Lemm,K. & Banaji,M.R. (1999). Unconscious attitudes and beliefs about women and men. In U.Pasero & F. Braun(eds.), Wahrnehmung und Herstellung von Geschlecht (Perceiving and performing gender) (pp 215-233). Opladen: Westdutscher Verlag.

Unconscious attitudes and beliefs about women and men[1]

Kristi Lemm and Mahzarin R. Banaji
Department of Psychology, Yale University

In the last four decades, analyses of gender have produced entirely new fields of study and created deep transformations in traditional disciplines of scholarly inquiry. Whatever their orientation, these analyses share a recognition of historically prevalent inequities that place sharp limits on women's access to knowledge, power, and wealth. Among social scientists, examinations of such inequities have shaped the debate about gender differences and assumptions that difference (actual or assumed) implies the superiority of what is male, or at least that what is male is normative.

With blinding gender discrimination as the background, it is not surprising that psychologists expected and found that beliefs about and attitudes toward women were more negative than those toward men. In a mid-century examination of attitudes toward men and women, Fernberger (1948) observed that males were rated as "all around" superior to females. This report was supported by McKee and Sherriffs' (1957) observation that 93% of male participants rated men as slightly to greatly superior to women on a question about the "overall general worth" of men relative to women. Female participants concurred with this attitude, with 86% indicating that men as a group were indeed of greater worth than women. This overall favorability toward men appeared to persist in spite of the consciousness-raising impact of the women's movement of the 1970s. Using data collected in 1978, Werner and LaRussa (1985) replicated the McKee and Sherriffs (1957) study and observed that men were still rated as superior to women overall, although the gender gap had narrowed somewhat.

Social, political, and economic discrimination against women, along with evidence supporting the belief in women's inferior nature, conspired to produce the consensus that in general, beliefs about and attitudes toward women are more negative than those toward men. Thus it came as a surprise when Eagly and her colleagues posed an empirically-based challenge to this view, with their observation that both males and females rated women quite favorably, even more so than they rated men (Eagly/Mladinic 1989). Although the observation that women are held in positive regard went against conventional wisdom and previous research, this finding has been replicated in subsequent independent research (Eagly/Mladinic/Otto 1991, Glick/Fiske 1996, Haddock/Zanna 1994).

---

1 The research reported in this paper was supported by grants to Mahzarin Banaji from the National Science Foundation (SBR-9422241 and SBR-9709924) and fellowships from the John Simon Guggenheim Foundation and the Cattell Fund, and a National Science Foundation Graduate Research Fellowship to Kristi Lemm. We thank Siri Carpenter and Richard Hackman for their helpful comments on a previous draft.

What accounts for this difference between the conclusion of several decades of research and the opposing findings of Eagly and colleagues? Eagly and Mladinic (1989) proposed that such apparent discrepancies may stem from a failure in the research literature to have distinguished between overall evaluations of women as a social group versus evaluations of women's social roles. Women's roles and professions are typically afforded lower status than men's, and evaluation of these roles is correspondingly lower than evaluation of men's roles. But overall positive feelings toward women may be observed nonetheless. Women as mothers, friends, and partners are liked and loved, even if women are less desirable when they acquire the label of feminist, engineer, or senator. Eagly and Mladinic designed their measures carefully to separate these two components of attitudes, and explained why indeed they had observed genuinely favorable attitudes toward women, although men were rated higher on particular qualities, specifically those linked to agentic and instrumental traits.

Alternatively, the favorability toward women obtained by Eagly and Mladinic (1989) may simply derive from a change in attitudes toward women over time. Such a thesis is supported by Werner and LaRussa's (1985) finding that the gap between attitudes toward men and women was smaller in 1978 than it was twenty years earlier. In fact, it may be that attitudes toward women turned even more positive in the decade before Eagly and Mladinic collected their data. However, because attitude change was measured by accessing explicit attitudes and beliefs, i.e., those that are consciously available, it is not possible to know the boundary conditions of such a finding. It is possible that conscious and unconscious attitudes toward men and women had both undergone change. But it may also be the case that the nature of the data collection had allowed respondents to monitor their responses on questionnaires to avoid appearing sexist to themselves, to the experimenter, and to society at large. Explicitly expressed attitudes may well have masked attitudes that were unconsciously present.

In the decade since Eagly/Mladinic's proposal that attitudes toward women are not uniformly negative, at least two other analyses have also suggested a more complex view of attitudes toward gender. Glick/Fiske (1996) proposed that rather than being simply negative, sexist attitudes toward women are often ambivalent, encompassing positive as well as negative evaluations, with sexist attitudes falling into two components; hostile and benevolent sexism. Hostile sexism corresponds to the traditional definition of sexism as antipathy toward women. Benevolent sexism, in contrast, involves endorsing stereotypes about women that are subjectively positive, at least to those who hold the stereotypes, although not to the women who are the targets of these stereotypes. Benevolent sexists may endorse beliefs such as "Women should be cherished and protected by men", or "A good woman should be set on a pedestal". Although these statements are evaluatively positive, they are believed to reflect attitudes that undermine gender equality and preserve systems of subordination.

Fiske et al. (in press) have expanded the concept of benevolent sexism by theorizing that the attitudinal dimensions of respect and liking may be unrelated. That is, a dominant group may respect members of a subordinate group even though they do not like them (e.g., businesswomen), or they may like members of a subordinate group whose abilities they do not respect (e.g., housewives). Fiske et al. point out that overall ratings of women may be ambiguous because they elicit liking and disrespect toward some subtypes of women (e.g., "sexy chicks") and disliking but respect for other subtypes (e.g., feminists).

Jackman (1994) also claims that attitudes toward women can include positive and negative components. Most critically, she argues that positive attitudes toward women constitute the foundation of male dominance, with paternalism being the agent of subordination, which is justified out of concern and even love for the dominated. In such analyses, members of dominant groups, such as men, hold positive attitudes toward those whom they dominate, such as women, however self-interested the origin of such positivity. Such positions point out the complexity of social attitudes, both in the interpretation of the reason for positive regard and in the dissociations in attitude toward various subtypes of the larger social group.

This paper offers an analysis of beliefs and attitudes about women and men that diverge from previous investigations in that the focus is primarily on thoughts and feelings that lie outside conscious awareness or conscious control. Whereas most measures assume explicit or conscious access to the contents of the attitude or belief, the work we describe takes advantage of recently developed measures that tap automatically evoked beliefs and attitudes. At the turn of the century, it seems appropriate to assess the state of beliefs and attitudes toward women and men based on new learning that has emerged about unconscious forms of automatic attitudes and beliefs more generally and about the measurement of such processes itself. By the term *belief* we refer to assessments of specific attributes of men and women (e.g., on dimensions such as strong-weak). By *attitude*, we refer to assessments of global evaluation of men and women (e.g., such as good-bad, pleasant-unpleasant).

Unconscious beliefs and attitudes

The inference that a person is male or female happens rapidly, without conscious control, and possibly without conscious awareness. To intuitively appreciate the routinized and habitual nature of this particular mental operation, consider situations in which a person's sex identity is forced into conscious awareness because it is not readily discernable. The slower and more deliberate search for information that such a situation provokes reveals its otherwise spontaneous occurrence. For the past several decades, psychologists have investigated the nature of mental processes that lie outside conscious awareness and conscious control, and in recent years, such analyses have expanded to include the nature of social judgment as well. How automatically are social categories (e.g., gender, race, class) activated? What distinguishes judgments that arise from conscious versus unconscious assessments? Does the social category of the perceiver moderate the effects of automatic categorization, perception, and memory?

In pursuit of answers to these questions, research in experimental social psychology is transforming the manner in which attitudes and beliefs more generally are viewed. For more than a hundred years, the dominant theories and methods of experimental psychology assumed a conscious mode of operation – a tradition that may have its roots, in part, in an articulated rejection of psychodynamic theory, and in part, in the unavailability of tractable methods to investigate unconscious processes. Thoughts and feelings, including those that concern social objects (persons, social groups) have traditionally been assessed as if they operated largely in a conscious mode. In a paper advocating the need to measure unconscious or implicit modes of thinking and feeling about self and social groups, Greenwald and Banaji (1995) defined unconscious *belief (stereotype)* as:

> (...) introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category. (1995:15)

and unconscious *attitude* as:

> (...) introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects. (1995:8)

These definitions have their origin in two assumptions about thought, feeling, and behavior. First, the bounded rationality of mental systems place obvious limits on the ability to introspect accurately. Second, a desire to be morally and socially acceptable to one's self and others further diminishes access to thoughts and feelings. While the same assumptions also form the basis of the psychoanalytic unconscious, the similarity ends there. Analyses of the cognitive unconscious as reflected in the experiments we report rely on working knowledge of theories of associative learning and memory, perception and categorization processes.

Measurement of unconscious beliefs and attitudes

Standard measures of attitudes and beliefs typically ask questions in a direct manner. For example, to assess the extent to which men and women are associated with qualities such as good/bad, strong/weak, or emotional/unemotional, verbal self-report measures are commonly obtained in which numerical values are assigned to reflect the strength of such associations. Or, to assess beliefs, responses to statements such as "Do you agree that men and women ought to have equal opportunities for employment?" are often sought. Although such probes are useful if the goal is to document the state of consciously available attitudes and beliefs that reflect personal and social standards, they are ineffective means of reaching attitudes and beliefs that remain hidden from conscious awareness and that lie beyond introspective reach. If, as in many fields of inquiry, there is a growing recognition that large portions of thought and feeling systems operate in an unconscious mode, such a mode cannot be ignored in investigations of implicit social cognition (Banaji/Lemm/Carpenter in press, Greenwald/Banaji 1995).

It has been far easier to recognize a need for alternative measures of attitude and belief than to generate tractable measures that assess the output of the complex mental operations that are involved. How should attitudes and beliefs of which one is not aware be measured? Some families of indirect measures have involved unobtrusive observation of behavior, such as noting how often help is offered to a woman in physical difficulty (Latane/Rodin 1969) or using a hidden video camera to observe reactions to the telling of sexist jokes (LaFrance/Woodzicka 1998). Another class of measures has relied on projective techniques, in which participants generate stories in response to ambiguous drawings or photographs that are shown to predict behavior better than self-reported values and intentions (Biernat 1989, McClelland 1985).

Our own research has been directed toward relatively new measures of unconscious gender beliefs and attitudes. Such tools belong to a family of measures of unconscious mental processes that are based on well-worn principles of associative learning and memory. If two concepts have come to be associated through experience (e.g., female-*delicate*, male-*rough*), the presence of one (e.g., female) should produce relative facilitation of the other (e.g., *delicate* compared to *rough*). Such facilitation can be measured in a variety of ways. As an example, a word such as *delicate* may be presented for a brief duration, with the task being to judge if the word that follows it is representative of the category male (e.g., John) or female (e.g., Susan). To the extent that delicate is more strongly associated to female than male, responses to "Susan" should be faster than responses to "John" when they follow *delicate*. The uncontroversial assumption that such differences in response speed reveal the differential strength of association between *categories* (male, female) and *attributes* (good/bad, delicate/rough) is extended in research on social cognition to include the more controversial assumption that the strength of such mental associations is a meaningful indicator of automatic attitude (or evaluation) and automatic belief (or stereotype). The primary research from our laboratory includes two measures of time to respond (measured in milliseconds) that have been used successfully to study unconscious beliefs and attitudes – a variant of priming procedures mentioned above, and a newer task to measure the strength of implicit associations.

Measuring the unconscious component of gender beliefs and attitudes is a relatively new enterprise, and as a result, only a modest amount of research on unconscious stereotyping and prejudice with a focus on gender even exists. In this chapter, we do not provide a comprehensive review of past research, but rather, we focus primarily on recent research in a single laboratory as illustrative of the approach. Specifically, we describe three types of evidence for the unconscious operation of gender concepts. First, we discuss evidence for the activation of gender beliefs (stereotypes) that occur without conscious control, through brief exposure to gender-linked information, and reveal consensually shared knowledge. While our framing of the introductory comments have focused on attitude (evaluative, liking), investigations of related constructs of belief or stereotype allow interesting comparison of similarity and overlap as well as difference and dissociation. Next, we discuss research on automatic attitudes toward women and the degree to which they are a function of one's

own group membership and the particular subtypes that are the target of social attitudes (e.g., mothers, women leaders). Finally, we show the attributes of automatic gender identity and how it is related to beliefs and attitudes toward men and women.

Unconscious gender stereotypes

In a now classic report of the consensus regarding the psychological separation of male from female, Broverman et al. (1972) obtained ratings of the extent to which selected traits were descriptive of men and women and found that men were more often associated with competency-related traits whereas women were more associated with warmth- and expressiveness-related traits. The Brovermans' technique for assessing gender stereotypes is a common one and not restricted to the measurement of gender alone; it assumes that stereotypes are simply attributes associated with a particular group, with the strength of the stereotype being reflected through a numerical score that reflects the conscious assessment. In measuring the content of unconscious stereotypes, we assume similarly that gender stereotypes are attributes associated with social groups, but we measure them without soliciting a conscious judgment of the relationship between social group and attribute.

Instead, as an example, Banaji and Hardin (1996) conducted an experiment in which they showed participants gender-specific pronouns (e.g., *he, she*) one at a time on a computer screen. Participants were asked to judge as quickly as possible whether each pronoun (the *target*) was male or female, with only about one half of a second (500 milliseconds) needed to make such judgments. Prior to seeing the pronoun to be judged, they saw a word representing a gender-stereotypic occupation (*secretary, engineer*), or a gender-specific title (*Ms., Mr.*), kinship term (*mother, father*), or term with a sex-specific suffix (e.g., *chairman, salesgirl*). This word (the *prime*) appeared long enough for participants to read it, but not long enough for them to deliberate on its meaning (approximately 250 milliseconds.). Although the instruction was to ignore the prime words when judging the target pronouns that followed, the mere presence of these primes produced a notable effect on judgment of target pronouns. Specifically, responses were significantly faster when the prime and target were consistent in gender (however else they may differ) than when they were inconsistent. Thus, judgments of the pronoun *he* were made more quickly after seeing words such as *engineer* than words such as *secretary*, whereas the opposite was true for judgments of *she*. In other words, the judgment of the pronoun was facilitated or inhibited to the extent to which the prime was gender-consistent. The relative facilitation of gender-consistent than gender-inconsistent pairs reveals that gender is coded automatically in grasping meaning. The magnitude of this automatic gender stereotyping effect was unrelated to consciously expressed beliefs about gender and views regarding the use of gender-neutral language in speech and writing.

Blair and Banaji (1996) demonstrated variations of this phenomenon using a similar task. They measured the length of time it took participants to identify common Anglo-American first names (e.g., *Steve, Susan*) as male or female names. Prior to the judgment of each target name, participants saw a briefly presented prime that was either a trait stereotypically associated with gender (e.g., *gentle, strong*), or an attribute stereotypically associated with gender (e.g., *skirt, trousers*). As predicted, names were categorized more quickly when immediately preceded by a gender-congruent prime than by a gender-incongruent prime. Can such effects that reveal the automatic nature of gender stereotypes be avoided or changed? Blair and Banaji created conditions under which a strong expectation of gender mismatch between prime and target were created (e.g., *engineer-Jane, secretary-John*) and showed that only when the demands of the judgment was obtained at a much slower pace were gender inconsistent pairings readily accepted.

In a challenge to this view, Blair and Ma (1999) used a different procedure, one in which they first created the opportunity to imagine "a strong woman" and showed that such imagery significantly increased automatic stereotypes of women as strong (compared with a control imagery condition). This finding is supported by recent experiments on automatic anti-Black race attitudes. Dasgupta/Greenwald (1999) show a similar reduction in automatic prejudice (pro-White automatic attitudes) following exposure to admired Black Americans, raising the possibility that even long-standing automatic stereotypes and attitudes may be more malleable than previously assumed. However, experimental findings that raise the possibility of malleable unconscious stereotypes must be viewed in context – the blunt fact remains that social life does not easily provide the mental props that allow counterstereotypic or counterattitudinal associations to be activated, and as such the experimentally observed loosening of automatic gender stereotypes remains possible but not probable. Although experiments to explore unconscious stereotypes and prejudice would lead us to believe that they are elicited by linguistic (often word-level) representations of social information, it is quite obviously the case that dependence on textual representations largely reflects the superior theoretical development and technical convenience of verbal rather than nonverbal representations. There is little doubt that social perception, categorization, memory, and judgment occur in response to nonverbal representations of information.

Lemm et al. (1999) explored how pictorial representations activate gender stereotypes, as well as how stereotype activation can influence judgments about such pictorial representations. They showed participants pictures that either strongly or weakly evoked gender. For example, a female or male face or full body form was considered a strong gender prime (*denoting* female and male), whereas pictures of objects such as an oven mitt or a baseball mitt were regarded as weaker gender prime (*connoting* female and male). Using a variant of the priming procedure previously used with words, they observed strong facilitation for targets that followed gender-consistent primes. Remarkably, across 200 male and female pictures, not only did strong gender primes serve to elicit faster categorization of gender congruent targets, but so also did weaker gender connoting primes.

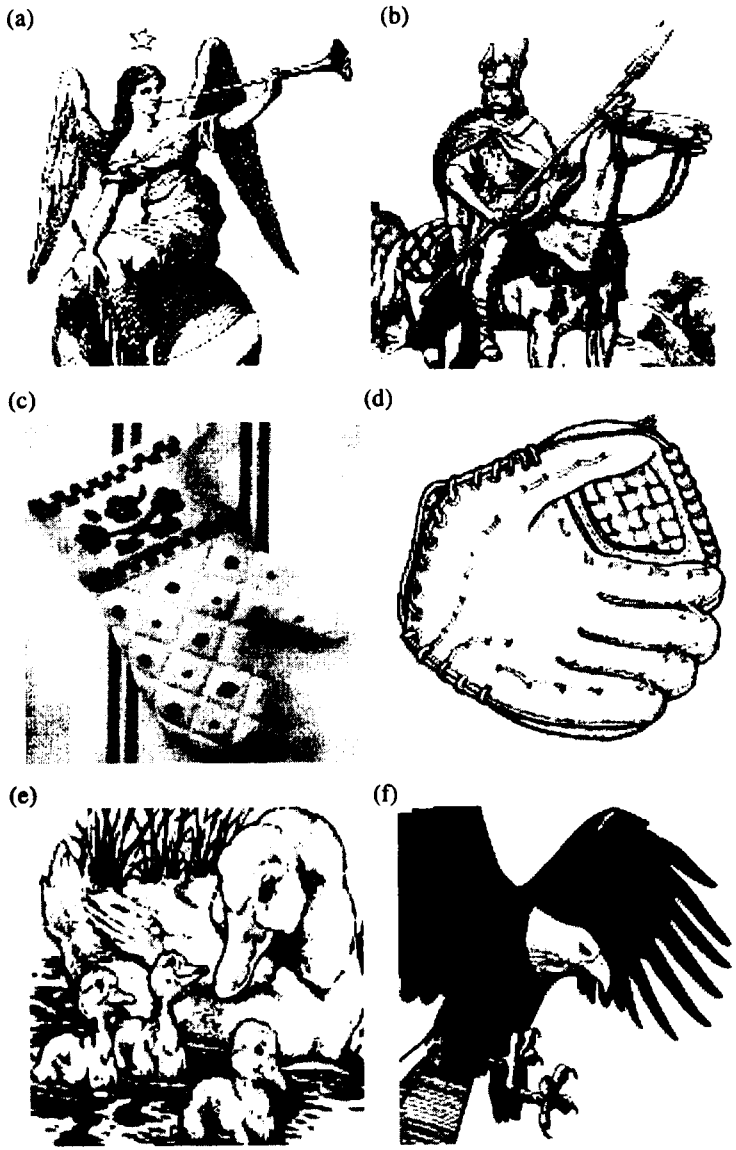

Figure 1: Sample stimulus items from Lemm et al. (1999)

Figure 1 shows samples of the pictures used in these experiments, with panels (a) and (b) representing strong gender primes, (c) and (d) representing a class of weaker gender primes (objects used more frequently by men or women) and (e) and (f) representing a different class of weaker gender primes (objects implying femininity or masculinity through remote association). That pictures in panels c-f served as sufficiently potent primes to automatically evoke judgments of female and male on targets that denote gender (e.g., a male or female face or full body) underscores the integral nature of gender as a component of meaning.

Lemm et al. (1999) used pictorial representations of gender to further explore the connection between gendered language and imagery. Psychologists and linguists have spoken about such a relation, although a firm empirical basis for concluding whether gender-specific language constrains subsequent ability to imagine gender-free thought is lacking. Lemm et al. (1999) demonstrated that words that are stereotypically masculine or that contain male-specific suffixes (e.g., *salesman*) lead to faster identification of male pictures, whereas stereotypically feminine words or words that contain female-specific suffixes (e.g., *businesswoman*) lead to faster identification of female pictures. That is, having learned that engineers and senators are likely to be men appears to cognitively constrain thought so as to include men more readily than women as possible candidates. Such research links discoveries about cognitive functioning (i.e., the constraints on automatic thought) to questions of equal opportunity and access. If social discourse automatically constrains the ability to imagine women as doctors, senators, and CEOs (and complementarily, constrains the ability to imagine men as nurses, secretaries, and sex workers) it raises the question of the impact of granting or denying cognitive "admission" in the early stages of information processing on downstream judgments such as evaluating a person's performance.

A relatively new measure of automatic association is the Implicit Association Test (IAT, Greenwald/McGhee/Schwartz 1998), which has two properties that distinguish it from other measures. First, it appears to be a particularly sensitive measure of strength of association. For example, almost 80% of participants show a relationship between *male-career* and *female-family* and over 90% of participants show an association between *old-bad* and *young-good.* Second, the IAT allows the subjective experience that some associations are cognitively easier (e.g., *female-home)* than others (e.g., *female-work*) and as a result, can be a tool that readily reveals the dissociation between consciously held and unconsciously exhibited responses. The form of the IAT applied to investigate stereotypes and attitude toward a variety of social groups may be sampled at www.yale.edu/implicit.

The idea underlying the IAT is simple: if two concepts are associated in memory, they will be easier to associate in judgment or behavior compared to concepts that are less associated or counterassociated in memory. To grasp the idea that produces the IAT effect, imagine sorting a randomly ordered stack of pictures of men and women into two piles, one containing men, the other containing women. Also imagine sorting another stack of pictures of say, power tools and kitchen utensils, again

into separate piles. Now, imagine shuffling the two stacks and sorting female faces and kitchen utensils into a single pile, and male faces and power tools into another pile. Measure the time it takes to do this. Now, imagine an alternative arrangement in which female faces and power tools are sorted into a single pile and male faces and kitchen utensils make up the other pile. Again, measure the time to do this sorting. Research with the IAT suggests that such a sorting task should be mentally harder than the previous one, and as a result it should take longer to complete. Measured typically by sorting on a computer that records responses in milliseconds, the two conditions of the IAT permit a consciously palpable experience of differential mental ease when performing the male+power tools versus female+power tools sortings. In several experiments to date, this differential ease of pairing has been used as an indicator of the strength of automatic gender stereotypes.

Beliefs about gender differences continue to persist even in the face of knowledge about their instability and reduction over time. One such belief concerns men and women's differential ability in fields of academic achievement – men are believed to possess the ability that permits excellence in math and science, and women are believed to have greater affinity for language and arts. Nosek, Banaji, and Greenwald (1998) studied the strength of this stereotype at the automatic level (its presence at the conscious level is also easily detected) using the IAT, and observed that both men and women show strong unconscious association between male and math/science and female and arts/language. Knowledge of such difference is widespread in the culture, and automatic stereotypes reflect the extent of individual learning of such an association. If such associations are simply a reflection of cultural knowledge, and are not valid measures of the individual's unconscious beliefs, such effects should not predict other judgments. However, as Rudman and Glick (1999) showed, the stronger the unconscious stereotype, the more respondents tended to downgrade the skills of a female job applicant.

Nosek, Wang, and Banaji (1999) more specifically examined unconscious beliefs about gender roles in the professional context. They observed that male and female college students show automatic beliefs linking men with leadership roles (e.g., *boss, CEO, director*) and women with "helper" roles (e.g., *assistant, attendant, secretary*), reflecting gender distributions. In addition, men were linked more strongly with career roles (*business, profession, work*), whereas women were more strongly linked with the home (e.g., *domestic, family, household*). Even as more women have entered the workforce, it appears that as long as they continue to be associated with the home, such beliefs will continue to be visible on measures of strength of association. Goodwin and Banaji (1999) substituted the category *powerful* and *powerless* for Wang et al's *leader* and *helper* categories. They too observed that men and women both endorsed the gender stereotype, responding faster when *male+powerful* was paired, and slower when *female+powerful* was paired.

Rudman and Glick (1999) used the IAT to assess the stereotype that women are relatively more communal and that men are relatively more agentic. Both men and women were faster to respond when female names+communal terms and male names+agentic terms were paired than when the pairing was reversed. In part, because measures such as priming and the IAT are discordant with prevailing notions about the meaning of "belief" and "stereotype" or "attitude" and "prejudice", in particular that they inherently refer to conscious constructs, there is uniquely high demand for demonstrations of the validity of such measures. We point to Rudman and Glick's (1999) finding that the strength of automatic association between female+communal and male+agentic was related to judgment of male and female job candidates. The greater the magnitude of the unconscious gender stereotype of agentic/communal, the lower were ratings of the social skills of a job applicant, and the lower the willingness to hire her relative to an otherwise identical male applicant. Notably, the strength of the unconscious gender stereotype was not related to participants' conscious endorsement of gender stereotypes. The lack of relationship between conscious and unconscious stereotyping, observed in many experiments in this tradition, is theoretically expected (i.e., conscious and unconscious representations are assumed to be independent and separate) but socially problematic. That is, even those who do not explicitly endorse gender stereotypes may nonetheless harbor sufficiently strong implicit stereotypes and such beliefs may guide behavior.

Research on unconscious gender stereotypes leads to three conclusions. First, splitting humans into female and male is a sufficiently fundamental cognitive act that categorization by gender is automatic, and in the sense used here, unavoidable. Many dimensions of attributes (physical, psychological, social) are linked firmly enough to female and male that such attributes automatically activate associated gender. As the examples from the picture priming study showed, an oven mitt activated "female" more swiftly just as baseball mitt activated "male". Whatever one's conscious desire to avoid categorizing as female and male, this research suggests that such categorization is inevitable and with cognitive and social consequences for those who perceive and are perceived.

Second, relevant research, of which we have reviewed only a small sample, suggests that the automatic activation of stereotypes does not sit in any obvious relation to consciously held or desired beliefs. Often, correlations between conscious and unconscious measures are close to zero, and when they are more robust, it is unclear what produces such relationships (Blair in press). While conscious attempts to fake or shape automatic associations are not successful, the next wave of research will reveal the extent to which automatic associations are indeed malleable through brief new experiences.

Finally, both men and women hold similar unconscious associations about gender. Not surprisingly, if women and men have shared learning experiences regarding the nature and magnitude of gender associations, such similarity in learning and memory should be reflected in performance. Conscious expressions of stereotypes, in so far as they reflect a desire to represent what "ought" to be or desirable response to self and others, may be more sensitive to group differences because such conscious demands may be differentially placed on members of social groups. But in so far as unconscious gender beliefs reflect the state of the world as filtered through personal

experience, they may not distinguish between members of social groups. However, this finding of similar patterns of males' and females' implicit gender stereotypes will cease to hold when we turn our attention to implicit gender attitudes. Here, group differences will be visibly present showing that "knowing" and "liking" vary in their sensitivity to self and social group.

Unconscious attitudes toward gender

Research on unconscious stereotypes shows that the mere presence of a gender signifying stimulus in the environment – represented through ordinary symbols such as a name, a common noun, an adjective, a picture – activates knowledge of the category and its attributes. It is important to note, however, that although these gender stereotypes can be positive, negative, or neutral, we do not consider them to represent an *attitude*. Our understanding of the term stereotypes refers to *beliefs* or knowledge about what the category is, not an evaluation (favorable or unfavorable) of it.

Just as gender categories can automatically evoke knowledge of attributes associated with them, they can also elicit automatic evaluations or attitudes. In this section, we describe research that has enabled the measurement of unconscious attitudes toward social groups including men and women. We show that attitudes toward women and men do exist at an unconscious level, and that these unconscious attitudes do not always parallel the attitudes that men and women consciously express.

As noted in the introduction, it is only in the last decade that researchers have observed that consciously expressed attitudes toward women may actually be quite positive (Eagly/Mladinic 1989). This observation was supported by Carpenter and Banaji (1998a), who also showed that both men and women provided more favorable ratings of women than men. In both investigations, not only were attitudes favorable toward women, there appeared to be no difference between men's and women's expression of attitude.

In recent research, we extended Eagly and Mladinic and Carpenter and Banaji's findings by measuring unconscious attitudes. Like unconscious stereotypes, we think of unconscious attitudes in the form of associations between a category and an attribute, with the difference being that the attribute in the case of attitude is a global evaluation along an evaluative dimension – *good* versus *bad* or *pleasant* versus *unpleasant*. Unconscious attitudes can vary in direction (positive or negative) as well as strength, with some attitudes held more strongly than others. By our definition, an attitude toward *female* would be represented by the degree to which the attribute *good* or *pleasant* is relatively easily paired with the category *female* or *woman*. To the extent that pairings of *female* and *good* exist, we conclude that a favorable attitude toward *female* exists.

Carpenter and Banaji (1998a) used the IAT to assess college students' implicit attitudes toward women and men. Their participants sorted typical Anglo-American male or female first names at the same time as evaluatively pleasant or unpleasant items (e.g., *hostile, maggot; gentle, heaven*), in both combinations, female+pleasant and male+pleasant. They found that women were evaluated more favorably overall, particularly by female participants. The finding of favorable implicit attitudes toward women is consistent with other findings of explicit attitudes. However, whereas recent research using explicit measures shows little or no subject sex effect in explicit evaluation of men and women, implicit attitudes do differentiate between male and female participants. Females show vastly more positive attitudes toward women. Males, on the other hand, hover around the implicit neutral point, with some samples showing slight positivity toward women relative to men.

Mitchell, Nosek, and Banaji (1999) used a similar IAT to assess unconscious gender bias, but they used a wider range of stimulus names, including names typical of Black Americans as well as White Americans. They observed that men and women both showed a preference for female over male, but women showed this preference more strongly than did men. Thus, it appears that implicit and explicit attitudes toward gender diverge – whereas men and women hold similar attitudes toward gender at an explicit level, at an implicit level, women tend to show much stronger favorability toward women and corresponding implicit negativity toward men.

Explicit and implicit attitudes toward women are further complicated by the finding that not all groups of women are rated equally favorably or unfavorably. For example, lesbians and feminists tend to be explicitly rated negatively relative to more stereotypically feminine groups of women, such as housewives (Haddock/Zanna 1994, Saris/Johnston/Lott 1995). A subtype that has received substantial research attention is female leaders, women who compete for positions of power in society. Explicit ratings of females in leadership roles (e.g., managers, politicians) show that female leaders are evaluated more negatively than male leaders (Heilman/Block/Martell 1995), particularly when such women assume a stereotypically male leadership style (Eagly/Makhijani/Klonsky 1992).

Carpenter and Banaji (1997) adapted the IAT measure to compare ratings of subtypes of women and men. In an experiment in which mothers were evaluated relative to female leaders and fathers relative to male leaders, they observed that mothers were rated more positively than female leaders, whereas fathers were seen to be just as positive as male leaders. Although the subjects in this experiment may harbor no explicit negative attitudes toward women leaders (and many of the women may themselves aspire to be women leaders), these data suggest that males are equally positively regarded both as fathers and leaders, whereas women elicit greater favorability as mothers than as leaders. Carpenter and Banaji further observed that the implicit attitudes they assessed were largely unrelated to explicitly assessed attitudes, suggesting that the measure of unconscious attitude may in fact tap contents that differ from those assessed by conscious measures of attitude.

In a follow-up study, Carpenter and Banaji (1998b) pitted attitudes toward male and female leaders against one another, rather than against mother and father. They again observed substantial differences between men and women's responses, with women indicating a stronger automatic preference for female leaders, and males indicating an opposite but weaker preference for male leaders. Interestingly, men explicitly re-

ported greater conscious preference for female leaders, thus placing self-reported, conscious attitude in contradiction to the unconscious one.

In a third experiment, Carpenter and Banaji (1998b) presented a female and male candidate for election and obtained implicit and explicit measures of liking for them. Both men and women showed an implicit preference for the female candidate. Thus, although they showed preference for a specific female candidate, males showed less liking for female leaders as a general group. Just as one may like one's local politician but not the category "politician", men's responses to a particular female candidate may well be positive even though attitudes toward the group as a whole are not.

In overlapping conditions of a few different experiments, Carpenter and Banaji and Mitchell, Nosek and Banaji have shown one persistent effect: Female participants show strongly positive attitudes toward all representations of *female* over *male*. Females show greater automatic associations of good to mother (than father), to female (than male), to specific newly-encountered female political candidates (than male candidates), to the general category "female leader" (than "male leader"). In none of several experiments conducted to date have these investigators failed to find a robust implicit positive attitude toward women by women. The findings for male participants are less clear. In most cases, men show weaker preferences, with most conditions showing a small preference for female over male. Carpenter and Banaji predicted that the strong and consistent implicit positive attitude of women by women should result in the stronger predictive power of such attitudes on behavior. In support, they found that female participants' strength of implicit attitude toward specific male and female leaders predicted their voting for that leader. For men, whose implicit attitudes toward both gender groups were less strong and consistent, the strength of automatic attitude did not predict voting. Further research is needed to understand the origin of women's strong positivity toward women and men's lack of strong attitude toward either group. It is possible that the implicit association of *female* to *good* is widely shared, and that for those who are themselves members of the culturally favored group, i.e., women, the implicit attitude of positivity is intensified.

Measurement of unconscious attitudes toward women and men represents a considerably smaller body of research than measurement of unconscious stereotypes. However, even with this relatively limited research, two conclusions can be drawn. First, it is quite clear that attitudes about women and men are held at an unconscious level, and that differences in the strength and direction of these unconscious attitudes can be measured. Only in the future will the full implications of implicit attitudes about social groups be revealed in research on the predictive validity of such measures.

Second, implicit attitudes toward gender do not always parallel explicit attitudes toward gender. In particular, there are differences between male and female respondents in their implicitly expressed attitudes toward men and women, and this is in contradiction to findings of implicit beliefs (stereotypes). Whereas typically no differences between women's and men's implicit stereotypes about gender are found, there appears to be considerable difference between women's and men's implicit attitude toward gender. Implicit beliefs may reflect knowledge of social groups at some

distance from self, which would explain the similarity in implicit stereotypes held by males and females. Attitudes, on the other hand, with their foundation in feeling rather than thought, may partly reflect the culture's assessment of social groups (explaining why both women and men show positive attitudes toward women), but importantly, may reflect additionally, an affinity for attributes attached to self (explaining the stronger positive attitude of women toward women). If this reasoning is accurate, White Americans and Black Americans should show positive implicit attitudes toward White rather than Black, with White Americans showing such an attitude more strongly. Cultural evaluations (i.e., the shared evaluation of a group within a culture) and the location of self (i.e., one's own membership in the group or not) are both feeders to implicit attitudes, with groups that are culturally favored and associated to self through membership eliciting the most positive implicit attitudes.

Unconscious gender identity

Perhaps not surprisingly, just as unconscious beliefs and attitudes about gender influence the judgments we make about others, they also have influence on the judgments we make about ourselves. Even those who consciously reject notions of gender differences must find themselves unconsciously adopting culturally imposed conceptions of masculinity and femininity. The extent to which unconscious identification with gender groups can influence choice and opportunity, and the extent to which choices that appear to be consciously made are unconsciously driven by group membership are questions of interest here, although the data are not yet up to the task of answering them.

Psychologists have debated about the meaning and measurement of gender identity. In the first test specifically designed to assess gender identity, Terman and Miles (1936) developed a 910-question scale, called the M-F scale, that tapped feminine characteristics with questions such as "Do you like people to tell you their troubles?" and masculine characteristics with questions such as "Were you ever fond of playing with snakes?" The M-F scale, and most of the scales that have followed, are based on three assumptions: That masculinity and femininity exist but cannot always be identified by observation, that masculinity and femininity are distinct from each other but related to traits and outcomes, including psychological adjustment, and that reporting about one's masculine and feminine attributes is so susceptible to demand that measures of them must be subtle (Morawski 1985). Terman and Miles were cognizant of the difficulty of direct, explicit measures, but nevertheless, their questionnaire obtained fully conscious measures of gender identity out of necessity, and with few exceptions this is the case with most standard measures.

Our research has used a measure of gender identity that circumvents conscious control. We make use of the same reaction time measure described earlier to assess self-identification with cultural stereotypes about masculinity and femininity. It should be noted that our use of the terms masculinity and femininity differs from that of our predecessors. In particular, early scales sought to measure masculinity and feminini-

ty as "pure" constructs, free of cultural bias. In contrast, for our measure, masculinity and femininity are defined as the cultural stereotype of what is appropriate or common for males versus females. We measure the extent to which men and women incorporate the culturally defined view of masculinity and femininity into their self-concept.

In a series of experiments, Lemm and Banaji (1998) measured unconscious gender identity using the IAT. Participants categorized "me" and "not me" words, which were either words specifically relevant or irrelevant to the participants (e.g., their name and hometown or some other person's name and hometown) or pronouns relevant to self versus other (e.g., *I, me; they, them*). Simultaneously, they categorized masculine and feminine words, which were present tense verbs moderately associated with men or women (*e.g., feels, cooks; employs, fixes*). The measure of gender identity was obtained, similar to stereotype and attitude, from the relative speed of pairing self with feminine verbs and traits versus with masculine verbs and traits.

Overwhelmingly, responses indicated that men and women associated their self-concept with the gender associated with their biological sex: Men showed strong implicit links with masculinity and females with femininity. Interestingly, this pattern did not mimic the responses given by the same participants on an explicit measure of gender identity. When asked to rate masculine and feminine verbs for the extent to which they applied to themselves, female participants reported endorsing the masculine verbs almost as much as the feminine verbs, indicating a more androgynous conscious gender identity. Thus, for women, the pattern of unconscious gender identity differed from the pattern of conscious gender identity. The participants in this experiment were college students, from a subculture in which it is socially acceptable, even encouraged, for women to endorse positive masculine behaviors such as *owning* and *governing*, in addition to feminine behaviors. As such, acceptance of masculine actions is visible on measures that enable conscious endorsement. But on less uncontrollable measures, women are less able to escape traditional gender roles and attributes. In contrast, perhaps because masculine acts are more socially valued, men showed a more congruent pattern of responses on implicit and explicit measures (automatic associations are in alignment with what is also consciously desired), and indicated a more masculine gender identity on both.

Nosek et al. (1998) also used the IAT to measure gender identity, substituting stimuli that are denotative of male or female (e.g., *he, father; she, mother*) to represent male and female categories for a measure of sex identity. The pattern of responses they observed was very similar to that by Lemm/Banaji, with women exhibiting a stronger link between self and female, and males showing the reverse pattern of a stronger automatic link between self and male. Nosek et al. found that, for women, this implicit sex identity was related to gender stereotypes about mathematics and arts. Women who were strongly female-identified tended also to hold the implicit association that math is male and arts is female. Further, the more strongly women were female-identified, the less they identified with math. Thus, women who have an implicit female identity also have the automatic association that math is more appropriate for men than for women, and more appropriate for others than for themselves.

The same relationships regarding sex identity and academic preferences were not observed among male participants. Men tended to identify themselves with math more so than women did, but this identification was not related to their endorsement of the math-arts gender stereotype, nor their identity as male. However, men's identification with math and the strength of the automatic belief that math is male was related to their performance on a standardized math test (Scholastic Aptitude Test). The greater their belief that math is male-associated, and the greater their association of math and self, the better they performed on a test of math ability. Such implicit attitudes may be the foundation of the behaviors, both conscious and unconscious, that explain the exodus of women from math and science fields during college and beyond.

Goodwin and Banaji (1999) examined the link between self and power, to better understand the link between gender and power. Implementing an IAT measure with *me* and *not-me* items paired against the concepts of powerful and powerless (e.g., *manager, director; assistant, servant*), they observed that both men and women tended to associate powerful roles, rather than powerless roles, with themselves. In addition, both men and women showed an implicit association between powerful and good. Of particular interest was a positive relationship between power identity (me-power association) and power attitude (power-good association) suggesting that what is implicitly associated to self is also what is implicitly liked.

Lemm et al. (1999) studied the way that gender-specific language can be incorporated into men's and women's self-concept. At the college at which their experiment was conducted, students in their first year of undergraduate study continue to be referred to as *freshmen*. The generic masculine term includes a male suffix (*men*), but is explicitly intended to refer to both male and female members of the first-year class. As expected, research participants reported believing that the terms *freshman* and *freshmen* are gender-neutral, because they are equally likely to be applied to male as to female students. However, performance on the IAT in which pairing of me and not-me against generic masculine terms (*freshman, freshmen*) and gender neutral terms (*frosh, first-year*) were used showed clear gender differences. Men exhibited a stronger identification with generic masculine terms (e.g., *freshmen*) than did women. Women showed just the opposite pattern of responses, indicating a stronger implicit identification with gender neutral terms (e.g., *first-year*). It appears that even though women routinely refer to themselves as "freshmen", they implicitly associate terms of more neutral origin to themselves. Thus, even when generic masculine terms are automatically associated with the groups male and female equally, men and women may not apply generic masculine terms equally to themselves – at least not at levels that lie outside conscious control. Identifying the implications of such differences in unconscious cognition of commonly used terms for psychological outcomes such as subtle experiences of inclusion remain to be explored.

The research described shows that gender is an integral part of the self-concept even, and perhaps especially, when assessed on measures that avoid conscious control. Women may explicitly seek a more masculine/androgynous gender identity, but unconscious associations with femininity are strong. Implicit gender identification is related to unconscious attitudes about gender as well as to unconscious stereotypes about gender. Such associations have implications for performance on gender-related tasks such as academic orientation, choice of career, and commitment to career.

Conclusion

Any analysis of the nature of humans and the structure of societies must immediately confront questions of gender. What is the nature of female and male, their relation to each other, their aspirations for the future? Of the many fundamental questions about women and men that scientists have posed, we focused on beliefs and attitudes. Of course, beliefs and attitudes about women and men are obtained from members of the very groups about whom we seek appraisal. This problem, of the knower also being the known, as William James would have framed it, has led to investigations such as ours in an attempt to bypass conscious awareness and control. This chapter surveys research, largely from a single laboratory, on unconsciously held beliefs about female and male, unconsciously expressed evaluations of positivity toward female and male, and unconscious identification with male and female, masculine and feminine.

Male and female reside not only as physical realities in the outer world, but they have psychological presence and meaning in the mental world as well. Their difference in mental status is detected through the beliefs we hold about them, our feelings toward them, and our identification with them. Research on gender stereotypes show that male and female are mentally distinct categories that are automatically and uniquely differentiated. Additionally, objects that have come to be associated with male and female (sometimes even weakly) reflect gendered meanings. In so far as such associations are learned, recent research is beginning to show how new learning about social groups (e.g., *female-strong*) that contradicts past learning (e.g., *female-weak*) can change existing associations that underlie unconscious attitudes and beliefs.

The automatic associations we examined under the term "identity" point out a similar difficulty with assumptions that one's association to female and male is freely chosen. More than measures that rely on consciously accessed gender identity are able to detect, our experiments show the almost isomorphic mental representation of sex and gender – at levels of thinking that are outside conscious control, men strongly identify as masculine, females as feminine. Such a finding has sobering implications as we consider the discrepancies that can arise between conscious aspirations for qualities that are socially valued and desired (as did female subjects in the study in which they consciously endorsed actions such as *governing*), and unconscious lack such endorsement (as did female subjects when they did not show implicit association between self and *governing*). The invisible fences of group membership need new and urgent attention.

Implicit attitudes toward men and women do not show patterns that are consistent with early research summarized at the outset, for there is no evidence of implicit negativity toward women. In fact, in keeping with Eagly's research, the representation of female is positive. But implicit attitudes toward men and women do not fully fit the patterns revealed in contemporary research using conscious measures either. Our data suggest that simply being female predisposes one to have strong implicit attitudes of positivity toward female over male. Understanding the basis of such a predisposition and its implications for judgment and decision-making is the challenge reserved for new research.

References

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

Banaji, M.R./K.M. Lemm/S.J. Carpenter (in press): The social unconscious. In: *Handbook of Social Psychology.* Ed. A. Tesser/N. Schwarz. Oxford, UK: Blackwell.

Biernat, M. (1989): Motives and values to achieve: Different constructs with different effects. In: *Journal of Personality* 57 (1). pp. 69-95.

Blair, I.V. (in press): Implicit stereotypes and prejudice. In: *Future directions in social cognition*. Ed. G. Moskowitz. Hillsdale, NJ: Erlbaum.

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

Blair, I.V./J. Ma (1999): *Imagining stereotypes away: The moderation of automatic stereotypes through mental imagery.* Unpublished manuscript. University of Colorado, Boulder, CO.

Broverman, I.K./S.R. Vogel/D.M. Broverman/F.E. Clarkson/P.S. Rosenkrantz (1972): Sex-role stereotypes: A current appraisal. In: *Journal of Social Issues* 28. pp. 59-77.

Burns-Glover, A.L./D.J. Veith (1995): Revisiting gender and teaching evaluations: Sex still makes a difference. In: *Journal of Social Behavior and Personality* 10 (6). pp. 69-80.

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

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

Carpenter, S.J./M.R. Banaji (1998a): *Unpublished raw data*. Yale University.

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

Dasgupta, N./A.G. Greenwald (1999): *Exposure to admired members of stigmatized groups reduces automatic negativity.* Unpublished manuscript, University of Washington, Seattle, WA.

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

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

Eagly, A.H./A. Mladinic/S. Otto (1991): Cognitive and affective bases of attitudes toward social groups and social policies. In: *Journal of Experimental Social Psychology* 30. pp. 113-137.

Fernberger, S.W. (1948): Persistence of stereotypes concerning sex differences. In: *Journal of Abnormal and Social Psychology* 43. pp. 97-101.

Fiske, S.T./J. Xu/A.J. Cuddy/P. Glick (in press): Respect versus liking: Status and interdependence underlie ambivalent stereotypes. In: *Journal of Social Issues*.

Glick, P./S.T. Fiske (1996): The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. In: *Journal of Personality and Social Psychology* 70 (3). pp. 491-512.

Goodwin, S./M.R. Banaji (1999): *Gender and power: Evidence for an implicit personal/group discrimination discrepancy.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

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

Greenwald, A.G./D.E. McGhee/J.L.K. Schwartz (1998): Measuring individual differences in implicit cognition: The implicit association test. In: *Journal of Personality and Social Psychology* 74 (6). pp. 1464-1480.

Haddock, G./M.P. Zanna (1994): Preferring "housewives" to "feminists". In: *Psychology of Women Quarterly* 18. pp. 25-52.

Heilman, M.E./C.J. Block/R.F. Martell (1995): Sex stereotypes: Do they influence perceptions of managers? In: *Journal of Social Behavior and Personality* 10 (6). pp. 237-252.

Jackman, M.R. (1994): *The velvet glove.* Berkeley, CA: University of California Press.

LaFrance, M./J.A. Woodzicka (1998): No laughing matter: Women's verbal and nonverbal reactions to sexist humor. In: *Prejudice: The target's perspective.* Eds. J.K. Swim/C. Stangor. San Diego, CA: Academic Press. pp. 61-80.

Latane, B./J. Rodin (1969): A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. In: *Journal of Experimental Social Psychology* 5. pp. 189-202.

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

Lemm, K.M./M. Dabady/M.R. Banaji/J. Kluewer (1999): *The influence of gendered language on thought.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

McClelland, D.C. (1985): How motives, skills, and values determine what people do. In: *American Psychologist* 40 (7). pp. 812-825.

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

Mitchell, J./B.A. Nosek/M.R. Banaji (1999): *Dissociated implicit attitudes: Examples from race, gender, and profession.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Morawski, J.G. (1985/1998): The measurement of masculinity and femininity: Engendering categorical realities. In: *The gender and psychology reader.* Eds. B.M. Clinchy/J.K. Norem. New York: New York University Press. pp. 199-226.

Nosek, B.A./M.R. Banaji/A.G. Greenwald (1998): *Me = female, math = male, therefore math ≠ me.* Paper presented at the annual meeting of the American Psychological Society, Washington, DC.

Nosek, B.A./C. Wang/M.R. Banaji (1998): *Unpublished raw data.* Yale University.

Rosenkrantz, P./S. Vogel/H. Bee/I. Broverman/D. Broverman (1968): Sex-role stereotypes and self-concepts in college students. In: *Journal of Consulting and Clinical Psychology* 33. pp. 287-295.

Rudman, L.A./P. Glick (1999): *Implicit gender stereotypes and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of managers.* Unpublished manuscript, Rutgers University.

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

Terman, L.M./C.C. Miles (1936): *Sex and personality: Studies in masculinity and femininity.* New York, NY: Mcgraw-Hill.

Ware, R./J.J. Cooper-Studebaker (1989): Attitudes toward women as managers with regard to sex, education, work, and marital status. In: *Psychological Reports* 65. pp. 347-350.

Werner, P.D./G.W. LaRussa (1985): Persistence and change in sex-role stereotypes. In: *Sex Roles* 12 (9-10). pp. 1089-1100.