

There are many people without whose help I would have been unable to complete this dissertation. My very warm thanks go to Mahzarin Banaji, who offered steady guidance, encouraged me to follow an unconventional path, and extended her friendship. The indefatiguable Trish Devine inspired me to be a scientist, a gift for which I will always remain indebted. I thank the members of my committee, Geoffrey Cohen, Marianne LaFrance, Peter Salovey, and Karen Wynn, for their thoughtful comments and suggestions. Tracy Banaszynski, Wil Cunningham, Buju Dasgupta, Thierry Devos, Jack Glaser, Nicole Gleason, Stephanie Goodwin, Aiden Gregg, Kristin Lane, Kristi Lemm, and Brian Nosek made invaluable contributions, and this dissertation would be greatly diminished were it not for their ideas and suggestions. Two outstanding research assistants, Cheryl Conner and Hyura Choi, recruited participants and collected data, and I greatly appreciate their enthusiastic help. My thanks to the capable and cheerful staff of the Yale psychology department business and registrar's offices for their help on occasions too plentiful to count, and to Yale University for providing me with five years of stipend and research support. For the friendship of Kristi Lemm and Tracy Banaszynski, I am profoundly grateful. Wendi Walsh pushed me to be a person I would not have suspected I could be, and her effect on me cannot be overstated. My parents, Ann and Peter Delwiche, showed me the way. Finally, my deep and loving thanks go to Joseph Carpenter, for his unshakeable confidence in me—and for uncomplainingly cleaning the house more often than I did.

Implicit Gender Attitudes

The study of attitudes has held a dominating position in social psychology for almost 100 years, and even by the 1930s, the attitude construct was viewed as "social psychology's most indispensable concept" (Allport, 1935). Over the decades, the definition of the construct has remained largely consistent, in large part because its generality has allowed distinct approaches to measurement to be included within its boundaries. In 1931, Thurstone described attitude as "the affect for or against a psychological object" (p. 261). Similarly, Sarnoff (1960) wrote that an attitude is "a disposition to react favorably or unfavorably to a class of objects" (p. 261). More recently, Eagly and Chaiken (1993) defined attitudes thus: "An attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor."

Of the many aspects of attitude structure and function that have been studied, a particular family of attitudes has held lasting attention. In this dissertation, these attitudes will be referred to as *attitudes toward social groups* (although the research that concerns such attitudes is typically described as research on *prejudice*, especially when it concerns attitudes toward different racial and ethnic groups). One reason for social psychologists' keen interest in attitudes toward social groups may be sociocultural. Social psychology hit its stride as an experimental science in the early decades of the 20th century, an era steeped in intergroup conflict of frightening proportions. Beginning in the 1930s and continuing throughout the middle part of the century, social scientists were increasingly

conscious of the often dangerous consequences of such conflict and the need to understand the origins, operation, and sequelae of attitudes toward social groups.

The issues that form the core of the present research—the role that group membership plays in shaping intergroup attitudes, the sources of attitudes' strength, and attitude consistency and stability—are issues that have received sustained theoretical and empirical attention. Questions surrounding the impact of group membership on attitudes have permeated social psychological theory, and group membership and identity have long been regarded as central, if incompletely understood, elements of intergroup attitudes (Brewer & Brown, 1998; Deaux, 1996; Eagly & Chaiken, 1993, 1998; Fiske, 1998; Tajfel & Turner, 1986). Similarly, attitude scholars have attempted to uncover the circumstances under which attitudes toward social groups can be changed. In an effort to understand the routes by which attitudes gain strength and stability, they have probed variables such as consistency between attitudes and beliefs, and the ambivalence, importance, and accessibility of attitudes (Eagly & Chaiken, 1993, 1998; Fazio & Zanna, 1981; Petty & Cacioppo, 1981; Petty & Krosnick, 1995; Petty & Wegener, 1998; Pratkanis, Breckler, & Greenwald, 1989; Rosenberg, 1960).

This dissertation addresses aspects of each of these classic questions. The focus of the present research is at the intersection of two lines of inquiry, each of which has considerably altered the landscape of social psychology in recent decades. First, social psychologists' interest in understanding prejudice toward social groups has grown rapidly, as is evidenced by a dramatic increase in the amount of journal space and the number of conference presentations dedicated to the topic. A count of journal articles whose titles include the terms "prejudice" or "stereotype" (or their variants) reveals an

escalation of research on those subjects, growing from 315 articles in the 1960s to 1,047 in the 1970s, 1,352 in the 1980s, and 1,781 in the 1990s. Recently, Fiske (in press) observed that at each of two recent scientific meetings, of the Society of Experimental Social Psychology and the European Association of Experimental Social Psychology, fully one-third of the presentations were on the topic of prejudice and intergroup relations.

Accompanying this rising interest in attitudes toward and stereotypic beliefs about social groups, methodological advances in the study of implicit social cognition have enabled psychologists to explore previously hidden aspects of perception, judgment, memory, belief, and attitude that function implicitly, or outside conscious awareness and control (see Kihlstrom, 1990).

Goals of the Present Research

A primary focus in the present research was on the functions that group membership and culturally shared construals of social groups serve in determining implicit attitudes about gender. Four preliminary experiments measured men's and women's implicit attitudes toward several construals of gender (women vs. men, mothers vs. fathers, female leaders vs. male leaders, specific exemplars of female leaders vs. male leaders), which vary in how favorably they are evaluated by the culture. Although it is a central focus of much of the research on social attitudes, the role of group membership (i.e., the inclusion of self in the group) in guiding implicit attitudes, in particular, has yet to be examined. Likewise, there has been no investigation of the relative importance of both group membership and the culture's shared evaluation of a given social group (i.e., cultural notions of whether a group is "good" or not) for implicit judgments of that group.

In the case of gender attitudes, such an investigation is particularly important, as previous research examining explicit attitudes toward women and men has not indicated that group membership is a principal determinant of gender attitude.

The second goal of the present research was to develop greater understanding of the relationship between implicit gender attitudes and explicitly expressed gender preferences. Experiment 1 addressed the link between implicit gender attitudes and explicit preferences for particular (fictional) political candidates. Of particular interest in Experiment 1 was the question of whether attitudes' usefulness for predicting candidate preference is greater when implicit gender attitudes possess greater *valence-consistency* across variations in gender construal, relative to when attitudes are less valence-consistent. It is emphasized that the construct *valence-consistency*, introduced in this report, is distinct from other forms of structural consistency that are well known in the attitude literature, such as evaluative-cognitive consistency (see Eagly & Chaiken, 1998). Nor is the construct synonymous with variability in a numerical or statistical sense; rather, the term is intended to convey the extent to which an individual's evaluation of an entity is consistently favorable or unfavorable across different representations, or construals, of the attitude object, as opposed to fluctuating between positivity and negativity toward that entity.

The third goal of the present research was to probe the malleability versus stability of implicit gender attitudes and beliefs, and the extent to which these two aspects of implicit social cognition operate independently. Experiment 2 addressed these questions by using a mild intervention to make a counterstereotypical construal of women

more accessible. The implications of this increased accessibility for implicit beliefs and attitudes were examined.

Implicit Attitudes

Recent years have witnessed a burgeoning interest among social psychologists in investigating attitudes and beliefs that operate implicitly (for reviews, see Banaji, 2000; Banaji & Greenwald, 1994; Banaji, Lemm, & Carpenter, in press; Bargh, 1996, 1997; Bargh & Chartrand, 1999; Fazio, in press; Greenwald & Banaji, 1995; Wegner & Bargh, 1998). Almost two decades ago, it was first established that the strength of association between an attitude object and evaluative meaning (i.e., favorability on the good-bad dimension) can provide a measure of evaluation of that object (Fazio, Chen, McDonel, & Sherman, 1982; Fazio, Powell, & Herr, 1983; Powell & Fazio, 1983; see Banaji, 2000). Like traditional definitions of attitude, this perspective also takes the position that attitudes are simply the predisposition to respond favorably or unfavorably—even if without necessary conscious reflection—toward an attitude object. The present research, examining the operation of implicit attitudes toward women and men, exists within this tradition.

Evaluative Associations as a Measure of Implicit Attitudes

In an early demonstration of what they termed automatic attitude activation,[1] Fazio, Sanbonmatsu, Powell, and Kardes (1986) argued that attitudes reflect the strength of association in memory between an attitude object and evaluative meaning (good-bad). To measure this association, Fazio et al. used a procedure that they termed *evaluative*

[1] Several terms have been used to describe attitudes and beliefs that reside outside of conscious awareness or control, including *automatic*, *implicit*, and *unconscious*. For consistency, the term *implicit* will be used henceforth in this report.

priming. Participants were presented with individual positive or negative prime words (e.g., *party*, *death*), each followed quickly by a target adjective (e.g., *delightful*, *awful*). Participants' task was to classify each target adjective as *good* or *bad* as quickly as possible. The speed with which participants responded to target adjectives that were congruent or incongruent with the evaluative prime was taken to be a measure of the strength of evaluative association. Prior to Fazio's use of this procedure, similar procedures had been used since the early 1970s to measure semantic priming of concepts in memory (Meyer & Schvaneveldt, 1971; Neely, 1977; Posner & Snyder, 1975).

Fazio et al. (1986) found that participants classified target adjectives more quickly, relative to a baseline no-prime condition, when they had first been presented with an evaluatively congruent prime (e.g., when the prime word *party* was quickly followed by the target adjective *delightful*) than they did when the prime and target were evaluatively incongruent (e.g., when the prime word *party* was quickly followed by the target adjective *awful*). Fazio et al. concluded that strongly held attitudes can be automatically activated upon the mere presentation of an attitude object (or its symbolic equivalent). Subsequent research has demonstrated the generality of Fazio et al.'s initial findings (Bargh, Chaiken, Govender, & Pratto, 1992; Bargh, Chaiken, Raymond, & Hymes, 1996) and has revealed the influence of implicit attitudes on other basic cognitive processes, including visual attention, categorization, decision making, and behavior (Fazio, Blascovich, & Driscoll, 1992; Fazio, Jackson, Dunton, & Williams, 1995; Roskos-Ewoldsen & Fazio, 1992; Smith, Fazio, & Cejka, 1996; for a review, see Fazio, in press).

Implicit Attitudes Toward Social Groups

Like attitudes toward non-social objects, attitudes toward members of social groups have also been shown to operate implicitly (Fazio et al., 1995; Glaser & Banaji, 1999; Greenwald, Klinger, & Liu, 1989; Greenwald, Draine, & Abrams, 1996; Hermans, De Houwer, & Eelen, 1994; Perdue, Dovidio, Gurtman, & Tyler, 1990; Wittenbrink, Judd, & Park, 1997; for a review, see Banaji et al., in press). For example, Fazio et al. (1995) examined participants' implicit attitudes toward Blacks and Whites using an evaluative priming technique similar to the one used in their previous research (Fazio et al., 1986). White participants were presented with Black and White faces, each quickly followed by a positive or negative target adjective. Fazio's group found that participants were quicker to classify positive adjectives as positive when they were preceded by White faces than when they were preceded by Black faces; likewise, participants classified negative adjectives as negative more quickly when they followed Black faces than when they followed White faces.

In addition to research that uses Fazio et al.'s (1986) evaluative priming method, a procedure developed more recently, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), has also assessed the strength of association between attitude objects (e.g., social groups) and evaluation. Like the evaluative priming method, the IAT (described more fully in a later section of this report) measures the strength of association between concepts in memory using a speeded response task. In the published article that introduced the method, Greenwald et al. (1998) demonstrated associations in memory between attitude objects that are widely well-liked (e.g., flowers) and favorable evaluation, and associations between objects that are widely disliked (e.g., insects) and

unfavorable evaluation. In several studies, the IAT has measured implicit evaluation of social groups, including those identified by race, gender, and ethnicity (Dasgupta & Greenwald, 2000; Greenwald et al., 1998; Lemm & Banaji, 1998; Nosek, Cunningham, Banaji, & Greenwald, 2000; Rudman, Greenwald, Mellot, & Schwartz, 1999; Rudman & Kilianski, in press). For example, Greenwald et al. (1998; Experiment 3) found that White participants showed an implicit preference for the social category White over the category Black, a finding that replicates other studies of implicit race attitudes (e.g., Fazio et al., 1995).

With the accumulation of evidence that attitudes toward social groups—and their cognitive counterparts, stereotypic beliefs (Banaji & Hardin, 1996; Banaji, Hardin, & Rothman, 1993; Blair & Banaji, 1996; Devine, 1989; Gilbert & Hixon, 1991; Lepore & Brown, 1997, 1999; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997; van Knippenberg & Dijksterhuis, 1996)—can be called into service without individuals' conscious consent, a new set of theoretical questions concerning implicit attitudes and beliefs has arisen. Among these are questions surrounding the circumstances under which implicit social cognitive processes are initiated (Blair & Banaji, 1996; Glaser & Banaji, 1999), the link between implicit and explicit processes (for a review, see Blair, in press), the malleability or context-dependence of implicit attitudes and beliefs (Bargh, 1999; Blair, Ma, & Lenton, 2000; Dasgupta & Greenwald, 2000; Mitchell, Nosek & Banaji, 1998), the role that group membership plays in guiding implicit social judgment (Richeson & Ambady, 2000), and the extent to which implicit attitudes and beliefs govern behavior (Chen & Bargh, 1997, 1999; Fazio et al., 1995).

Gender Attitudes: A Brief Historical Tour

The present research focused on attitudes toward women and men. Much of the influential attitude research of the 20th century has used gender as the social category of interest. This may be in part because women have long met with considerable social, economic, and political disadvantage. Even in contemporary American society, after the gains of the gender revolution of the 1970s, gender inequality persists (Rhode, 1997). For example, women are vastly under-represented in government and wield less political clout than do men (Kenworthy & Malami, 1999; Mattei, 1998). Compared with their male counterparts, women earn less money at every level of organizational hierarchies (Kay & Hagan, 1995; McGuire & Reskin, 1993). Women are less likely to attain top leadership positions than are men of the same qualifications (Fiske, Bersoff, Borgida, Deaux, & Heilman, 1991; Heilman, 1995; Kay & Hagan, 1995; McGuire & Reskin, 1993; Melamed, 1995; Morrison & von Glinow, 1990) and are more likely to be sexually harassed at work (Cortina, Swan, Fitzgerald, & Waldo, 1998; Fitzgerald**,** Drasgow, & Magley, 1999). In every sphere outside the home—and often there, too (see Monson & Langhinrichsen-Rohling, 1998)—women confront discrimination and disadvantage.

Not surprisingly, given the pervasiveness of gender discrimination, psychologists long assumed that the culture was fraught with negative attitudes and stereotypic beliefs about women, and the literature overwhelmingly reported that this was the case (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Goldberg, 1968; Greenglass, 1982; Lips, 1988; Matlin, 1987; McKee & Sherriffs, 1957; Rosenkrantz, Vogel, Bee, Broverman, & Broverman, 1968). Investigations consistently reported that women are seen as incompetent, immature, and dependent—in no uncertain terms, the

literature asserted that society viewed women as inferior to men. That position remained largely unchallenged until the late 1980s, when Alice Eagly and her colleagues argued that methodological flaws, combined with a misplaced eagerness to assume that discrimination was attributable to negative attitudes per se, may have masked a quite different state of affairs (Eagly & Mladinic, 1989; Eagly, Mladinic, & Otto, 1991; also see Deaux & Lewis, 1983; Del Boca, Ashmore, & McManus, 1986).

Much of the literature on gender attitudes and on the evaluative meaning of gender stereotypes, Eagly and her colleagues observed, had relied on methods used by Broverman, Rosenkrantz, and colleagues (Broverman et al., 1972; Rosenkrantz et al., 1968). In this research, the favorability or unfavorability of stereotypes was calculated by simply comparing the total number of positive versus negative traits ascribed to men and women. As Eagly's group noted, other researchers have questioned whether such a method is an appropriate means of assessing the overall favorability of stereotypic beliefs, suggesting that it may be more appropriate to instead consider the average evaluations of the traits that participants select as characterizing women and men. Indeed, when such an analysis was applied to the Broverman and Rosenkrantz data itself, no difference was observed in the favorability of stereotypes regarding men and women (Eagly et al., 1991).

A further source of confusion in the gender attitudes literature, Eagly's group contended, was the mistaken belief that a popular gender scale, the Attitudes Toward Women scale (AWS; Spence & Helmreich, 1978) was, in fact, a measure of attitudes toward or global evaluations of women (Eagly & Mladinic, 1989). Rather, as the scale's authors stated when the scale was published, the AWS is really a measure of the "rights,

roles, and privileges women ought to have or be permitted." Because of its deceptively simple title, however, the scale (which typically reveals a belief that women should have fewer rights and privileges than men should have) became widely perceived as a measure of gender attitudes (i.e., favorability toward women). Eagly and Mladinic (1989) argued that this occurrence further muddied the issue of the favorability of attitudes toward women and men.

Eagly and her colleagues argued that even if discrimination against women is alive and well, the culture's attitudes toward women need not be predominantly negative (Eagly & Mladinic, 1989; Eagly et al., 1991). They examined attitudes and beliefs about women and men using semantic differential attitude scales (e.g., good-bad, positive-negative, pleasant-unpleasant), gender-stereotype trait-rating scales, a free-response measure of gender stereotypes, and a free-response measure of the emotions elicited by each target group. Their results showed that, contrary to what had been asserted confidently in the literature, both male and female participants expressed more favorable attitudes toward women than toward men, a finding that was recently replicated by Carpenter and Banaji (1997). These findings underscore an important principle: that attitudes toward social groups, stereotypic beliefs, and discrimination, although often correlated, are distinguishable and even independent constructs. As Eagly and colleagues' research on attitudes toward women and men first demonstrated, negativity along one of the dimensions need not mean that negativity will likewise be observed in the others.

Three Competing Hypotheses Regarding Implicit Gender Attitudes

Liking for Women May Be Restricted to Explicitly Measured Attitudes

One possible explanation for Eagly and her associates' unprecedented finding that both male and female participants demonstrated favorable attitudes toward women may be that the measures that Eagly used required participants to reflect consciously on the categories *male* and *female* and to provide evaluations of each. Thus, Eagly's group may have obtained the effect they did—strong positive evaluation of women—either because participants' responses reflected self-presentational or impression management concerns, or because they were unable to accurately introspect on their gender attitudes or the sources of those attitudes (see Nisbett & Wilson, 1977). Eagly and her colleagues collected data from college students in the 1980s, and their participants were likely to be well aware of the gender revolution and the importance of gender parity; this awareness may have colored participants' responses on the explicit attitude measures.

If it is the case that Eagly's attitude measures captured participants' attitudes only at a conscious level, then implicit measures may not show the same pattern of findings that Eagly's measures did. Indeed, Fazio et al. (1995) suggested that measures of implicit or automatic attitudes, such as evaluative priming, bypass self-presentational and impression management efforts, providing a "bona fide pipeline" to individuals' social attitudes. Their view is that measures of implicit attitudes are a better indicator of individuals' "true" evaluation than are measures of explicit attitudes.

An equally plausible possibility is that Eagly's participants were responding candidly, but that their responses on the attitude measures—which assessed only conscious attitudes—revealed only one aspect of their evaluation of female leaders. This

has been the favored view of Banaji and her colleagues (e.g., Banaji, 2000; Banaji & Greenwald, 1994; Greenwald & Banaji, 1995), who argue that the value of implicit measures extends beyond their ability to bypass the social desirability and self-presentational concerns of explicit measures. They contend that implicit measures reflect an aspect of attitude and belief that is conceptually and perhaps functionally distinct from (albeit no more "true" than) explicit social cognition.

Numerous studies indicate that measures of implicit attitudes can be sensitive to negative attitudes and stereotypes undetected by conventional measures of explicit attitudes (Cunningham, Nezlak, & Banaji, 2000; Cunningham, Preacher, & Banaji, in press; Dasgupta, McGhee, Greenwald, & Banaji, in press; Devine, 1989; Dovidio et al., 1997; Fazio et al., 1995; Greenwald et al., 1998; Mitchell et al., 1998; Nosek et al., 2000). For example, a website that opened in September 1998 (www.yale.edu/implicit) has now assessed thousands of anonymous respondents' implicit and explicit attitudes toward a number of social groups (by August, 2000, more than 900,000 tasks had been completed). Although the website's main purpose at present is to serve as a demonstration site for educational purposes, the data that have emerged are instructive. Pertinent to the present discussion of dissociations between implicit and explicit evaluation, for example, analyses of the race attitude data show that respondents' implicit attitudes toward Blacks relative to Whites are reliably more negative than are their explicitly endorsed attitudes (Nosek et al., 2000).

Recently, Cunningham et al. (2000) examined explicit and implicit attitudes of college students toward social groups categorized by race, sexual orientation, social class, religion, and nationality, in order to explore the possibility of a general implicit

ethnocentrism. Their results were consistent with web data: For every social group, measures of explicit attitude revealed significantly more favorable evaluation of the disadvantaged group (i.e., Blacks, gays, poor people, Jews, or foreigners) than was the case for IAT measures of implicit attitudes.

Thus, several studies have revealed a discrepancy between individuals' explicitly endorsed egalitarian attitudes and their implicit evaluations of social groups, which appear to be more biased. This evidence that implicit attitudes are sometimes at odds with self-reported conscious attitudes raises a question about Eagly and her colleagues' (Eagly & Mladinic, 1989; Eagly et al., 1991) findings: Will the favorable attitudes toward women that they observed in self-reported attitudes also be observed on measures of implicit attitudes? Or, will measures of implicit gender attitudes reveal negative evaluation, in keeping with the discrimination that women experience? Although one possibility is that Eagly et al.'s findings with explicit attitudes may be mirrored by measurements of implicit attitudes, the research discussed above suggests an alternative possibility:

> *Hypothesis 1: On a measure of implicit gender attitudes, an implicit preference for men over women will be observed, among both male and female participants.*

Group Membership May Determine Attitude

Although Eagly and her colleagues (Eagly & Mladinic, 1989; Eagly et al., 1991), did not observe consistent participant sex differences in gender attitudes in their research, neither was that their emphasis. In the present research, participant sex differences in implicit gender attitudes—or the lack thereof—is of particular interest.

An alternative prediction to the one presented above finds its roots in social identity theory and its close relative, self categorization theory (Hogg & Abrams, 1988; Mullen, Brown, & Smith, 1992; Oakes & Turner, 1980; Tajfel, 1981; Tajfel & Turner, 1986; Turner, 1985; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). These perspectives posit that positive self-esteem and self-categorization, born out of identification with a social group, lead individuals to make biased evaluations of salient out-groups and to engage in discriminatory behaviors favoring the in-group.

Hundreds of studies have found support for social identity theory and self categorization theory's predictions that individuals demonstrate evaluative and behavioral bias in favor of their own in-groups, at the expense of out-groups (Abrams & Hogg, 1988; Bourhis, 1994; Bourhis, Turner, & Gagnon, 1997; Hogg & Abrams, 1990; Mullen et al., 1992; Perdue et al., 1990; Rubin & Hewstone, 1998; Tajfel & Turner, 1986). The great majority of these experiments have used the minimal group paradigm, a procedure in which participants are randomly assigned to membership in one of two arbitrary social groups (e.g., X-group vs. Y-group, overestimators vs. underestimators, Klee- vs. Kandinsky-likers). This corpus of research shows that, despite the triviality of the distinction between groups, individuals allot greater resources to members of their own group than to members of the out-group, give greater benefit of the doubt to members of their own group in ambiguous situations, and evaluate members of their own group more positively than members of the out-group (for reviews, see Brewer, 1979; Brewer & Brown, 1998). Like research using minimal groups, examinations of in-group bias among members of real social groups has further supported the predictions of social identity theory and self categorization theory (Lindeman & Sundvik, 1995). In short, the evidence

indicating that group members take care of their own, attitudinally and behaviorally, is persuasive.

Recent advances in implicit social cognition lend additional support to the notion that group membership may be an important determinant of implicit social attitudes (i.e., liking for one's own group). For example, in a recent examination of implicit evaluations that employed the minimal group paradigm, Reed (2000) randomly assigned participants to be members of one of two groups, the *Blue Jays* or the *Canaries*. On an IAT measure of implicit attitudes, participants in both groups showed an implicit preference for their in-group over the out-group. Thus, even on implicit measures of the association between the in-group and favorable evaluation, a small intervention of assigning oneself to membership in the group produced liking for the group's members.

According to a new unified theory of implicit social cognition (Greenwald, Banaji, Rudman, Farnham, Nosek, & Rosier, 2000a; Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, 2000b), social attitudes and beliefs are informed by the association of the concept of *self* with affective and cognitive representations of social groups. That is, to the extent that a given social group is strongly linked with the self, evaluation of that group will be favorable, and from this will also follow favorable orientation toward attributes associated with the group. A wide variety of investigations involving real social groups, like Reed's (2000) minimal group study, lend support to the unified theory's contention that group membership is central to implicit attitudes and beliefs. Among the many investigations that have successfully tested the theory's postulates are studies of Koreans versus Japanese (Greenwald et al., 1998), Blacks versus Whites (Greenwald et al., 1998; Nosek et al., 2000; Richeson & Ambady, 2000), women

versus men (Nosek, Banaji, & Greenwald, 1998; Rudman, Greenwald, & McGhee, 1999), Democrats versus Republicans (Nosek et al., 2000), East Germans versus West Germans (Schiessl, 2000), Bavarians versus North Germans (Neumann et al., 1998), vegetarians versus omnivores (Swanson, Rudman, & Greenwald, in press), and members of rival university residential colleges (Lane & Banaji, 2000).

Investigations that support the predictions of social identity theory, self categorization theory, and the unified theory of implicit social cognition inform a second hypothesis for the present research, emphasizing the role of group membership in shaping implicit attitudes. That is, individuals should show greater implicit liking for their own groups than for an out-group. Like the pattern delineated in Hypothesis 1, such a pattern would also diverge from Eagly and colleagues' result of parity in males' and females' liking for men and women.

> *Hypothesis 2: On a measure of implicit gender attitudes, an implicit preference for the category female over the category male will be observed among female participants, and the opposite preference will be observed among male participants.*

Group Membership and Cultural Construal: An Integrative Perspective

A third, integrative possibility is proposed, that there may be two independent sources of implicit attitudes. The first source, consistent with social identity theory, self categorization theory, and the unified theory of implicit social cognition, is through group membership. To the extent that group membership is a meaningful source of individuals' self-concept, greater implicit liking for their own group relative to an out-group should follow. Additionally, however, a second path to implicit liking is offered: Not all groups are liked equally, even by their own members. Instead, it is proposed, differences in

liking for the in-group are moderated by a culture's imposition of the quality of *good* or *bad* on the group. Thus, in this analysis, two mechanisms—group membership and culturally shared evaluation—may combine to determine implicit attitudes toward social groups.

There is some empirical evidence that membership in particular race, age, or religious groups may interact with the culture's evaluation of those groups to determine implicit evaluation (Rudman, Greenwald, Mellott, et al., 1999; Nosek et al., 2000). For example, web data revealed that although White respondents showed an implicit in-group preference, Black respondents showed no reliable preference for either Black or White (Nosek et al., 2000). Banaji, Greenwald, and Rosier (1998) have suggested that such a finding reflects a tension between group membership and the culture's evaluation of Black Americans.

The question of cultural construal is particularly relevant for the study of gender attitudes. Gender groups are easily and commonly divided into different subgroups (Carpenter, 1994; Carpenter & Trentham, 1998; Coats & Smith, 1999; Eckes, 1994a, 1994b; Kunda & Oleson, 1995; Mitchell et al., 1998; Taylor, 1981), which vary in how favorably they are evaluated by the culture. Many studies indicate that despite evidence of positive attitudes toward women in general (Carpenter & Banaji, 1997; Eagly & Mladinic, 1989; Eagly et al., 1991; Glick & Fiske, 1996), such favorable evaluations are tenuous, fading when women are evaluated in specific, non-traditional contexts, such as feminist and leadership roles (Burns-Glover & Veith, 1995; Butler & Geis, 1990; Deal & Stevenson, 1998; Eagly & Karau, 1999; Eagly, Makhijani, & Klonsky, 1992; Haddock & Zanna, 1994; Heilman, 1995; Heilman, Block, & Martell, 1995; Rudman & Glick, in

press; Rudman & Kilianski, in press; Saris, Johnston, & Lott, 1995; Swim, Borgida, Maruyama, & Myers, 1989).

Attitudes toward women who are leaders have received special research attention, and negativity toward female leaders has been shown in numerous investigations. For example, one study found that female managers were evaluated more negatively than were male managers (Heilman et al., 1995). In a group setting, another study showed, female leaders received more negative nonverbal responses than did male leaders who behaved identically (Butler & Geis, 1990). Likewise, an examination of students' attitudes toward prospective professors revealed more positive evaluations of male professors than of female professors (Burns-Glover & Veith, 1995). In a meta-analysis of 62 experimental studies of attitudes toward women leaders, Eagly et al. (1992) reported that women are evaluated significantly less favorably when they are portrayed or enacted in leadership positions than are men in such positions.

Thus, there is substantial evidence that explicit gender attitudes are contingent on the particular construal of gender being considered. When women fulfill more traditional roles and personify more communal, warm, and nurturing characteristics, the culture views them more favorably than when they occupy roles or personify characteristics that challenge traditional expectations (Glick & Fiske, 1996). This variation in cultural evaluation of different gender subgroups may interact with group membership to shape implicit attitudes. In the case of implicit gender attitudes, then, it may be that, as in Eagly and colleagues' results, women are evaluated more favorably than are men, but that group membership and the culture's construal of different gender subgroups qualifies that effect.

Hypothesis 3a (favorable evaluation of women): On measures of implicit attitudes, implicit preferences for women over men will be shown among both male and female participants.

Hypothesis 3b (influence of group membership): The preference for female will be stronger among female participants than among male participants.

Hypothesis 3c (influence of cultural construal): Less implicit liking for women will be observed for women construed as leaders than for construals that are favorably evaluated by the culture, such as mothers or the generic category, women.

Preliminary Experiments

Overview

Between 1997 and 1999, four preliminary experiments (PE1-PE4) were conducted to examine implicit attitudes toward different construals of men and women, using the IAT (Greenwald et al., 1998).[2] This method measures the strength of association between attitude objects (e.g., social groups such as males and females) and evaluation. The IAT requires participants to categorize stimuli using two designated keys on a computer keyboard. When strongly associated concepts are assigned to the same key, participants are expected to respond more quickly than when more weakly associated concepts are assigned to the same key. For example, an individual who holds favorable implicit attitudes toward the category *female* (relative to the category *male*) would be expected to respond more quickly to experimental trials in which female names are paired with pleasant words and male names are paired with unpleasant words (abbreviated as the female+pleasant key configuration), compared with when female

[2] PE1, PE2, and PE4 used Micro Experimental Laboratory (MEL) software, version 2.0 (Schneider, 1990). PE3 used Inquisit software (Draine, 1998). These two experimental programs yielded comparable results.

names and unpleasant words share the same key and male names and unpleasant words share the same key (abbreviated as the female+unpleasant key configuration).[3] The magnitude of the difference between the two key configurations, known at the IAT effect, indicates the strength of individuals' implicit preference for one social group over another (e.g., for women over men).

The present report of the preliminary experiments is brief. The main purpose is to communicate a pattern of findings across several representations, or construals, of gender groups (e.g., mother/father, female/male, female leader/male leader). These experiments then formed the basis of the two main experiments that follow. Additionally, small sample sizes in some of the preliminary experiments make it appropriate, when interpreting results, to focus attention on effect sizes (Cohen's *d*), rather than on tests of statistical significance (*p* values).

In PE1-PE4, participants were presented with male and female names (e.g., *Peter, Joseph, Jane, Barbara*) and with pleasant and unpleasant evaluative words (e.g., *Laughter, Vacation, Excellent, Cancer, Maggot, Terrible*). Evaluative words were selected based on the results of a pretest of 294 adjectives, nouns, and verbs and on norms established by Belleza, Greenwald, and Banaji (1986). For the IATs, word lists that were matched in mean word length and valence were constructed (for a complete list of evaluative stimuli used in the present research, see Appendix A). In experimental conditions that assessed participants' implicit attitudes toward particular gender subgroups (e.g., mothers vs. fathers, or male leaders vs. female leaders), names were accompanied by denotative labels (e.g., MOTHER, FATHER or MALE LEADER,

[3] Although this report abbreviates the two key configurations as female+pleasant and female+unpleasant, implicit evaluations made in the IATs involve simultaneous evaluation of males and females.

FEMALE LEADER). After a practice period, names and evaluative words were presented in alternating trials, in blocks of 30-100 trials (with variation between experiments). Stimuli in the IATs were presented randomly, with the constraint that no more than four consecutive correct responses were assigned to the same key.

Participants' task in the IATs in experiments PE1-PE4 was to classify each name as either male or female (or the appropriate gender subgroup) and to classify each evaluative word as either pleasant or unpleasant.[4] Responses on the IATs were made using keys on the left and right side of the computer keyboard. A reminder was positioned at the top of the computer screen, indicating which responses were assigned to the left key and which to the right. When participants made an incorrect response, the computer emitted a short (250 ms) tone to provide them with error feedback.[5] In all experiments, the order of the key configurations (i.e., female+unpleasant and female+ pleasant) was counterbalanced between subjects.[6] The main elements of each of the four preliminary experiments are summarized in Table 1.

Data Preparation

For each stimulus in the IATs, participants' response latencies (with precision to 1 ms) were recorded by the computer. As is customary to approximate a normal distribution, a log transformation was performed on the raw response latencies. Each participant's IAT effect score was computed by subtracting his or her mean response latency in the female+pleasant configuration from his or her mean latency in the

---

[4] In PE3, participants classified evaluative words as *good* and *bad* instead of as *pleasant* and *unpleasant*.

[5] In PE3, a red letter *X* appeared on the computer screen after incorrect responses.

[6] The order in which the two configurations were presented did not affect the pattern of results for the IATs in any of the experiments; therefore, analyses collapse across this variable.

female+unpleasant configuration. Thus, the IAT effect reflects the extent to which participants showed an implicit preference for female over male. Male and female participants' IAT effects for PE1-PE4 are shown in Figure 1.

Table 1

Summary of Elements of Preliminary Experiments

Experiment	Total N	IAT Comparison Groups
PE1	67	Female/Male Mother/Father
PE2	45	Female/Male Female Leader/Male Leader
PE3	66	Female Leader/Male Leader (no prime) Female Leader/Male Leader (leader prime)
PE4	44	Female/Male Specific Female Leader/Male Leader

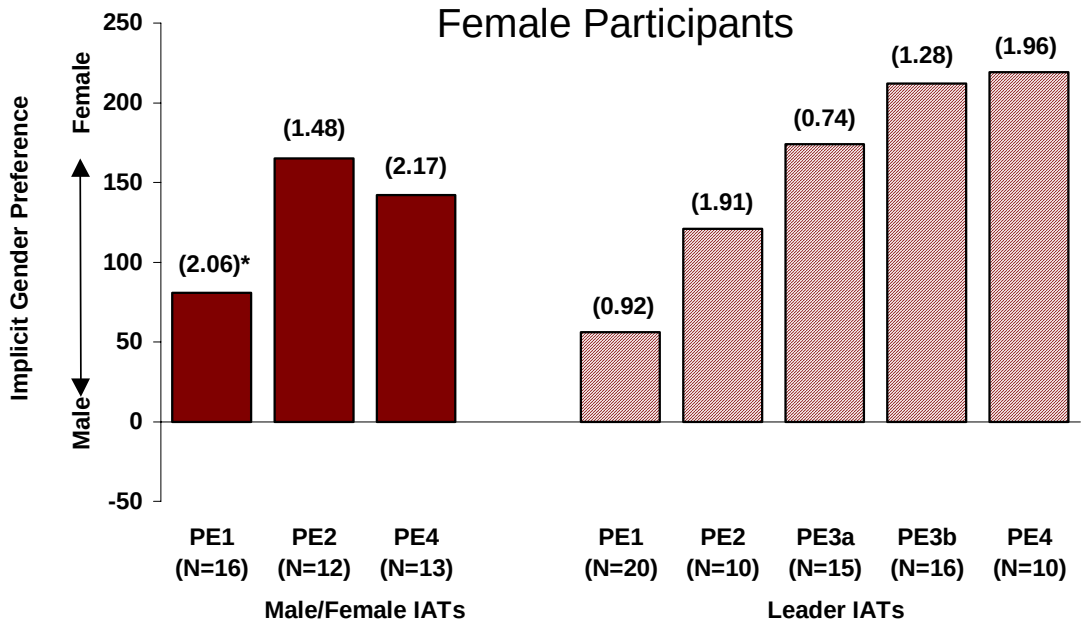
Preliminary Experiment 1 (PE1)

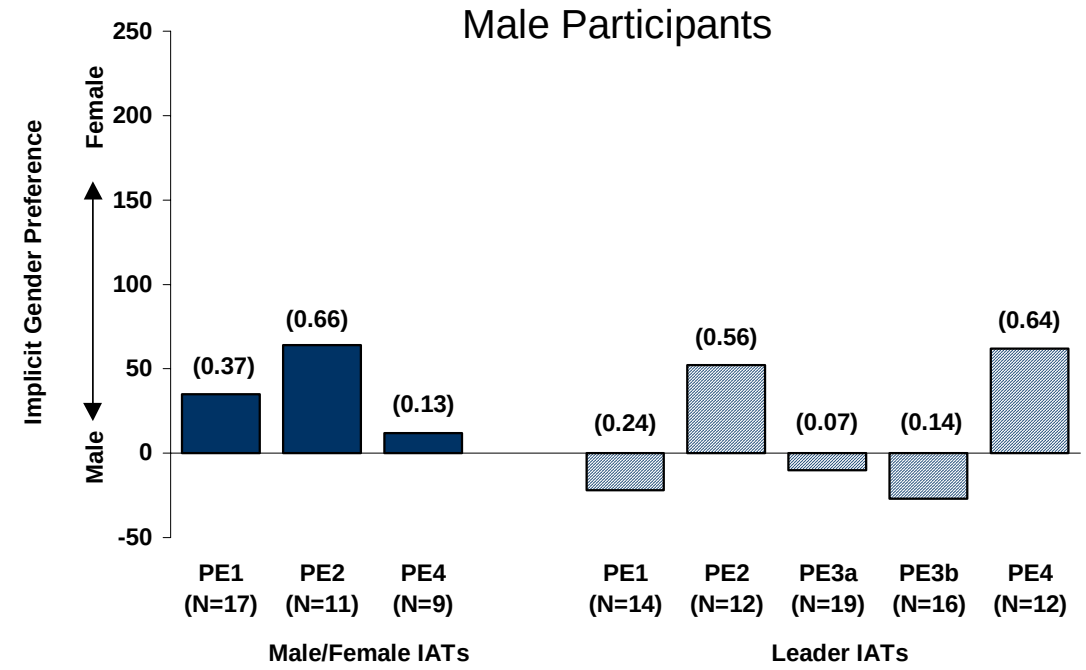
Goal

The goal of PE1 was to scrutinize males' and females' implicit attitudes toward two different gender subgroups: fathers versus mothers, and male leaders versus female leaders. Previous research on explicit gender attitudes (e.g., Carpenter & Banaji, 1997; Eagly & Mladinic, 1989; Eagly et al., 1991) has suggested that men and women hold similarly favorable attitudes toward women. PE1 examined whether this pattern of liking for women is also observed at an implicit level. PE1 also began to explore the extent to which different construals of women (e.g., mothers, female leaders) would elicit similarly positive implicit attitudes.

Figure 1

IAT Effects (in Milliseconds) for Preliminary Experiments (PE1 – PE4)





Note. *Values in parentheses above bars are Cohen's *d* effect sizes.

Method

Participants in PE1 were 36 female and 31 male undergraduates who participated in the experiment for course credit. Half the participants completed an IAT that assessed implicit attitudes toward mothers versus fathers (Father/Mother IAT). In this IAT condition, names were accompanied by the labels FATHER and MOTHER, and participants' task was to classify each name as a father name or a mother name. The other half of the participants completed an IAT with stimuli that were identical to those in the Father/Mother IAT, with one exception. In this IAT (Leader IAT), names were labeled MALE LEADER or FEMALE LEADER, and participants categorized each name as a male leader name or a female leader name. Because the stimuli presented in the Father/Mother and Leader IATs were identical except for the group labels that accompanied male and female names, any differences observed between the two IATs can be attributed to participants' construal of the gender subgroups.

Results

Father/Mother IAT. The dependent variable for PE1 (and for all subsequent IAT measures in the present research) is the *IAT effect*. This variable is obtained by subtracting each participant's mean log response latency in the female+pleasant key configuration from his or her mean latency in the female+unpleasant key configuration. Thus, the IAT effect is a difference score; a positive IAT effect score reflects an implicit preference for female over male, and a negative IAT effect score reflects the opposite implicit preference.

On the Father/Mother IAT, both male and female participants exhibited favorable implicit attitudes toward mothers relative to fathers, with participants responding more

quickly when mother names were paired with pleasant words than when mother names were paired with unpleasant words, F (1, 31) = 22.06, $p < .0001$ (see Figure 1). However, this overall effect was qualified by a significant participant sex effect, F (1, 31) = 4.28, $p < .05$. Female participants' implicit preference for mothers over fathers was statistically significant, F (1, 15) = 63.68, $p < .0001$, $d = 2.06$.[7] Male participants' implicit preference for mothers over fathers was not statistically significant and was considerably smaller in magnitude than was female participants' preference, F (1, 16) = 2.21, $p < .16$, $d = 0.37$.

Leader IAT. On the Leader IAT, a similar pattern of results emerged. Female participants demonstrated an implicit preference for female leaders over male leaders, F (1, 19) = 15.89, $p < .001$, $d = 0.92$. Male participants exhibited a small (but not statistically significant) preference for male leaders over female leaders, F (1, 13) = 0.78, ns, $d = 0.24$. This differential pattern of responding was captured by a significant participant sex effect, F (1, 32) = 8.87, $p < .01$.

Preliminary Experiment 2 (PE2)

Goal

The goal of PE2 was to extend the results of PE1, in which implicit attitudes toward male and female leaders were compared with attitudes toward mothers and fathers. PE2 measured implicit attitudes toward men and women in general as well as toward male leaders and female leaders.

Method

Participants in PE2 were 22 female and 23 male undergraduates who participated for course credit or for monetary compensation ($5). As in the Leader IAT in PE1, male

[7] Cohen's *d* was used as a measure of effect size. In this metric, .2, .5, and .8 are customarily considered small, medium, and large effects, respectively.

and female names were accompanied by the labels MALE LEADER and FEMALE LEADER. In the Male/Female condition of the experiment, male and female names were not labeled.

Results

Male/Female IAT. As in PE1, both male and female participants exhibited an implicit preference for women over men on the Male/Female IAT; the IAT effect was statistically significant, with participants responding more quickly when female names were paired with pleasant words than when female names were paired with unpleasant words, F (1, 21) = 25.02, $p < .0001$ (see Figure 1). However, this effect was qualified by a significant participant sex effect, F (1, 21) = 4.66, $p < .05$. Female participants' positive implicit preference for women over men was statistically significant, F (1, 11) = 24.34, $p < .001$, $d = 1.48$. For male participants, this preference was only marginally significant and was of smaller magnitude, F (1, 10) = 4.36, $p < .06$, $d = 0.66$.

Leader IAT. On the leader IAT, a significant IAT effect was also observed, with participants responding more quickly when female leader names were paired with pleasant words than when they were paired with unpleasant words, F (1, 20) = 24.89, $p < .0001$. However, as before, there was also a significant participant sex effect, F (1, 20) = 4.84, $p < .05$. Female participants' implicit preference for women over men was statistically significant, F (1, 9) = 32.78, $p < .001$, $d = 1.91$. Inconsistent with the results of PE1, in which males demonstrated an implicit preference for male leaders over female leaders, male participants in PE2 showed an implicit preference for female leaders over male leaders. Although the number of participants in PE2 did not allow this effect to

reach conventional levels of significance, F (1, 11) = 3.49, ns, the magnitude of the effect was medium in size, d = 0.56.

Preliminary Experiment 3 (PE3)

Goal

The primary goal of PE3 was to test whether the previous observation of male participants' implicit preference for female leaders over male leaders was reliable. A secondary goal was to explore whether presenting participants with exemplars of specific (fictional) female leaders would affect their implicit attitudes toward male versus female leaders by activating differing construals of those groups.

Method

Participants in PE3 were 31 female and 35 male undergraduates who participated for course credit. In PE3, implicit attitudes toward male versus female leaders were assessed, in two experimental conditions. About half of the participants were in a no-prime condition, in which they completed the Leader IAT immediately after the experiment began. This condition mirrors the procedures for PE1 and PE2. The remaining half of the participants were in a leader-prime condition, in which they were presented with an article, ostensibly taken from *Business Week* magazine, entitled "Top Women Executives of 1998." The article, which was actually fabricated, provided participants with profiles of four companies, all led by female executives who were described as strong and demanding leaders. Participants were instructed to spend several minutes reading the article and to form impressions of the companies and executives in the profiles. Following this task, participants completed a Leader IAT. As in the Leader IATs in PE1 and PE2, male and female names were accompanied by the labels MALE

LEADER and FEMALE LEADER. In PE3, only attitudes toward male and female *leaders* were measured; there was no measure of attitudes toward any other gender subgroups.

Results

No-Prime condition (PE3a). The participant sex effect in experiment PE3a was statistically significant, F (1, 32) = 7.01, $p < .05$ (see Figure 1). Consistent with experiments PE1 and PE2, female participants demonstrated favorable implicit attitudes toward female leaders relative to male leaders, responding more quickly in the female+pleasant key configuration than in the female+unpleasant configuration, F (1, 14) = 7.60, $p < .05$, d = 0.74. In contrast, male participants showed a small, nonsignificant implicit preference for male leaders over female leaders, F (1, 18) = 0.08, ns, d = 0.07. This finding resembles the results of PE1 but is inconsistent with PE2, in which males showed positive implicit attitudes toward female leaders relative to male leaders.

Leader-Prime condition (PE3b). As in PE3a, the participant sex effect in experiment PE3b was statistically significant, F (1, 30) = 13.06, $p < .01$. After being primed with female leader exemplars, female participants again demonstrated favorable implicit attitudes toward female leaders relative to male leaders, F (1, 15) = 24.40, $p < .001$, d = 1.28. In contrast, male participants showed a nonsignificant implicit preference for male leaders over female leaders, F (1, 15) = 0.29, ns, d = 0.14. This finding parallels the results of PE1 and PE3a, and is inconsistent with PE2.

Preliminary Experiment 4 (PE4)

Goal

The goal of PE4 was to further examine implicit attitudes toward different construals of women, comparing implicit attitudes toward men and women in general with implicit attitudes toward actual, specific male leaders and female leaders.

Method

Participants in PE4 were 23 female and 21 male undergraduates who participated for course credit or for monetary compensation ($5). Half the participants completed an IAT measuring attitudes toward males versus females, and half completed an IAT measuring attitudes toward specific male leaders and female leaders. Before completing the IAT, all participants first spent several minutes familiarizing themselves with a list of 10 male and 10 female names. For participants in the Male/Female condition of the experiment, this list contained only generic first names, with no accompanying information. For participants in the Leader condition of the experiment, the list contained first and last names of real individuals and their professional positions (e.g., *William Cohen, Secretary of Defense; David Murdock, CEO, Dole Food Corporation; Olympia Snowe, Senator from Maine; Jill Barad, CEO, Mattel Corporation*; see Appendix B). For each participant, the names on this list were presented in one of four random orders.

Next, participants completed the IAT portion of the experiment. Participants in the Male/Female condition completed an IAT concerning their implicit attitudes toward men and women in general, as in PE1 and PE2. Participants in the Leader condition completed an IAT concerning their implicit attitudes toward the specific male and female leaders with whom they had just familiarized themselves.

Results

Male/Female IAT. As in the previous experiments, analyses of the Male/Female IAT revealed positive implicit attitudes toward women relative to men, among both male and female participants (see Figure 1). This overall pattern is reflected in a significant overall IAT effect, F (1, 20) = 19.88, $p < .001$. This main effect was qualified by a significant participant sex effect, F (1, 20) = 14.56, $p < .001$. As in every previous preliminary experiment, this implicit preference was statistically significant and of large magnitude for female participants, F (1, 12) = 56.45, $p < .0001$, $d = 2.17$. Also consistent with the previous preliminary experiment, male participants likewise showed a pattern of implicit preference for females over males, although the effect was much smaller than was the case for female participants, F (1, 8) = 0.13, ns, $d = 0.13$.

Specific Leader IAT. On the IAT assessing attitudes toward specific male leaders and female leaders, both male and female participants showed an implicit preference for female leaders over male leaders; the IAT effect was statistically significant, F (1, 20) = 34.87, $p < .0001$. The participant sex effect was also significant, however, F (1, 20) = 9.82, $p < .01$. Female participants showed a strong preference for female leaders over male leaders, F (1, 9) = 34.45, $p < .001$, $d = 1.96$. Male participants showed a marginally significant implicit preference for female leaders over male leaders, F (1, 11) = 4.55, $p < .06$, $d = 0.64$. However, as the effect size indicates, the lack of significance for male participants is due to the small sample size.

Discussion of PE1-PE4

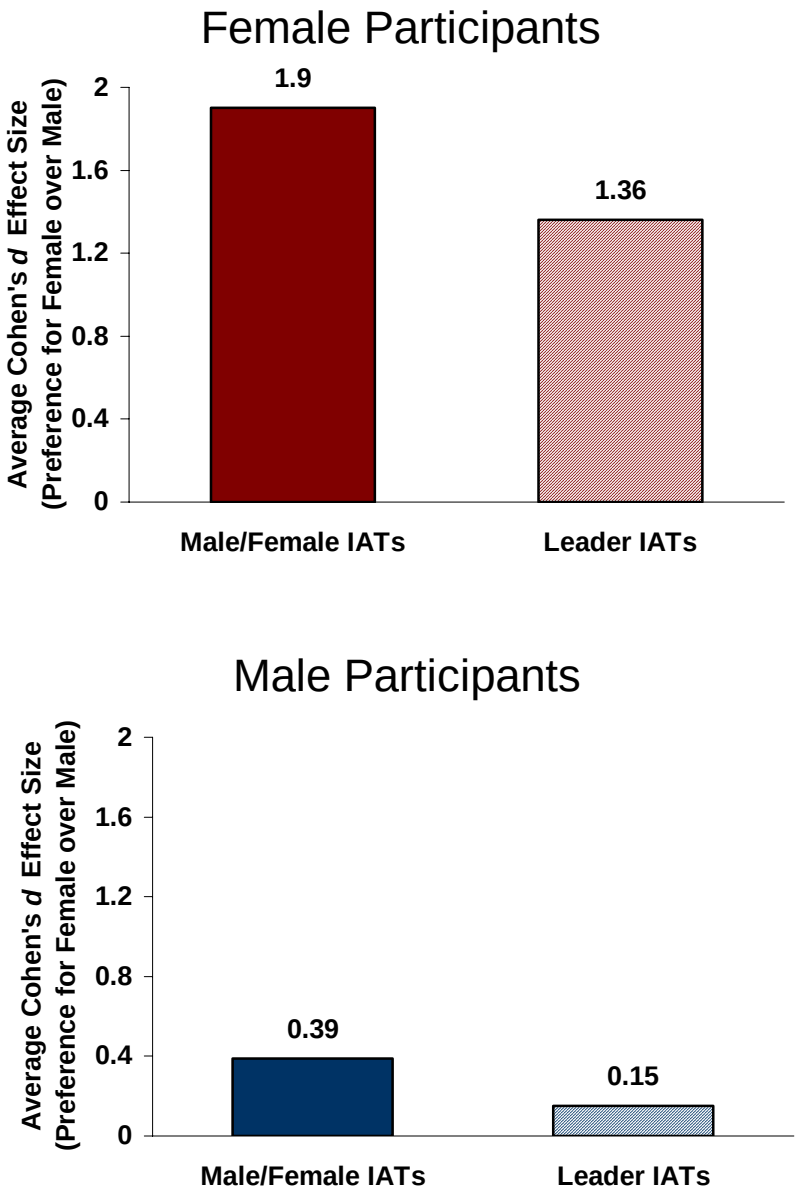
Summary of Results

The preliminary experiments reveal a striking pattern of results (see Figure 1). Across four experiments, female participants showed consistently strong implicit preferences for female over male, regardless of the context in which gender evaluations were made. That is, female participants showed an implicit preference for mothers over fathers, for women over men, for specific female leaders over male leaders, and for the generic category of female leaders over male leaders. What is more, women's implicit preferences were consistent in valence, revealing strong favorability toward the category female (relative to male) across the experiments; the mean effect sizes (Cohen's *d*) for the Male/Female IATs and the Leader IATs were 1.90 and 1.36, respectively (see Figure 2).

Like female participants, male participants exhibited positive implicit attitudes toward mothers relative to fathers and toward the general category women relative to the category men; however, male participants' implicit preferences for female over male were considerably smaller in magnitude than were those of female participants (mean Cohen's $d = 0.39$; see Figure 2). Moreover, measurements of men's attitudes toward female leaders versus male leaders revealed an inconsistent pattern of responding. In PE1, male participants showed a small implicit preference for male leaders over female leaders; in PE3a and PE3b, they showed no appreciable implicit preference for either male leaders or female leaders; and in PE2 and PE4, males demonstrated a moderate implicit preference for female leaders over male leaders (across all experiments assessing attitudes toward leaders, male participants' mean $d = 0.15$; see Figure 2). Overall, then,

men showed (a) no strong preference for either gender group, and (b) less valence-consistency in their implicit gender attitudes than did women (i.e., fluctuating between positive and negative implicit attitudes).

Figure 2
Average Effect Sizes for PE1-PE4



Implications

Overall implicit preference for female over male. The results of the preliminary experiments support reports that individuals indeed hold positive attitudes toward women (Carpenter & Banaji, 1997; Eagly & Mladinic, 1989; Eagly et al., 1991). The present pattern, like Eagly and her colleagues' work, contradicts the literature's earlier consensus that negative attitudes toward women are prevalent. This consensus may have emerged from a confusion between the vast discrimination that women face in political, social, and economic arenas, on one hand, and the favorability of attitudes and stereotypes concerning women, on the other hand. These new experiments are the first to replicate Eagly and colleagues' experiments using implicit measures, and they demonstrate that the observed positive attitude toward women extends to implicit attitudes as well.

Effect of group membership on implicit gender attitudes. The present data, showing positivity toward women, are qualified by a strong participant sex difference—a pattern that Eagly and her colleagues did not reliably observe. The average magnitude of female participants' implicit preference for female over male, across several preliminary experiments, was much larger than the magnitude of male participants' preference, indicating that participants' group membership (i.e., their sex) was an important aspect of their gender attitudes. This finding is congruent with several theoretical positions on group membership, self-concept, and attitudes, including social identity theory (Tajfel, 1981; Tajfel & Turner, 1981), self categorization theory (Turner, 1985; Turner et al., 1987), and the unified theory of implicit social cognition (Greenwald et al., 2000a, 2000b). The fact that men did not show strong, consistent implicit liking for

their own group may reflect an interaction between the influences of group membership and culturally shared evaluations of women and men, which pull for greater liking for women than for men.

An alternative possibility is that in-group bias was not responsible for the participant sex differences observed in the preliminary experiments. Rather, it may be that a more general mechanism was at work. For instance, perhaps women are likely to show greater liking than are men for all things feminine or "nice." The present research, which only assessed implicit evaluations of gender groups, cannot rule out this possibility.

Consistency in attitudes across construals of women and men. Another notable outcome of the preliminary experiments concerns the valence-consistency of implicit attitudes toward the different gender construals presented. For female participants, implicit evaluations of women, relative to men, were favorable across all construals of women presented (i.e., women, mothers, generic female leaders, specific exemplars of female leaders). For male participants, there was less valence-consistency, with men showing weakly favorable evaluations of construals of women that are positively valued by the culture (i.e., women, mothers) but less consistent favorability toward a construal that is known to be less culturally valued (i.e., female leaders). Thus, the results of the preliminary experiments suggest that group membership and cultural construal may combine to influence implicit gender attitudes.

The finding of greater valence-consistency in gender attitudes among women than among men has not been shown previously, and it raises many questions. How are individuals' implicit attitudes toward women related to their explicit preferences? Are

attitudes that are more consistent across varying construals of the same attitude object more closely linked to explicit preferences than those that are less consistent? How can implicit attitudes be changed? Two subsequent experiments turn to these questions and also address the relationship between implicit gender attitudes and implicit beliefs about the attributes of men and women.

Experiment 1

As the prevalence of implicit social cognition has become increasingly apparent, the relationship between implicit attitudes and beliefs and explicitly endorsed judgments has attracted growing attention (Brauer, Wasel, & Niedenthal, 2000; Dunton & Fazio, 1997; Karpinski & Hilton, in press; Rudman & Kilianski, in press; for a review, see Blair, in press). Theoretically, models of implicit and explicit attitudes, like established models of implicit and explicit memory, have been conceptualized as distinct processes (see Greenwald & Banaji, 1995). Indeed, some investigations suggested that the link between implicit and explicit processes was at best weak (e.g., Banaji & Hardin, 1996; Devine, 1989). More recent evidence, however, has suggested stronger relation between the two (e.g., Cunningham et al., 2000, in press; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1995; Kawakami, Dion, & Dovidio, 1998; Lepore & Brown, 1997; Wittenbrink et al., 1997).

It has been suggested that some measures, statistical methods, or individuals are more likely to show an association between implicit and explicit attitudes toward a given social group than are others. For example, Fazio et al. (1995) proposed that individuals' motivation to control prejudiced responding mediates the strength of the link between implicit and explicit attitudes. They found that individuals who scored low on a measure

of motivation to respond without prejudice to Blacks demonstrated a stronger link between implicit and explicit attitudes toward Blacks than did participants who scored high in motivation to control prejudice.

In another vein, Kawakami et al. (1998) observed that even in research in which implicit and explicit measures are associated, the association is relatively weak. They proposed that extremely sensitive procedures may be necessary to pick up relationships between implicit stereotyping and explicit beliefs and attitudes. In a recent analysis of implicit and explicit ethnocentrism (i.e., a constellation of biased attitudes toward multiple social groups), Cunningham et al. (in press) observed that, when random measurement error was statistically controlled, implicit and explicit attitudes were modestly related—an effect that remained hidden without accounting for such error.

The present research is the first to examine both implicit and explicit attitudes toward different construals of gender. The findings of the preliminary experiments, that both male and female participants showed implicit favorability toward women, relative to men, raised the question of how such implicit preferences are linked to explicit judgments. This issue was explored in Experiment 1.

Method

Overview

To provide a test of the extent to which implicit gender attitudes are related to explicit judgments, Experiment 1 presented participants with brief descriptions of fictional male and female political candidates. Following this presentation, participants expressed a candidate preference (by indicating which of the two candidates they would vote for in an upcoming election). Additionally, they made explicit judgments about each

of the candidates. Finally, participants completed IAT measures of their implicit attitudes toward the specific candidates and toward generic male leaders and female leaders.

Although previous research has measured implicit attitudes toward known social groups (e.g., Black-White, male-female, Japanese-Korean, etc.), Experiment 1 represents a first effort to compare implicit attitudes toward a known social group with evaluations of specific (artificially created) exemplars of that group. Thus, one aim of Experiment 1 was to compare implicit attitudes toward specific political candidates with implicit attitudes toward generic male leaders and female leaders. A second goal of Experiment 1 was to examine whether implicit attitudes predict candidate preference and explicit judgments better for members of groups who show greater valence-consistency in their evaluations of varying representations of the same general social group, as female participants were shown to do in the preliminary experiments (see Figure 1).

Several tentative predictions are offered for Experiment 1. First, it is predicted that, as in the preliminary experiments, female participants will show greater implicit positivity toward female leaders than will male participants. It is expected that this pattern will be especially pronounced when attitudes toward male and female leaders as a generic social group are assessed, versus when attitudes toward particular exemplars of the group are assessed. Second, it is predicted that neither male nor female participants will exhibit an explicit bias against female leaders relative to male leaders. Third, it is predicted that both explicit and implicit attitudes will be correlated with with the measure of candidate preference. Finally, it is tentatively hypothesized that the link between implicit attitudes and explicit preferences will be stronger for female participants, who

demonstrated greater valence-consistency in the preliminary experiments and thus may also be expected to show greater attitude consistency of other varieties.

Participants

The participants were 70 male and 59 female undergraduates who participated in the experiment for course credit.[8]

Materials and Procedure

After reading and signing an informed consent form, participants were seated at a computer to begin the experiment.

Candidate descriptions. Participants were instructed that they would be presented with biographical descriptions of two fictional political candidates for an upcoming election. Because participants would receive little information upon which to base their voting decisions, there was a risk that they would be unwilling to make a meaningful voting choice. Previous research has indicated that people feel more entitled to rely on their stereotypes in making social judgments when those stereotypes are not perceived as being the main source of information (Yzerbyt, Leyens, & Schadron, 1997; Yzerbyt, Schadron, Leyens, & Rocher, 1994). Thus, in Experiment 1 a cover story was constructed in order to heighten participants' sense of the candidates' *social judgeability*. As part of this cover story, participants were instructed that in addition to the descriptions of the candidates, further relevant information about the candidates would be flashed on the computer screen so quickly that they would not be able to consciously perceive it.

[8] The data reported here as Experiment 1 were collected in two separate, consecutive data collections. The two sets of data are combined here because there are only minor differences in the materials used for the two collections. In addition to the explicit attitude measures described in this report, two additional explicit, paper-and-pencil attitude measures were used, but each was used in only one of the two data collections. These two measures were not developed with as much care as were the other measures used in the present research, and their results were not clear. Because each was used for only about half of the participants, they are not discussed further in this report.

However, participants were told, they would be able to unconsciously perceive such information, and that information would help them form impressions of the candidates, which they could use in their voting decisions.

Next, the descriptions of one male and one female candidate were presented to participants on the computer screen, one sentence at a time (see Appendix C). For half the participants, the candidates' names were *Brian Reilly* and *Karen Nichols*; for the other half of the participants, the candidates' names were *Gary Nichols* and *Lisa Reilly*. Each set of male and female first names had been previously rated as equally likable (Kasof, 1993). Each sentence of the descriptions was presented for 5.5 s. To support the cover story, after each sentence disappeared from the screen, a string of nonsense syllables (e.g., *jru op evg uqw idf ar*) was presented for 16 ms. This presentation was long enough for participants to see a flicker on the screen, but not long enough for them to identify the letters.

The candidate descriptions were equal in length. Each description contained information about the candidates' political backgrounds, the issues they had focused on in their political careers, and how they were regarded by their colleagues. The order in which the two candidates were described and the assignment of each of the two descriptions to the male or the female candidate were counterbalanced between subjects.

Candidate preference. After reading the descriptions of both candidates, participants expressed a candidate preference, by indicating on the computer which candidate they would vote for in an upcoming election.

Explicit judgments. Next, participants responded to several explicit items presented on the computer, using a 1-7 scale to rate each candidate's intelligence,

strength, compassion, cooperativeness, overall favorability, and likelihood of being successful as a leader, if elected.

Implicit attitude measures.[9] After completing the explicit attitude measures, participants completed IAT measures of their implicit attitudes toward the specific candidates and toward male and female leaders as a generic social category. The IAT procedures in Experiment 1 were similar to those in the preliminary experiments, with several modifications, as described below. The evaluative stimuli were also similar to those used in the preliminary experiments (see Appendix A). The IATs used Micro Experimental Laboratory (MEL) software, version 2.0 (Schneider, 1990).

The IAT concerning the specific male and female candidates (Candidate IAT) measured participants' strength of association between each candidate and favorable or unfavorable evaluation. On this measure, participants were presented with the candidates' first names, last names, or first and last names (e.g., *Lisa, Reilly, Lisa Reilly; Gary, Nichols, Gary Nichols*). Participants' task was to classify each stimulus as representing one of the two candidates (e.g., *Lisa Reilly, Gary Nichols*). Thus, the task on the Candidate IAT involved making implicit judgments of the specific political candidates, rather than of the generic category, leaders.

The Candidate IAT consisted of two blocks of trials, each including 20 practice trials and 40 data collection trials. Participants were presented with evaluative words and with the names of the male and female candidates about whom they had just read. Their

[9] All participants completed the explicit measures, including indicating which candidate they would vote for in an upcoming election, before completing the measure of implicit attitudes. This fixed order was established because it was viewed as important that the explicit measures, over which participants had conscious control, not be contaminated by participants having first completed the implicit attitude measures.

task was to categorize the evaluative words (as pleasant and unpleasant) and the names (as belonging to one or the other candidate) by pressing either the left key (the 'A' key) or the right key (the '5' key on the numeric keypad) with their left and right index fingers. Throughout the trials, a reminder was positioned at the top of the screen, indicating which responses were to be made with the left key and which were to be made with the right. The participants were instructed to classify each word as quickly as possible, avoiding mistakes. When participants made incorrect responses, the computer emitted a short (250 ms) tone to provide them with error feedback. For half the trials, the female candidate name was paired on the same key with pleasant words, and the male candidate name was paired on the same key with unpleasant words (abbreviated as the female candidate+pleasant configuration). For the other half of the trials (abbreviated as the female candidate+unpleasant configuration), these pairings were reversed. The order in which these two key configurations were presented was counterbalanced between subjects.

After completing the Candidate IAT, participants next completed a measure of implicit attitudes toward the generic categories of *male leaders* and *female leaders* (Generic Leader IAT). This task was identical to the Candidate IAT except that male leaders and female leaders were represented not by the names of the specific candidates about whom participants had read, but by generic male and female first names accompanied by the label MALE LEADER or FEMALE LEADER, as in the preliminary experiments. Participants were presented with stimuli in two different key configurations (abbreviated as the female leader+pleasant and the female leader+unpleasant configurations). Their task was to categorize each name as a male leader name or a

female leader name, and to categorize each evaluative word as a pleasant or unpleasant word. The order in which the two key configurations were presented was counterbalanced between subjects.

Finally, participants completed a brief demographic questionnaire and were thoroughly debriefed. The entire procedure took approximately one hour.

Results and Discussion

Candidate Preference

When asked which of the two candidates they would vote for in an upcoming election, 46% of male participants indicated a preference for the female candidate and 54% indicated a preference for the male candidate. Among female participants, 56% indicated a preference for the female candidate and 44% indicated a preference for the male candidate. Although this pattern suggests a pattern of slight bias in favor of candidates whose gender matches participants', the Participant Sex x Candidate Sex interaction was not significant, $\chi^2 (1, 127) = 1.35$, ns.

Explicit Judgments

After reading the candidate descriptions and indicating which of the two candidates they would vote for in an upcoming election, participants responded to several items measuring their explicit attitudes and beliefs about the particular candidates. To create bias scores for participants' explicit ratings of how favorably they felt toward each candidate, how successful they believed each candidate would be if elected, and how intelligent, strong, cooperative, and compassionate each candidate was, each participant's rating of the male candidate for each item was subtracted from his or her rating of the female candidate for that item. Thus, higher scores reflect a greater bias in favor of the

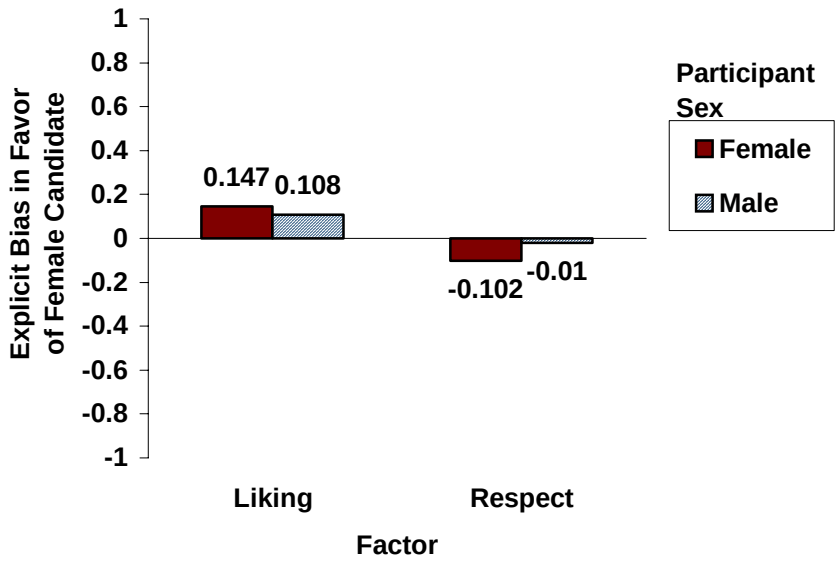
female candidate over the male candidate. Next, participants' preference scores on each trait were entered into a principal components factor analysis. An examination of the scree plot, eigenvalues, and factor loadings supported a two-factor, varimax rotated structure. The first factor, labeled *Liking*, included the items cooperative, compassionate, and favorable. The second factor, labeled *Respect*, included the items strong, intelligent, and successful.

To test whether participants showed an explicit preference for one candidate over the other, individual t-tests were performed. As shown in Figure 3, these analyses showed that on the Liking factor, no preference for either the male candidate or the female candidate was shown either among male participants, t (68) = .57, ns, or among female participants, t (59) = .60, ns. Likewise, on the Respect factor, no preference was found either among male participants, t (68) = .07, ns, or among female participants, t (59) = .55, ns. Additionally, separate one-way ANOVAs on each of the two factors revealed no participant sex effects on either Favorability, F (1, 126) = 0.02, ns, or Respect, F (1, 126) = 0.16, ns.

The results of the explicit attitude measure used in Experiment 1 should be interpreted with caution. It is possible that the social judgeability aspect of the experiment—in which participants were led to believe that they were exposed to subliminal information about the candidates that could help them in their judgments—introduced some reactivity into the experiment. This manipulation may have heightened participants' self-attentiveness, biasing their responses. Notwithstanding this possibility, the results of the explicit measures used in Experiment 1 provide initial evidence that

neither males nor females are inclined to express negative attitudes or beliefs about particular female leaders.

Figure 3
Explicit Bias in Favor of Female Candidate



Implicit Attitudes

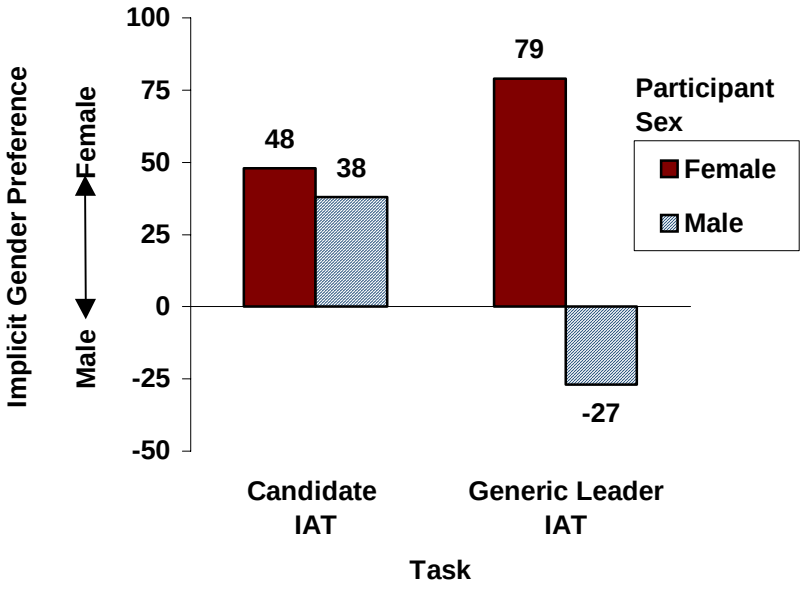
Each participant's IAT effect score was computed by subtracting the mean response latency in the female candidate+pleasant configuration (or, in the Generic Leader IAT, the female leader+pleasant configuration) from the mean latency in the female candidate +unpleasant (or female leader+unpleasant) configuration. Thus, the IAT effect reflects the extent to which participants showed an implicit preference for the female candidate over the male candidate (or for female leaders over male leaders).

Candidate IAT. On the Candidate IAT, both male and female participants showed an overall implicit preference for the specific female candidate over the specific male candidate, responding faster in the female candidate+pleasant configuration than in the female candidate+unpleasant configuration, F (1, 127) = 32.62, $p < .0001$ (see Figure 4).

The effect of an implicit preference for the female candidate over the male candidate was significant both for female participants, F (1, 58) = 19.77, $p < .0001$, $d = 0.58$, and for male participants, F (1, 69) = 13.32, $p < .001$, $d = 0.44$. The participant sex effect was not significant, F (1, 127) = 0.45, ns.

The finding of an implicit preference for the female candidate over the male candidate, among both male and female participants, parallels previous research on explicit gender attitudes, which found evidence of positivity toward some construals of women (Carpenter & Banaji, 1997; Eagly & Mladinic, 1989; Eagly, Mladinic, & Otto, 1991). Further, the results are consistent with those of the preliminary experiments, which demonstrated strongly favorable implicit attitudes toward women among female participants, and among male participants, to a lesser degree and with less consistency across experiments.

Figure 4
IAT Effects (in Milliseconds) for Candidate and Generic Leader IATs



Generic Leader IAT. The Generic Leader IAT measured implicit attitudes toward the general categories of *female leaders* and *male leaders*. Consistent with the results of the preliminary experiments, female participants showed more favorable implicit attitudes toward female leaders than did male participants; the participant sex effect was significant, F (1, 125) = 35.87, $p < .0001$ (see Figure 4). Female participants demonstrated an implicit preference for female leaders over male leaders, F (1, 57) = 35.27, $p < .0001$, $d = 0.79$. In contrast, male participants demonstrated an implicit preference for male leaders over female leaders, F (1, 68) = 5.35, $p < .05$, $d = 0.28$.

Overall, the results of the Candidate and Generic Leader IATs revealed a pattern similar to the results of the preliminary experiments. Female participants showed favorable implicit attitudes both toward generic female leaders (relative to male leaders) and toward the specific female political candidate (relative to the male candidate); that is, their implicit attitudes toward women and men were consistent in valence across the two different construals of women. In contrast, male participants showed an inconsistent pattern of responding, demonstrating negative implicit attitudes toward female leaders as a generic social category (relative to male leaders), but positive implicit attitudes toward the specific female candidate (relative to the male candidate). As in the preliminary experiments, then, male participants' implicit gender attitudes were less consistent in valence than were female participants' attitudes. As discussed above, the sources of this differential valence-consistency are not known, nor are its consequences.

Relations Between Implicit and Explicit Attitudes

Correlations between implicit attitudes and explicit favorability and respect judgments. To test the association between implicit and explicit judgments of the male

and female candidates, zero-order correlation analyses were performed for both male and female participants. For male participants, implicit attitudes toward the candidates were significantly correlated with the explicit Liking factor, r (68) = .35, $p < .01$, but not with the explicit Respect factor, r (68) = .20, ns. For female participants, implicit attitudes toward the candidates were significantly correlated both with the explicit Liking factor, r (57) = .58, $p < .0001$, and with the explicit Respect factor, r (57) = .46, $p < .001$. Thus, the results of Experiment 1 indicated that male participants showed an implicit preference for the female candidate over the male candidate to the extent that they viewed her as cooperative, compassionate, and generally favorable. Female participants showed implicit liking for the female candidate to the extent that they evaluated her as not only cooperative, compassionate, and generally favorable, but also to the extent that they viewed her as intelligent, strong, and likely to succeed as a leader if elected.

Correlation between implicit attitudes and candidate preference. The finding that participants' scores on the explicit measures in Experiment 1 were correlated with their implicit attitudes toward the male and female candidates provides the opportunity to test the resilience of the relation between implicit attitudes and another explicit preference measure—namely, participants' choice of which candidate they would vote for in a forthcoming election. Toward this end, analyses of the zero-order correlation between participants' implicit attitudes and candidate preferences were conducted. These analyses revealed that candidate preference was not significantly associated with implicit attitudes toward the generic categories of male and female leaders for male participants, r (67) = .09, or for female participants, r (56) = .06. Moreover, for male participants, implicit attitudes toward the *specific candidates* were likewise unrelated to candidate preference, r

(68) = .17. For female participants, however, implicit attitudes toward the specific candidates were significantly correlated with candidate preference, r (57) = .58, $p <$ .0001. Tests of the differences between the zero-order correlations revealed that this last correlation (between female participants' expressed candidate preference and their implicit attitudes toward the specific candidates) was significantly different from (a) female participants' candidate preference and their implicit attitudes toward generic male and female leaders, $z = 3.10$, $p < .001$, (b) male participants' candidate preference and their implicit attitudes toward generic male and female leaders, $z = 3.09$, $p < .001$, and (c) male participants' candidate preference and their implicit attitudes toward the specific candidates, $z = 2.73$, $p < .01$.

A hierarchical regression analysis was conducted to test whether for female participants, the relation between implicit attitudes toward the particular candidates and expressed candidate preference remained strong even after explicit judgments were accounted for.[10] The regression analysis showed that female participants' implicit attitudes toward the particular candidates continued to significantly predict their candidate preference even after statistically controlling for participants' explicit Liking and Respect scores (see Table 2).

The findings in Experiment 1 are interesting on several counts. First, the finding that women's implicit attitudes toward the specific leaders predicted their expressed likelihood of voting for either the male or the female candidate, even though their attitudes toward generic male and female leaders did not, suggests that the correspondence between implicit attitudes and explicit preferences—and perhaps by

[10] Because the zero-order correlation between implicit attitudes and voting behavior was not significant for male participants, their data were not submitted to regression analysis.

extension, behavior—is likely to be strongest when the attitudes measured are similar in specificity. This notion is captured (in behavioral terms) by Ajzen and Fishbein's compatibility principle (1977). Although the compatibility principle was posited in reference to attitudes that are measured explicitly, the present results suggest that it may also be brought bear on implicit social cognition.

Table 2

Summary of Hierarchical Regression Analysis for Variables Predicting Candidate Preference, for Female Participants

Variable	*B*	*SE B*	β	*T* for H0	*P*
Step 1					
Liking	.131	.036	.487	3.61	.001
Respect	.080	.047	.227	1.68	.099
Step 2					
Liking	.092	.037	.343	2.428	.019
Respect	.063	.046	.181	1.389	.171
Candidate IAT	2.857	1.144	.302	2.497	.016

Note. R^2 = .43 for Step 1; R^2 = .49 for Step 2.

Second, it is significant that female participants' implicit attitudes predicted their explicit candidate preferences, but male participants' did not. As demonstrated in the preliminary experiments (see Figure 1), and corroborated by the results of the Candidate and Generic Leader IATs in Experiment 1, women's implicit gender attitudes tend to be more consistent across varying construals of women and men than do male participants' attitudes. The current finding that women's implicit attitudes toward the specific female and male candidates predicted their candidate preferences, whereas male participants'

implicit attitudes did not, suggests that this particular type of structural consistency may be an important element of implicit attitudes.

Finally, Experiment 1 demonstrated that implicit attitudes predicted candidate preference even after taking into account participants' explicit judgments of the candidates' overall favorability, cooperativeness, compassion, likelihood of succeeding if elected, intelligence, and strength—measures that would seem to closely resemble one's choice of which candidate to vote for. This finding further bolsters the idea that the implicit gender evaluation assessed in the present research is a meaningful psychological construct, distinct from consciously expressed attitudes and beliefs. Additionally, although the present research did not test ecologically valid behavior, these results hint that implicit attitudes may have value for predicting action.

Experiment 2

Malleability in Implicit Attitudes and Beliefs

As evidence has accrued that implicit judgments are nearly ubiquitous—if sometimes insidious—features of social cognition, the question of how readily implicit attitudes and beliefs can be changed has garnered increasing attention (Blair et al., 2000; Dasgupta & Greenwald, 2000). It is not surprising that investigators should turn to such questions; indeed, the much longer tradition of research on prejudice and stereotyping processes that operate consciously has been marked by efforts to understand the circumstances under which these often destructive mental processes can be thwarted.

A prevailing view of implicit attitudes and beliefs is that they are slow in forming, forged gradually through experience and learning in the social world. By extension, it is believed, implicit attitudes and beliefs must be more stable and more resistant to change

than are their explicit counterparts (see Gregg, 2000; Smith & DeCoster, 1999). In support of this claim, several investigations have indeed suggested that conscious efforts to suppress or override implicit attitudes and stereotypes can easily backfire (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994).

However, several recent studies—most of which are currently unpublished—have suggested that implicit attitudes and beliefs may be more malleable than had previously been expected. This research has indicated that implicit associations, although resistant to conscious attacks, may be susceptible to more subtle avenues of change. Especially promising are methods that do not rely on the social perceiver's conscious willingness and ability to effect change, but instead attempt to alter the accessibility of stereotypes or negative attitudes in memory without participants' conscious awareness (Blair & Banaji, 1996; Blair et al., 2000; Dasgupta & Greenwald, 2000; Goodwin & Banaji, 2000; Lowery, Hardin, & Sinclair, 1999; Macrae et al., 1997).

In a recent examination of implicit race attitudes, for example, Dasgupta and Greenwald (2000) found that presenting participants with exemplars of admired Black individuals (e.g., *Martin Luther King, Jr.*, *Denzel Washington*) and disliked White individuals (e.g., *Jeffrey Dahmer*, *Ted Kaczynski*) was able to reduce implicit anti-Black attitudes—an effect that remained detectable for at least 24 hours. In another experiment (Experiment 3), Dasgupta and Greenwald replicated this finding in the domain of attitudes toward the elderly.

Like implicit attitudes, implicit stereotypic beliefs have also shown surprising malleability. Blair et al. (2000) recently used a mental imagery task to make counterstereotypes about gender more accessible in memory. In this priming intervention,

they instructed participants to imagine and write brief essays about strong women. On a subsequent IAT measure of implicit stereotyping, Blair et al. found that participants who were in the counterstereotypical imagery condition showed significantly less implicit stereotyping (i.e., less of an association between the concepts *female* and *weak*) than did participants in gender-stereotypic imagery, neutral, and no-prime conditions.

Do Implicit Attitudes and Beliefs Function Independently?

The discovery that implicit attitudes and beliefs may each be malleable in lawful ways raises the related question of whether the two constructs are conceptually distinct or are sufficiently intertwined that changes in one will induce changes in the other. Decades of research that has attempted to clarify the extent to which attitudes and beliefs regarding social groups operate independently has met with mixed results (see Glaser, 1999). Although some have argued that attitudes toward social groups (i.e., prejudice) are simply the "hot" offshoots of the "cold" categorization processes that produce stereotypic beliefs, others have maintained that attitudes and beliefs—including those that operate outside conscious awareness—function independently.

Greenwald and Banaji (1995) defined implicit attitudes as "…introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects." Distinguishing implicit attitudes from their cognitive counterparts, Greenwald and Banaji defined implicit stereotypic beliefs as "…introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category." The distinction that Greenwald and Banaji drew between implicit attitudes and belifs parallels a more longstanding distinction between evaluation, or attitudes, and cognition,

with the former involving a global assessment of the favorability or unfavorability of an attitude object, and the latter involving assignment of specific attributes—most of which undoubtedly also carry evaluative meaning—to that object.

In an examination of the independence of implicit attitudes and beliefs, Glaser (1999) produced biases in the implicit attitudes toward and beliefs about novel (i.e., fictional) social groups and experimentally demonstrated that beliefs and attitudes concerning the novel groups operated independently.[11] O'Connor, Cunningham, Banaji, Gore, Gatenby, & Phelps (2000) further bolstered the argument for viewing implicit attitudes and beliefs as distinct systems. While in an fMRI scanner, participants completed IAT measures that paired Black and White faces either with evaluative (good/bad) words or with words related to semantic beliefs (scholar/athlete). O'Connor et al. found the expected pattern of biased implicit attitudes and beliefs regarding Blacks versus Whites. Importantly, however, the fMRI data also revealed different patterns of brain activation in the anterior cingulate—a part of the brain pathway known to be involved in response competition—depending on whether participants were making evaluative (attitude) or semantic (belief) judgments. This evidence of apparently separate brain pathways for attitudinal and belief judgments provides additional support for the position that implicit attitudes and beliefs may operate independently.

In the domain of gender, Goodwin and Banaji (2000) recently found further evidence that implicit attitudes and beliefs may be dissociable. Consistent with the preliminary experiments reported in the present research, they found that women and

[11] Interestingly, Glaser's (1999) demonstration of the separability of implicit attitudes and beliefs revealed an unexpected twist: Manipulations of attitudes and stereotypes produced a "cross-resonating" effect, with the two constructs influencing each other more than they influenced themselves.

men differed in their implicit gender attitudes, with female participants showing a stronger implicit preference for women over men than did male participants. However, male and female participants in Goodwin and Banaji's experiment shared similar implicit gender stereotypes, both showing stronger associations between the concepts *female* and *family* and the concepts *male* and *career* than the opposite. This finding is corroborated by other research in which male and female participants were equally biased in assigning the quality of fame more readily to males than to females (Banaji & Greenwald, 1995).

Experiment 2 addressed the question of the malleability of implicit gender-stereotypic beliefs and attitudes and examined whether changing implicit beliefs induces second-order changes in implicit gender attitudes, or vice versa. Understanding the extent to which these two processes are amenable to change and the extent to which they are dependent on each another will help resolve theoretical questions of how implicit attitudes are formed and how readily they are influenced by stimuli in the social world. Further, such an understanding is likely to be an important step in developing interventions to short-circuit stereotypical beliefs and negative attitudes toward women.

Method

Overview

In order to address the questions of whether implicit gender beliefs and attitudes can be altered and the extent to which changing one effects change on the other, Experiment 2 provided half the participants with different construals of women, using a mental imagery priming intervention similar to that used by Blair et al. (2000). The other half of the participants were assigned to a neutral imagery control condition. The goal of the imagery task was to heighten the accessibility of a particular construal of women, as

strong leaders. Following this priming intervention, participants completed two IATs, one measuring implicit attitudes toward women and men (Good/Bad IAT), and the other measuring implicit beliefs about the relative strength and weakness of males and females (Strong/Weak IAT).

Participants

Participants in Experiment 2 were 59 male and 59 female undergraduates who participated in the experiment for course credit or for monetary compensation ($7).

Materials

Prime. Participants were randomly assigned to one of two different prime conditions (see Appendix D). In the strong female leader prime condition, participants were given a one-page form that requested that they spend several minutes imagining what women who are tough and aggressive leaders are like, and to write down their thoughts about the characteristics, behaviors, and decision-making strategies that illustrate these women's power and toughness. In the neutral prime condition, participants were instructed to imagine and write about the noteworthy sights on a tour of the university campus.

Implicit Association Tests. The two IATs—one measuring implicit liking for women and men (Good/Bad IAT) and one measuring implicit beliefs about the strength versus weakness of women and men (Strong/Weak IAT)—were conducted on Pentium processor computers using Inquisit software (Draine, 1998). In both IATs, the gender stimuli were words that denote either male (e.g., *He, Him, Male*) or female (e.g., *She, Her, Female*). In the Good/Bad IAT, participants were presented with evaluatively positive (e.g., *Laughter, Excellent, Vacation*) and negative (e.g., *Cancer, Awful, Maggot*)

words. In the Strong/Weak IAT, participants were presented with words that connote strength (e.g., *Bold, Powerful, Mighty*) or weakness (e.g., *Tender, Delicate, Dainty*). The order in which participants completed the Good/Bad and Strong/Weak IATs was counterbalanced between subjects.

Procedure

After reading and signing an informed consent form, participants began the experiment.

Priming. A female experimenter gave participants the priming intervention (i.e., strong female leader or neutral prime) and set a timer for five minutes. She instructed participants to turn to the computer when the timer sounded and to follow the instructions on the screen to continue with the experiment.

Good/Bad IAT. Each participant made implicit evaluative judgments concerning the categories *male* and *female*. All instructions were presented on the computer. In the Good/Bad (attitude) IAT, evaluative and gender-denoting words appeared on the computer screen one at a time, and participants categorized each word by pressing either the left key (the 'A' key) or the right key (the '5' key on the numeric keypad) with their left and right index fingers. Throughout the trials, a reminder was positioned at the top of the screen indicating which responses were to be made with the left key and which were to be made with the right. Participants were instructed to classify each word as quickly as possible, avoiding mistakes. When participants made an incorrect response, a red *X* appeared on the computer screen as error feedback. In order to continue to the next trial, participants were required to enter the correct response for that trial. Stimuli were

presented randomly, with the constraint that evaluative and gender-denoting words were presented in alternating trials.

The Good/Bad IAT consisted of two blocks of trials, each including 10 practice trials and 30 data collection trials. In one block of trials, participants' task was to press the left key when a *male* or *good* word appeared on the screen, and to press the right key when a *female* or *bad* word appeared on the screen. This block of trials is abbreviated as the female+bad configuration. In the second block of trials, participants were instructed to press the left key when a *male* or *bad* word appeared on the screen, and to press the right key when a *female* or *good* word appeared on the screen. This block of trials is abbreviated as the female+good configuration. The order in which the two key configurations were presented was counterbalanced between subjects.

The reversal of the keys representing good and bad words allowed measurement of the relative ease with which individuals performed the categorization judgments in one key configuration versus the other, thus providing a within-subjects measure of the strength of association between the concepts of *male* and *female* with evaluative meaning (good/bad).

Strong/Weak IAT. The Strong/Weak (belief) IAT was identical to the Good/Bad IAT except that words that connote strength or weakness were used instead of evaluatively good or bad words. In one block of trials (abbreviated as the female+weak configuration), participants' task was to press the left key when a *male* or *strong* word appeared on the screen, and to press the right key when a *female* or *weak* word appeared on the screen. In the other block of trials (abbreviated as the female+strong configuration), participants were instructed to press the left key when a *male* or *weak*

word appeared on the screen, and to press the right key when a *female* or *strong* word appeared on the screen. As in the Good/Bad IAT, the order in which the female+strong and female+weak key configurations were presented was counterbalanced between subjects. Again, the reversal of the keys representing strong and weak words provided a within-subjects measure of the strength of association between the concepts of *male* and *female* with the gender-stereotypic dimension of strength versus weakness.

After completing the priming intervention and the two IATs, participants completed a demographic questionnaire and were thoroughly debriefed. The entire procedure took approximately 30 minutes.

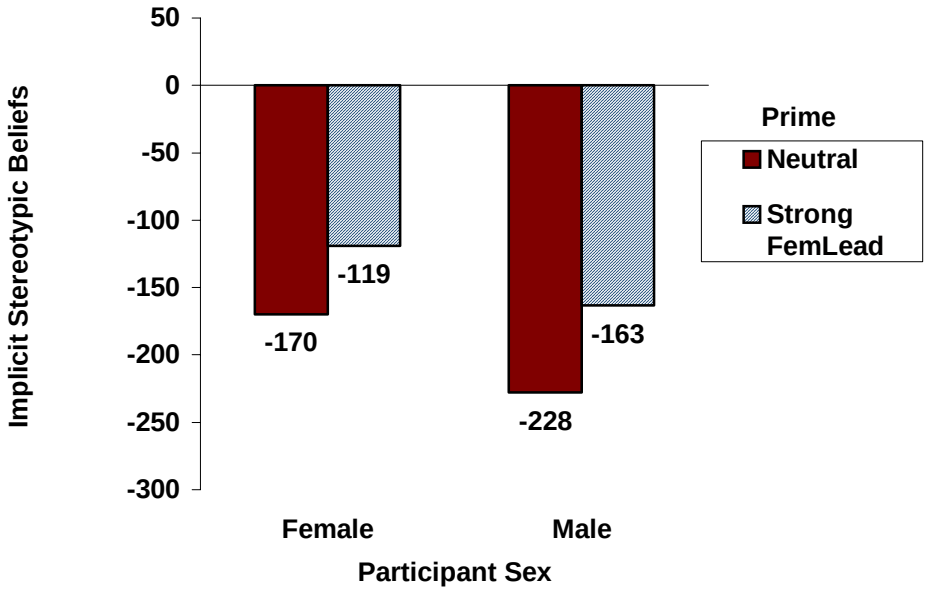
Results and Discussion

In Experiment 2, participants were primed either with a construal of women as strong leaders or with a neutral imagery task. Following this priming intervention, they completed IAT measures of (a) their implicit beliefs (i.e., stereotypes) about the relative strength versus weakness of women and men, and (b) their implicit attitudes toward women and men. It was predicted that, as in the preliminary experiments, participants would show favorable implicit attitudes toward women relative to men, and that the magnitude of this effect would be greater for female participants than for male participants. Further, it was predicted that male and female participants would both demonstrate evidence of implicit stereotypic beliefs about women, showing greater strength of association between the concepts *female* and *weak* and *male* and *strong* than the reverse pattern of association. Finally, Experiment 2 tested whether implicit beliefs and attitudes were susceptible to change through a priming intervention.

Strong/Weak IAT

On the Strong/Weak IAT, both male and female participants demonstrated a robust implicit belief that men are stronger than are women, F (1, 116) = 181.00, p < .0001 (see Figure 5). Congruent with previous evidence that men and women share similar implicit stereotypic beliefs about gender (Banaji & Greenwald, 1995; Blair et al., 2000; Goodwin & Banaji, 2000), the effect of participant sex on the Strong/Weak IAT was not statistically significant, F (1, 113) = 1.62, p < .21, d = 0.29, nor did participant sex interact with the prime to produce implicit stereotypic beliefs, F (1, 113) = .03, ns. Thus, it appears that stereotypic beliefs about the attributes of women and men are shared, at an implicit level, by men and women alike.

Figure 5
IAT Effects (in Milliseconds) for Strong/Weak IAT



A principal goal of Experiment 2 was to examine the extent to which implicit beliefs (and attitudes) were susceptible to change through a priming intervention. Consistent with Blair et al.'s (2000) findings, the Strong/Weak IAT revealed a significant

main effect of priming, F (1, 113) = 5.44, $p < .03$, $d = 0.44$. Participants who were primed with the construal of women as strong female leaders revealed a less robust implicit belief that women are weak (relative to men), F (1, 57) = 65.39, $p < .0001$, $d = 1.07$, compared with participants in the neutral priming condition, F (1, 58) = 126.85, $p < .0001$, $d = 1.48$. Finally, the order in which participants completed the Strong/Weak and Good/Bad IATs did not alter the effect of the priming intervention on implicit beliefs, F (1, 113) = .13, ns.

Good/Bad IAT

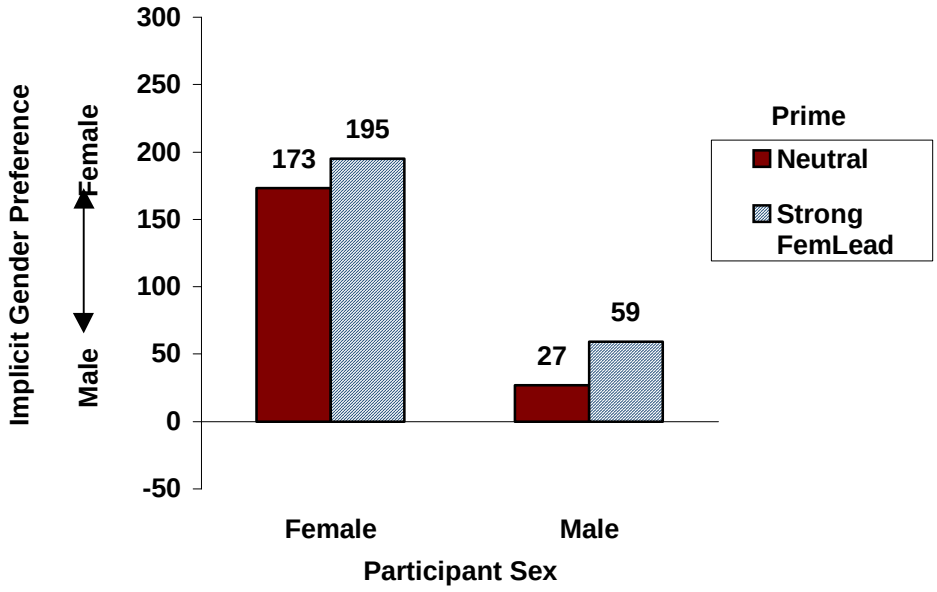
The Good/Bad IAT revealed a markedly different pattern of results from the Strong/Weak IAT (see Figure 6). Overall, participants showed a significant preference for women over men, F (1, 116) = 66.98, $p < .0001$. However, a significant main effect of participant sex was also observed, F (1, 115) = 39.73, $p < .0001$. Whereas female participants showed a very strong implicit preference for women over men, F (1, 57) = 121.32, $p < .0001$, $d = 1.46$, male participants' implicit preference for women over men was much smaller, F (1, 58) = 5.28, $p < .05$, $d = 0.30$. The order in which participants completed the Strong/Weak and Good/Bad IATs did not alter the effect of participant sex on implicit liking for women relative to men, F (1, 113) = .22, ns.

The effect of participant sex found for the Good/Bad IAT parallels the results of the preliminary experiments, in which a robust difference in men's and women's implicit gender attitudes was shown. Like the preliminary experiments, this experiment suggests that group membership plays an important role in shaping intergroup attitudes.[12]

[12] As part of the priming intervention, participants wrote either about their thoughts concerning women who are strong leaders or about a neutral topic. Content analyses of the essays concerning strong female leaders revealed overall positivity toward female leaders, with female participants' essays expressing greater positivity toward female leaders than did male participants. There was no correlation between essay

One of Experiment 2's primary goals was to test whether changes in implicit beliefs, induced through a priming intervention, would exert second-order effects on implicit gender attitudes. Analysis of the Good/Bad IAT data indicate that this was not the case; the main effect of prime was not significant for the Good/Bad IAT, F (1, 113) = 1.06, ns, nor was the Participant Sex x Prime interaction, F (1, 113) = 0.06, ns. It is possible that the lack of response to the priming intervention reflects a ceiling effect for female participants, whose baseline implicit attitudes reveal strong liking for the category female relative to male. However, this was not the case for male participants, yet they were equally unaffected by the priming intervention.

Figure 6
IAT Effects (in Milliseconds) for Good/Bad IAT



positivity and implicit beliefs, and there was a small positive correlation between essay positivity and implicit attitudes, r (57) = .27. Analysis of covariance showed that when essay positivity was partialled out of the analysis of implicit attitudes, the effect of participant sex on implicit attitudes remained significant, F (1, 55) = 12.10, $p < .001$.

Relation Between Implicit Attitudes and Beliefs

Overall, no correlation was observed between participants' implicit attitudes toward women and men and their implicit belief in the relative strength and weakness of women and men, r (115) = .06, ns. Further, separate correlational analyses were conducted for male and female participants and for participants in each of the two prime groups; none of these analyses revealed a significant relation between implicit attitude and belief.

Experiment 2 investigated the question of the stability versus malleability of implicit gender attitudes and beliefs. The results of the experiment indicate that presenting individuals with a counterstereotypical construal of women (i.e., as strong female leaders) induces change in the strength of association between the category *female* and the gender-stereotypic attribute of weakness. Moreover, this malleability was shown for both male and female participants, suggesting that implicit beliefs are not contingent upon group membership.

Despite the observed change in implicit gender beliefs, however, Experiment 2 did not demonstrate a change in implicit gender attitudes as a function of the construal of women that was made accessible via the priming intervention. Additionally, implicit attitudes and beliefs were uncorrelated, for both male and female participants and irrespective of priming condition, providing further evidence for the independent operation of implicit gender beliefs and attitudes. Thus, the results of Experiment 2, like Glaser's (1999) investigation of implicit attitudes and beliefs regarding novel groups, provide support for the view that the two components of implicit social cognition—attitude and belief—are conceptually and functionally distinct.

General Discussion

Throughout the past century, questions surrounding how attitudes toward social groups are formed, their structural consistency, their influence on behavior, and their malleability have held social psychologists' lasting attention. The present research addressed these questions within the framework of a particular class of attitudes: those that reside outside conscious awareness and control. As social psychological recognition of the prevalence of implicit attitudes has matured in recent years, a new set of questions surrounding their development and function has emerged. To gain fuller understanding of such implicit social cognition, this dissertation focused on four questions: What role does group membership play in determining attitudes toward social groups? Are effects of group membership moderated by cultural construals of the group? How are implicit attitudes related to explicit attitudes? Are implicit attitudes susceptible to intervention? The results of the present experiments yield some clear and some tentative answers to these questions and suggest directions for further exploration.

How Do Group Membership and Cultural Construal Shape Attitudes?

Four preliminary experiments in the present research examined implicit attitudes toward varying construals of women and men: mothers and fathers, female leaders and male leaders, specific exemplars of female leaders and male leaders, and the generic categories, women and men. The results of these experiments provide corroboration for earlier findings of explicit positivity toward women and further show that group membership (i.e., whether one is female or male) and cultural construal (i.e., the extent to which a particular representation of women is evaluated favorably by the culture) are important elements of implicit attitudes toward women and men.

The preliminary experiments revealed an overall implicit preference for women and men among both male and female participants, congruent with earlier findings of favorable attitudes toward women in research that used explicit measures of attitude, such as semantic differential and open-ended response measures (Carpenter & Banaji, 1997; Eagly & Mladinic, 1989; Eagly et al., 1991). However, this finding was qualified in each of the preliminary experiments by a participant sex effect, with female participants showing a far greater preference for the category *female* over the category *male* than did male participants. Thus, the results of the preliminary experiments indicate that, consistent with the predictions of social identity theory, self categorization theory, and the unified theory of implicit social cognition, group membership (i.e., one's own sex) plays an important role in shaping implicit attitudes toward women and men—perhaps more so than is the case for explicit attitudes.

The results of the preliminary experiments also indicate that cultural construals of women and men likewise help determine implicit gender attitudes. Previous research has demonstrated that women are liked to a greater extent, by both men and women, when they fulfill more traditional roles and personify more communal, warm, and nurturing characteristics than when they occupy more non-traditional roles (Eagly et al., 1991, 1992; Glick & Fiske, 1996). The preliminary experiments' comparison of implicit attitudes toward women versus men, mothers versus fathers, female leaders versus male leaders, and specific exemplars of female leaders and male leaders revealed that implicit gender attitudes also depend on cultural construal, for both male and female participants (see Figure 2). For male participants, implicit attitudes toward women when represented as mothers or women showed a larger preference (measured in effect size) than when

represented as leaders. For female participants, although overall preference for female over male was large in all experiments, we observed the same pattern, with female participants showing greater favorability toward mothers and women in general than toward female leaders.

These findings are congruent with other research that directly compared implicit attitudes toward the category *mothers* with attitudes toward the category *female leaders* (Carpenter, 1997). The results showed greater implicit liking for mothers than for female leaders among both male and female participants. In contrast, neither male nor female participants demonstrated an implicit preference for fathers over male leaders, or vice versa. Thus, there is growing evidence that, for both men and women, cultural construals of women help guide implicit attitudes.

Valence-Consistency of Attitudes Toward Women

The preliminary experiments also showed that female participants' implicit attitudes were more consistent in valence than were male participants' attitudes, across various representations of the categories female and male. That is, women showed a robust implicit preference for the category female, irrespective of the particular construal of women being evaluated (see Figure 1). In contrast, the valence of men's attitudes fluctuated, sometimes revealing an implicit preference for female, sometimes for male, and sometimes for neither gender group.

The source of this observed participant sex difference in valence-consistency is not clear; however, one potential explanation rests in the literatures on status, group identity, and attitude strength. For groups that are of lower status in social hierarchies (as women are), group identity tends to be more salient and to affect judgment and behavior

more than is the case for high-status groups (Abrams, Thomas, & Hogg, 1990; Gordon, 1968; Stangor, Carr, & Kiang, 1998; Steele, 1997; Steele & Aronson, 1995; Taylor, Fiske, Close, Anderson, & Ruderman, 1977; see Fiske, 1998 for a review). Because women's lower status confers heightened salience to their gender than is the case for men, gender identity is likely to be more chronically accessible for women than it is for men. Other research has indicated that attitudes tend to be more unwavering to the extent that they are more accessible, complex, and elaborated (Doll & Ajzen, 1992; Eagly & Chaiken, 1995, 1998; Fazio, 1989; Petty & Wegener, 1998; Powell & Fazio, 1984). Thus, heightened accessibility of women's gender identity may lend greater consistency to women's favorable in-group attitudes than is the case for men (see Mullen et al., 1992). Although this explanation is speculative, it suggests directions for further research. First, future experiments may manipulate the salience or accessibility of gender identity in both male and female participants and measure the consistency of attitude valence under conditions of greater or lesser identity salience or accessibility. Second, investigation of the development of gender identity and attitudes, from early childhood through adulthood, may help clarify the connection between gender identity salience and accessibility, on the one hand, and the gender attitudes' valence-consistency, on the other hand.

What Are the Consequences of Valence-Consistency in Implicit Attitudes?

The implicit-explicit attitude link. Whatever the reasons for men's and women's differential consistency in evaluation of women and men, the phenomenon itself is of interest, particularly if valence-consistency is a reliable predictor of explicit social judgments. The results of Experiment 1 provide initial evidence that this may in fact be

the case. In Experiment 1, participants were presented with descriptions of fictional male and female political candidates, cast a vote for one candidate or the other, and completed measures of their implicit and explicit attitudes. Replicating the results of the four preliminary experiments, female participants showed greater implicit liking for female leaders (relative to male leaders) than did male participants. However, both female and male participants showed an implicit preference for the female candidate over the male candidate. This finding is also consistent with the results of the preliminary experiments, which found that although women showed implicit favorability toward all construals of women (relative to men), male participants varied in their evaluation of these different construals.

Women's implicit attitudes toward the candidates were significantly correlated with two aspects of explicit judgment, namely Respect and Liking. In contrast, men's implicit attitudes were significantly correlated with only one dimension of explicit judgment: Liking. Importantly, Experiment 1 also revealed a stronger link between implicit attitudes and explicit candidate preferences among women than among men. That is, female participants, who were shown in the preliminary experiments to be highly consistent in their favorable evaluation of women relative to men, also showed greater consistency between their implicit attitudes toward a female political candidate and their selection of which candidate they would vote for in an upcoming election. In contrast, male participants, who showed valence-inconsistency in their gender attitudes in the preliminary experiments, also showed no relation between their implicit attitudes and their expressed candidate preference in Experiment 1.

Attitudes and behavior. Linked to questions concerning the relation between implicit and explicit attitudes is the question of how well implicit attitudes predict behavior. The nature of the link between attitudes and behavior is an issue with which social psychologists have wrestled for more than half a century (Ajzen & Fishbein, 1977, 1980; Eagly & Chaiken, 1993, 1998; Fazio, 1986, 1990; Fishbein & Ajzen, 1974; LaPiere, 1934; Millar & Tesser, 1989; Wicker, 1969; Zanna & Fazio, 1982; for a review, see Eagly & Chaiken, 1998).

After early failures to demonstrate a clear link between the two constructs (e.g., LaPiere, 1934), research on the subject stalled and the field adopted the discouraging view that attitudes do not predict behavior consistently (Wicker, 1969). In the 1970s, Fishbein and Ajzen (1974) began to turn the tide of scholarly opinion on the attitude-behavior relation. They observed that attitudes do, in fact, reliably predict behavior when the behavioral criterion is an aggregated set of actions, rather than a single act. Further, they noted that the correlation between attitudes and behavior is larger when the two constructs are equivalent in specificity (Ajzen & Fishbein, 1977). Social psychology's modern understanding of the attitude-behavior relation has been influenced quite strongly by Fishbein and Ajzen's expectancy-value model, the theory of reasoned action, and its descendant theory of planned behavior.

Likewise, the contemporary view of the attitude-behavior relation has been shaped by Fazio's automatic processing model of attitudes and behavior (Fazio, 1990; Fazio et al., 1982, 1983), which holds that attitudes that are accessible in memory can be automatically activated and influence subsequent behavior without perceivers' conscious control or even awareness. Although the Fishbein-Ajzen and Fazio perspectives differ in

the extent to which they view behaviors as resulting from rational and deliberate processes versus automatic activation of accessible constructs in memory, they share the principle that attitudes are indeed meaningful predictors of behavior to the extent that they are measured at similar levels of specificity.

Implicit attitudes and behavior. Despite the voluminous literature on the link between attitudes and behavior, little is known about the circumstances under which attitudes that reside outside conscious awareness direct behavior, and few studies have investigated how implicit attitudes toward social groups influence behavior toward members of those groups (cf. Fazio et al., 1995). In the present research, the measure of candidate preference that was used is not characterized as a measure of behavior, in that participants were aware that the preferences they expressed in the laboratory context would not affect a real (or indeed even a fictional) political election. Nonetheless, one can speculate that such preferences may have relevance for more ecologically valid behaviors, including real voting behavior. In order to more fully understand the social implications of implicit attitudes in future research, it will be important to implement ecologically valid behavioral measures.

To the extent that a more ecologically valid measure of behavior would coincide with the candidate preference results of the present research, the findings of Experiment 1 may provide initial evidence of a new condition under which attitudes may guide behavior. In particular, when attitudes are more consistent in valence (across varying construals of the attitude object), they are more likely to guide explicit judgments—and perhaps also behavior—than when they are less consistent. Although the reasons for this effect remain uncertain, one possible explanation is that, like attitudes that are rooted in

direct experience (Fazio et al., 1982; Fazio & Zanna, 1981), attitudes that are more consistent in valence are more accessible in memory; this heightened accessibility may in turn heighten valence-consistent attitudes' likelihood of influencing judgment and behavior.

Are Implicit Gender Attitudes and Beliefs Malleable?

Experiment 2 investigated the malleability of implicit gender attitudes and stereotypes. Participants were primed either with a construal of women as strong leaders or with a gender-neutral topic. Implicit gender attitudes and stereotypic beliefs were measured. Results showed that the priming intervention successfully altered implicit strong/weak beliefs, with participants who were primed with the construal of women as strong leaders showing less pronounced implicit stereotypic beliefs than did participants in the neutral prime condition. The effect of this intervention was, however, restricted to a change in stereotype; it did not influence women's or men's implicit gender attitudes.

Recent research has suggested that implicit attitudes and beliefs may be more susceptible to change through intervention than had previously been thought (e.g., Blair et al., 2000; Dasgupta & Greenwald, 2000). The present research, which bolsters those findings, is the first to test the malleability of both implicit attitudes and implicit stereotypic beliefs and to examine whether changes in one implicit process exert second-order influences on the other. The results of Experiment 2, in addition to showing that a mild intervention can change implicit gender stereotypes, also provide evidence for the independent function of implicit evaluation and beliefs.

The findings of Experiment 2 suggest several avenues for future research. First, further experiments should address whether an intervention directed at evaluative (or

affective) processes would change implicit attitudes as readily as the present intervention's focus on stereotypes influenced implicit beliefs. One possibility is that, in such an experiment, implicit attitudes would shift but a concurrent change in implicit beliefs would not be observed—a finding that would parallel the results of Experiment 2. That is, it may be that attitude-based and stereotype-based interventions each exert influence on the corresponding implicit judgments. For example, an intervention that required participants to imagine women in a positive (or negative) light may effect change in implicit attitudes toward women, but not alter implicit stereotypes.

Alternatively, however, it is also possible that the effects of interventions to change implicit attitudes and beliefs are asymmetrical. That is, implicit attitudes may be more entrenched, and hence more difficult to change, than are implicit beliefs. A related question concerns whether implicit attitudes are indeed more resistant to change than are implicit beliefs: Would changes in implicit attitudes be likely to induce corresponding changes in implicit beliefs, more so than the reverse? Future experiments should address the many possibilities raised by the results of the present research. Finally, other questions for future research, in the domain of attitude and belief malleability, include the persistence of changes induced in implicit attitudes and beliefs, the incremental effects of interventions that are more and less subtle, and the role that valence-consistency, discussed in much of the present research, plays in attitude and stereotype malleability.

Conclusion

Together, the results of the present research contribute to a developing theoretical understanding of how implicit attitudes are formed, maintained, and put to use in social categorization, judgment, and interaction. The four preliminary experiments and two

main experiments in the present research demonstrate that implicit attitudes toward social groups are guided by group membership and by cultural construals of the group. Second, the present research introduces a concept that has not before been systematically examined: valence-consistency of attitudes across varying construals of the same social group. The results of the present research indicate that this form of attitude consistency may be important for predicting explicit social judgment and behavior, and for that reason, the construct deserves further study. Finally, the results regarding the malleability of implicit attitudes and beliefs are provocative, suggesting that implicit associations may not be as resistant to change—even through a mild intervention—as has been thought.

References

Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology, 18,* 317-334.

Abrams, D., Thomas, J., & Hogg**,** M. A. Numerical distinctiveness, social identity and gender salience. *British Journal of Social Psychology, 29*(1), 87-92.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888-918.

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

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Worcester, MA: Clark University Press.

Banaji, M. R. (2000). Implicit attitudes can be measured. To appear in H. L. Roedeger, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder*. Washington, DC: American Psychological Association.

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* (Vol. 7, pp. 55-76). Hillsdale, NJ: Erlbaum.

Banaji**,** M. R; & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology*, *68*(2), 181-198.

Banaji, M. R., & Greenwald, A. G., & Rosier, M. (1997). Implicit esteem: When collectives shape individuals. Paper presented at the Preconference on Self, Toronto, Canada.

Banaji, M. R., & Hardin, C. (1996). Automatic stereotyping. *Psychological Science, 7*(3), 136-141.

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65*(2), 272-281.

Banaji, M. R., Lemm, K. M., & Carpenter, S. J. (in press). The social unconscious. In A. Tesser & N. Schwarz (Eds.), *Handbook of social psychology: Intraindividual processes*. Oxford: Blackwell.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169-183). New York: Guilford.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *The automaticity of everyday life: Advances in social cognition, Vol. 10* (pp. 1-61). Mahwah, NJ: Lawrence Erlbaum.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York: Guilford.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893-912.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*(1), 104-128.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462-479.

Belleza, 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, Instruments, & Computers*, *18*, 299-303.

Blair, I. V. (in press). Implicit stereotypes and prejudice. Chapter to appear in G. Moscowitz (Ed.), *Cognitive social psychology: On the future of social cognition*. Mahwah, NJ: Erlbaum.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*(6), 1142-1163.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2000). Imagining stereotypes away: The moderation of automatic stereotypes through mental imagery. Unpublished manuscript.

Bourhis, R. Y. (1994). Power, gender, and intergroup discrimination: Some minimal group experiments. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium* (Vol. 7, pp. 171-208). Hillsdale, NJ: Erlbaum.

Bourhis, R. Y., Turner J. C., & Gagnon, A. (1997). Interdependence, social identity and discrimination. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life*. Cambridge, MA: Blackwell.

Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, *4*(1), 79-101.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307-324.

Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 554-594).

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, *28*(2), 59-78.

Burns-Glover, A. L., & Veith, D. J. (1995). Revisiting gender and teaching evaluations: Sex still makes a difference. Special Issue: Gender in the Workplace. *Journal of Social Behavior and Personality*, *10*(6), 69-80.

Butler, D., & Geis, F. L. (1990). Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social Psychology*, *58*(1), 48-59.

Carpenter, S. (1994). Characteristics of gender subtypes: Group and individual differences. *Sex Roles*, *31*(3/4), 167-184.

Carpenter, S., & Trentham, S. (1998). Subtypes of women and men: A new taxonomy and an exploratory categorical analysis. *Journal of Social Behavior and Personality*, *13*(4), 679-696.

Carpenter, S. J. (1997). Implicit and explicit attitudes toward female leaders. Unpublished master's thesis, Yale University.

Carpenter, S. J., & Banaji, M. R. (1997, May). Implicit and explicit attitudes toward female leaders. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Carpenter, S. J., & Banaji, M. R. (1998, April). Implicit attitudes and behavior toward female leaders. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33*(5), 541-560.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224.

Coats, S., & Smith, E. R. (1999). Perceptions of gender subtypes: Sensitivity to recent exemplar activation and in-group/out-group differences. *Personality and Social Psychology Bulletin*, *25*(4), 515-526.

Cortina, L. M, Swan, S., Fitzgerald**,** L. F., & Waldo, C. (1998). Sexual harassment and assault: Chilling the climate for women in academia. *Psychology of Women Quarterly, 22*(3), 419-441.

Cunningham, W. A., Nezlak, J. B., & Banaji, M. R. (2000). Conscious and unconscious ethnocentrism: Revisiting the ideologies of prejudice. Unpublished manuscript.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (in press). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science.*

Dasgupta, N., & Greenwald, A. G. (2000). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired outgroup members. Unpublished manuscript.

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

Deal, J. J., & Stevenson, M. A. (1998). Perceptions of female and male managers in the 1990s: Plus ça change… *Sex Roles, 38*(3/4), 287-300.

Deaux**,** K. (1996). Social identification. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 777-798). New York: Guilford.

Deaux, K., & Lewis, L. L. (1983). Assessment of gender stereotypes: Methodology and components. *Psychological Documents, 13*(2), 25.

Del Boca, F. K., Ashmore, R. D., & McManus, M. A. (1986). Gender-related attitudes. In R. D. Ashmore & F. K. Del Boca (Eds.), *The social psychology of female-male relations: A critical analysis of central concepts* (pp. 121-163). Orlando, FL: Academic Press.

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

Doll, J., & Ajzen, I. (1992). Accessibility and stability of predictors in the theory of planned behavior. *Journal of Personality and Social Psychology, 63*, 754-765.

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*(5), 510-540.

Draine, S. C. (1998). Implicit [computer software]. Available at: http://www.millisecond.com.

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

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

Eagly, A. H., & Chaiken, S. (1995). Attitude strength, attitude structure, and resistance to change. In R. E. Petty & J. A. Krosnick (Eds.), Attitude strength: Antecedents and consequences (pp. 413-432). Mahwah, NJ: Erlbaum.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 269-322).

Eagly, A. H., & Karau, S. J. (1999). Few women at the top: Is prejudice a cause? Unpublished manuscript.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*(1), 3-22.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, *15*(4), 543-558.

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? *Psychology of Women Quarterly*, *15*, 203-216.

Eckes, T. (1994a). Explorations in gender cognition: Content and structure of female and male subtypes. *Social Cognition*, *12*(1), 37-60.

Eckes, T. (1994b). Features of men, features of women: Assessing stereotypic beliefs about gender subtypes. *British Journal of Social Psychology, 33*, 107-123.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204-243). New York: Guilford.

Fazio, R. H. (1989). On the power and functionality of attitudes: The role of attitude accessibility. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 153-179). Hillsdale, NJ: Erlbaum.

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

Fazio, R. H. (in press). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*.

Fazio, R. H., Blascovich, J., & Driscoll, D. M. (1992). On the functional value of attitudes: The influence of accessible attitudes upon the ease and quality of decision making. *Personality and Social Psychology Bulletin, 18,* 388-401.

Fazio, R. H., Chen, J., McDonel, E. C., Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, *18*, 339-357.

Fazio, R. H., Jackson, J. R., Dunton, 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*(6), 1013-1027.

Fazio, R. H., Powell, M. C., & Herr, P. M. (1983). Toward a process model of the attitude-behavior relation: Accessing one's attitude upon mere observation of the attitude object. *Journal of Personality and Social Psychology*, *44*, 723-735.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229-238.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 161-202). San Diego, CA: Academic Press.

Fishbein, M., & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*, 59-74.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 357-411). Boston, MA: McGraw-Hill.

Fiske, S. T. (in press). Stereotyping, prejudice, and discrimination at the seam between the centuries: Evolution, culture, mind, and brain. (Millennial Series). *European Journal of Social Psychology*.

Fiske, S. T., Bersoff, D. N., Borgida, E., Deaux, K., & Heilman, M. E. (1991). Social science research on trial: The use of sex stereotyping research in *Price Waterhouse v. Hopkins. American Psychologist, 46*, 1049-1060.

Fitzgerald**,** L. F, Drasgow, F., & Magley, V. J. (1999). Sexual harassment in the armed forces: A test of an integrated model. *Military Psychology*, *11*(3), 329-343.

Fox, R. L., & Smith, E. (1998). The role of candidate sex in voter decision-making. *Political Psychology*, *19*(2), 405-419.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509-517.

Glaser, J. C. (1999). The relation between stereotyping and prejudice: Measures of newly formed automatic associations. Unpublished manuscript, Yale University.

Glaser, J. C., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology, 77*(4), 669-687.

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*(3), 491-512.

Goldberg, P. (1968). Are women prejudiced against women? *Trans-action, 5*, 28-30.

Goodwin, S. A., & Banaji, M. R. (2000). Unpublished raw data.

Gordon, C. (1968). Self-conceptions: Configurations of content. In C. Gordon & K. Gergen (Eds.), *The self in social interaction* (pp. 115-136). New York: Wiley.

Greenglass, E. R. (1982). *A world of difference: Gender roles in perspective*. Toronto: Wiley.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Rosier, M. (2000a). Prologue to a unified theory of attitudes, stereotypes, and self-

concept. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition and behavior*. New York: Cambridge University Press.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2000b). Unified theory of implicit social cognition. Unpublished manuscript.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273*(5282), 1699-1702.

Greenwald, A. G., Klinger, M. R., & Liu, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory & Cognition, 17*(1), 35-47.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74* (6), 1464-1480.

Gregg, A. P. (2000). The Hare and the Tortoise: The Origins and Dynamics of Explicit and Implicit Attitudes. Unpublished Doctoral Dissertation, Yale University.

Haddock, G., & Zanna, M. P. (1994). Preferring “housewives” to “feminists.” *Psychology of Women Quarterly, 18*, 25-52.

Heilman, M. E. (1995). Sex stereotypes and their effects in the workplace: What we know and what we don’t know. Special issue: Gender in the Workplace. *Journal of Social Behavior and Personality, 10*(6), 3-26.

Heilman, M. E., Block, C. J., & Martell, R. F. (1995). Sex stereotypes: Do they influence perceptions of managers? Special issue: Gender in the Workplace. *Journal of Social Behavior and Personality, 10*(6), 237-252.

Hermans, D., De Houwer, J., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition and Emotion, 8*, 515-533.

Hogg, M. A., & Abrams, D. (1988). *Social identification: A social psychology of intergroup relations and group processes*. London: Routledge.

Hogg, M. A., & Abrams, D. (1990). Social motivation, self-esteem and social identity. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 28-47). New York: Harvester Wheatsheaf.

Karpinski, A., & Hilton, J. L. (in press). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*.

Kasof, J. (1993). Sex bias in the naming of stimulus persons. *Psychological Bulletin*, *113*(1), 140-163.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin, 24*(4), 407-416.

Kay, F. M., & Hagan, J. (1995). The persistent glass ceiling: Gendered inequalities in the earnings of lawyers. *British Journal of Sociology*, *46*(2), 279-310.

Kenworthy, L., & Malami, M. (1999). Gender and inequality in political representation: A worldwide comparative analysis. *Social Forces, 78*(1), 235-268.

Kihlstrom**,** J. F. (1990). The psychological unconscious. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 445-464). New York: Guilford.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*(4), 565-579.

Lane, K. L., & Banaji, M. R. (2000). Unpublished raw data.

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

Lemm, K. M., & Banaji, M. R. (1998, May). Implicit and explicit gender identity and attitudes toward gender. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*(2), 275-287.

Lepore, L., & Brown, R. (1999). Exploring automatic stereotype activation: A challenge to the inevitability of prejudice. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 141-163). Malden, MA: Blackwell Publishers.

Lindeman, M., & Sundvik, L. (1995). Evaluative bias and self-enhancement among gender groups. *European Journal of Social Psychology*, *25*(3), 269-280.

Lips, H. M. (1988). *Sex and gender: An introduction*. Mountain View, CA: Mayfield.

Lowery, B. S., Hardin, C. D., & Sinclair, S. S. (1999, May). Contingency of automatic prejudice on interpersonal relationships. Paper presented at the annual meeting of the American Psychological Society, Denver.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*(5), 808-817.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33*(5), 471-489.

Matlin, M. W. (1987). *The psychology of women*. New York: Holt, Rinehart, and Winston.

Mattei, L. R. W. (1998). Gender and power in American legislative discourse. *Journal of Politics, 60*(2), 440-461.

McGuire, G. M., & Reskin, B. F. (1993). Authority hierarchies at work: The impact of race and sex. *Gender and Society, 7*(4), 487-506.

McKee, J. P., & Sherriffs, A. C. (1957). The differential evaluation of males and females. *Journal of Personality, 25*, 356-371.

Melamed, T. (1995). Barriers to women's career success: Human capital, career choices, structural determinants, or simply sex discrimination. *Applied Psychology: An International Review*, *44*(4). 295-314.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Millar, M. G., & Tesser, A. (1989). The effects of affective-cognitive consistency and thought on the attitude-behavior relation. *Journal of Experimental Social Psychology, 25*, 189-202.

Mitchell, J., Nosek, B. A., & Banaji, M. R. (1998). A rose by any other name? Dissociated attitudes toward social group members. Paper presented at the annual meeting of the American Psychological Society, Washington, DC.

Monson, C. M, & Langhinrichsen-Rohling, J. (1998). Sexual and nonsexual marital aggression: Legal considerations, epidemiology, and an integrated typology of perpetrators. *Aggression & Violent* Behavior, *3*(4), 369-389.

Morrison, A. M., & von Glinow, M. A. (1990). Women and minorities in management. *American Psychologist*, *45*(2), 200-208.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103-122.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226-254.

Neumann, R., Ebert, M., Gabel, B., Gülsdorff, J., Kranich, H., Lauterbach, C., & Wiedl, K. (1998). Vorurteile zwischen Bayern und Norddeutschen: Die Anwendung einer neuen Methode zur Erfassung evaluativer Assoziationen in Vorurteilen [Prejudice between Bavarians and Northern Germans: The use of a new method for understanding evaluative associations of prejudice]. *Zeitschrift für Experimentelle Psychologie*, *45*, 99-108.

Nisbett**,** R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231-259.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (1998, May). Math = Bad + Male, Me = Good + Female, Therefore Math ≠ Me. Paper presented at the annual meeting of the American Psychological Society, Washington, DC.

Nosek, B. A., Cunningham, W., Banaji, M. R., & Greenwald, A. G. (2000). Measuring implicit attitudes on the Internet. Paper presented at the 1st annual meeting of the society for personality and social psychology, Nashville, TN.

Oakes, P.J., & Turner, J.C. (1980). Social categorization and intergroup behaviour: Does minimal intergroup discrimination make social identity more positive? *European Journal of Social Psychology, 10,* 295-301.

O'Connor, K. J., Cunningham, W. A., Banaji, M. R., Gore, J., Gatenby, C., & Phelps, E. A. (2000, February). The implicit association test and fMRI: Neural basis of response competition. Poster presented at the 1st annual meeting of the Society for Personality and Social Psychology, Nashville, TN.

Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology, 59*(3), 475-486.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches,* Dubuque, IA: William C. Brown.

Petty, R. E., & Krosnick, J. A. (Eds.) (1995). *Attitude Strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.

Petty, R. E., & Wegener, D. T. (1998). Attitude Change: Multiple roles for persuasion variables. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 323-390).

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Powell, M. C., & Fazio, R. H. (1984). Attitude accessibility as a function of repeated attitudinal expression. *Personality and Social Psychology Bulletin, 10*(1), 139-148.

Pratkanis, A. R., Breckler, S. J., & Greenwald, A. G. (Eds.) (1989). *Attitude structure and function.* Hillsdale, NJ: Erlbaum.

Reed, T. (2000). Automatic in-group preference: Evidence from the implicit minimal group paradigm. Unpublished manuscript.

Rhode, D. L. (1997). *Speaking of sex: The denial of gender inequality*. Harvard University Press.

Richeson, J. A., & Ambady, N. (1999). Implicit associations in context: Effects of group membership and situational roles. Unpublished manuscript, Harvard University.

Roper Starch Worldwide (1999, December). Women in Leadership Poll: Women in Elected Office. Available from Deloitte & Touche LLP, 1633 Broadway, New York, NY, 10011.

Rosenberg, M. J. (1960). An analysis of affective-cognitive consistency. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 15-64). New Haven, CT: Yale University Press.

Rosenkrantz, P. S., Vogel, S. R., Bee, H., Broverman, I. K., & Broverman, D. M. (1968). Sex role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology, 32,* 287-295.

Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology, 63*(2), 198-211.

Rubin, M., & Hewstone, M. (1998). Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification. *Personality and Social Psychology Review, 2*, 40-62.

Rudman, L. A., & Glick, P. (in press). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*.

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (1999). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. Unpublished manuscript.

Rudman, L. A., Greenwald, A. G., Mellot, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition*, *17*(4), 1-29.

Rudman, L. A., & Kilianski, S. E. (in press). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*.

Ruggiero, K. M., Mitchell, J. P., Krieger, N., Marx, D. M., & Lorenzo, M. L. (in press). Now you see it, now you don't: Explicit versus implicit measures of the personal/group discrimination discrepancy. *Psychological Science.*

Saris, R. N., Johnston, I., & Lott, B. (1995). Women as cues for men's approach or distancing behavior: A study of interpersonal sexist discrimination. *Sex Roles*, *33*(3-4), 289-298.

Sarnoff, I. (1960). Psychoanalytic theory and social attitudes. *Public Opinion Quarterly*, *24*, 251-279.

Schiessl, M. (2000). Unpublished raw data.

Schneider, W. (1990). MEL user's guide: *Computer techniques for real time psychological experimentation*. Pittsburgh: Psychology Software Tools.

Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-process models. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 323-336). New York: Guilford.

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology, 71*(5), 888-898.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. Austin, TX: University of Texas Press.

Stangor**,** C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology, 75*(5), 1191-1197.

Steele**,** C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.

Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (in press). Using the Implicit Association Test to investigate attitude-behavior consistency for stigmatized behavior. *Cognition and Emotion*.

Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin, 105*(3), 409-429.

Tajfel, H. (1981). The social psychology of minorities. In H. Tajfel (Ed.), *Human Groups and Social Categories*, (pp. 309-362). Cambridge, MA: Cambridge University Press.

Tajfel, H., & Turner, J.C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W.G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.

Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83-114). Hillsdale, NJ: Erlbaum.

Taylor, S. E., Fiske, S. T., Close, M., Anderson, C., & Ruderman, A. S. (1977). Solo status as a psychological variable: The power of being distinctive. Unpublished manuscript, Harvard University.

Thurstone, L. L. (1931). *The measurement of attitude*. Chicago, IL: University of Chicago Press.

Turner, J.C. (1985). Social categorization and the self-concept: A social cognitive theory of group behaviour. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research, Vol. 2* (pp. 77-121). Greenwich, CT: JAI Press.

Turner, J.C., Hogg, M. A., Oakes, P.J., Reicher, S. D., & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.

van Knippenberg, A., & Dijksterhuis, A. (1996). A posteriori stereotype activation: The preservation of stereotypes through memory distortion. *Social Cognition, 14*(1), 21-53.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology, Vol. 2 (4th ed.)* (pp. 446-496). Boston, MA: McGraw-Hill.

Wicker, A. W. (1969). Attitude versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *23*(4), 41-78.

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*(2), 262-274.

Yzerbyt**,** V. Y., Leyens, J., Schadron, G. (1997). Social judgeability and the dilution of stereotypes: The impact of the nature and sequence of information. *Personality and Social Psychology Bulletin, 23*(12), 1312-1322.

Yzerbyt**,** V. Y., Schadron, G., Leyens, J., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology, 66*(1), 48-55.

Zanna, M. P., & Fazio, R. H. (1982). The attitude-behavior relation: Moving toward a third generation of research. In M. P. Zanna, E. T. Higgins, & C. P. Herman (Eds.), *Consistency in social behavior: The Ontario Symposium* (Vol. 2, pp. 283-301). Hillsdale, NJ: Erlbaum.

Appendix A

Evaluative Stimuli for All Experiments

Negative Words	Positive Words	Negative Words	Positive Words
PE1		*Experiment 1*	
Abnormal	Beach	Abnormal	Birthday
Awful	Birthday	Awful	Excellent
Bad	Convenient	Broken	Happy
Broken	Dream	Cancer	Heaven
Cancer	Excellent	Grotesque	Laughter
Cemetery	Good	Hurt	Liberty
Cyclone	Great	Maggot	Perfect
Defect	Happy	Mildew	Star
Displeasing	Heaven	Poison	Treasure
Grotesque	Honest	Ugly	Vacation
Hurt	Justice		
Inferior	Laughter		
Lice	Liberty		
Maggot	Nice		
Mildew	Palace		
Naughty	Paradise		
Pest	Party		
Poison	Perfect		
Poor	Pure		
Rabies	Refreshment		
Rusted	Star		
Scorn	Sunrise		
Ugly	Treasure		
Waste	Vacation		
PE2, PE3, PE4		*Experiment 2*	
Death	Freedom	Awful	Excellent
Devil	Gift	Cancer	Gift
Maggot	Laughter	Maggot	Laughter
Measles	Paradise	Terrible	Paradise
Toxic	Vacation	Toxic	Vacation

Appendix B

Names and Professional Positions of Leaders in PE4

Male Leaders

John Dasburg	CEO, Northwest Airlines
Bill Daley	Secretary of Commerce
Mark Yudof	President, University of Minnesota
Bruce Babbitt	Secretary of the Interior
David Murdock	CEO, Dole Food Corporation
Al Lerner	CEO, MBNA Corporation
Fred Thompson	Senator from Tennesee
George Fisher	CEO, Eastman Kodak Company
Pete Domenici	Senator from New Mexico
William Cohen	Secretary of Defense

Female Leaders

Alexis Herman	Secretary of Labor
Olympia Snowe	Senator from Maine
Judith Rodin	President, University of Pennsylvania
Jill Barad	CEO, Mattel Corporation
Donna Shalala	Secretary of Health and Human Services
Kay Hutchison	Senator from Texas
Ann Fudge	President, Maxwell House Coffee
Elaine Chao	CEO, United Way
Janet Reno	U.S. Attorney General
Patricia Barron	President, Xerox Engineering Systems

Appendix C

Candidate Descriptions for Experiment 1

Description 1

[Lisa Reilly] has been in the public limelight for almost two decades. [She] began [her] political career in the local government of [her] city, where [she] served for several years as a judge and a member of the city council. In 1984, [Reilly] mounted a successful campaign for the state senate, where [she] has now served several terms. As a state senator, [she] has developed expertise in a number of important areas including education and employment issues. [Reilly] has led a number of efforts to revive the state's urban areas and to draw employers to impoverished areas, efforts that have begun to see some successful results. Although [Reilly] has been criticized for [her] style of interaction in the political arena, [she] is well respected by liberals and conservatives alike as a competent and socially conscious political leader. In [her] current campaign, [Reilly] has focused on education, unemployment, entitlement spending, and international trade issues.

Description 2

[Gary Nichols] has been a fixture in the state government for 18 years. [He] began [his] political life as a state senate aide, where [he] was instrumental in the creation of a number of urban development programs across the state. From 1980-88, [Nichols] served as an advisor to the governor of [his] state and then helped direct the governor's bid for [his] party's presidential ticket. [He] then successfully ran for the state senate, where [he] has now served several terms and has become an active member of a number of senate committees. [Nichols] has been a leader on issues ranging from urban development and environmental issues to health care and education. [His] colleagues recognize [Nichols] as an able leader and a politician with integrity, despite some of [his] controversial political decisions. In [his] current campaign, [Nichols] has emphasized issues pertaining to welfare and health care reform, the global economy, and campaign finance.

Appendix D

Imagery Primes for Experiment 2

Female Leader Prime

Thank you for participating in this experiment.

We would like you to spend the next several minutes imagining women who are tough and aggressive leaders (in business, government, etc.). What are these leaders like? In the space below, please write some examples of traits and behaviors that might describe women who are authoritative and demanding leaders. There are no right or wrong answers, of course—you should use this opportunity to express whatever thoughts come to mind. Your thoughts might include ways in which these women's characteristics, behaviors, and decision-making strategies illustrate their power and toughness.

When the timer sounds, please turn to the computer and follow the instructions on the screen to continue with the next part of the study.

Neutral Prime

Thank you for participating in this experiment.

We would like you to spend the next several minutes imagining the sights on a Yale campus tour. What are some of the most noteworthy elements of the campus? In the space below, please write some examples of buildings, lawns, and other landmarks that are part of a tour of the Yale campus. There are no right or wrong answers, of course—you should use this opportunity to express whatever thoughts come to mind.

When the timer sounds, please turn to the computer and follow the instructions on the screen to continue with the next part of the study.