

RACE AND GENDER CATEGORIZATION:  
THE SPECIAL CASE OF BLACK WOMEN

By

Amanda K. Sesko

Submitted to the graduate degree program in Psychology  
and the Faculty of the Graduate School of the University of Kansas  
In partial fulfillment of the requirements for the degree of  
Master's of Arts

---

Monica Biernat, Ph.D., Chair

---

Nyla Branscombe, Ph.D.

---

Christian Crandall, Ph.D.

Date defended: May 29, 2007

The Thesis Committee for Amanda K. Sesko certifies  
That this is the approved version of the following thesis:

RACE AND GENDER CATEGORIZATION:  
THE SPECIAL CASE OF BLACK WOMEN

---

Monica Biernat, Ph.D., Chair

---

Nyla Branscombe, Ph.D.

---

Christian Crandall, Ph.D.

Date approved: \_\_\_\_\_

### Acknowledgements

I would like to thank Monica Biernat for her advice and support with the planning, conducting, and writing of this thesis. Special thanks to Nyla Branscombe and Christian Crandall for their comments throughout the work of this project as well as for serving on this committee. In addition, I would like to thank the faculty and graduate students of the University of Kansas for their always helpful and supportive contribution to this research. Finally, I would like to thank the Biernat research lab for their comments and help conducting this work.

## Abstract

The present study investigated categorization along race and gender lines, with a particular emphasis on how Black women are categorized by Whites. As an alternative to the double jeopardy hypothesis (Beale, 1970), I proposed an *invisibility hypothesis*, suggesting that Black women's non-prototypicality (with regard to both the race and gender categories) leads to their decreased visibility and increased involvement in categorization errors. Participants listened to a group discussion among 2 Black female, 2 White female, 2 Black male, and 2 White male targets. Categorization errors were assessed when participants had to recall "who said what" during the group discussion (Taylor et al., 1978). I predicted that participants would be most likely to incorrectly match statements said by Black women with other Black women, White women and men, and Black men. Across all targets (Black/White, female/male,) participants tended to use subtyping, categorizing by both race and gender. Supportive of the invisibility hypothesis, Black women were most implicated and confused with all other targets, specifically White men, such that they were the least likely to be given credit for statements they said than any other group.

## Table of Contents

|                        |     |
|------------------------|-----|
| Title Page.....        | i   |
| Acceptance Page.....   | ii  |
| Acknowledgements.....  | iii |
| Abstract.....          | iv  |
| Table of Contents..... | v   |
| List of Tables.....    | vi  |
| List of Figures.....   | vii |
| Introduction.....      | 1   |
| Method.....            | 22  |
| Results.....           | 26  |
| Discussion.....        | 38  |
| References.....        | 49  |
| Table.....             | 56  |
| Figures.....           | 57  |
| Appendix A.....        | 67  |
| Appendix B.....        | 68  |

## List of Tables

|  |    |
|--|----|
| Table 1: Possible Error Types for Each Target..... | 56 |
|--|----|

## List of Figures

|  |    |
|--|----|
| Figure 1: Overall category errors.....   | 57 |
| Figure 2: Within-race/within-gender errors.....                                  | 58 |
| Figure 3: Within-race/between-gender errors.....                                 | 59 |
| Figure 4: Between-race/within-gender errors.....                                 | 60 |
| Figure 5: Participant genders effects on between-race/between-gender errors..... | 61 |
| Figure 6: Overall accuracy by target and participant gender.....                 | 62 |
| Figure 7: Mean foil statement errors by target race and gender.....              | 63 |
| Figure 8: Mean reaction times for correct matches by target race and gender..... | 64 |
| Figure 9: Evaluation ratings by target race and gender.....                      | 65 |
| Figure 10: Gender in hiring decisions by target race and gender.....             | 66 |

## Race and Gender Categorization: The Special Case of Black Women

It has been over a hundred and fifty years since Sojourner Truth delivered her famous “Ain’t I a woman?” speech to a women’s rights convention. While these words are highly valued and discussed among scholars and laypersons alike, Black women are still struggling for their voices to be heard (Bell, 1992; Brown-Collins & Sussewell, 1986; Jones & Shorter-Gooden, 2003; King, 1988). As a discipline, social psychology has been dedicated to studying racial and gender stereotyping and prejudice. However, the majority of research investigates Black men as targets of racism, and White women as targets of sexism, disregarding Black women. While Black women are underrepresented in the literature, there has been a debate as to the relative disadvantage or advantage of Black women due to their unique dual category membership. The prominent theory arguing in favor of disadvantage is the *double jeopardy hypothesis*, or “double hit” of both racism and sexism (Beale, 1970). In other words, Black women are assumed to be worse off than White women and Black men because they are subject to both racism and sexism. Data supporting this hypothesis, however, have been sparse, and only recently have researchers begun to empirically test, challenge, and offer alternatives to the double jeopardy hypothesis (e.g., Levin, Sinclair, Veniegas, & Taylor, 2002; Sidanius & Pratto, 1999; Sidanius & Veniegas, 2000).

Double jeopardy was first introduced by Beale (1970) in a feminist piece addressing the unique experiences of Black women. The hypothesis itself does not directly make predictions as to the nature of the dual effect of gender and racial



discrimination (other than suggesting that things are bad for Black women), but two relevant lines of research have emerged. One assumes *additive* effects from both racism and sexism, the other assumes an *interaction* between the two (Sidanius & Veniegas, 2000). One potential problem with the double jeopardy hypothesis is its underlying assumption that racism and sexism toward Black women take the same form as the racism Black men experience and the sexism White women experience. Some researchers, however, have acknowledged that while gender and racial discrimination may take on many of the same features for Black women as it does for White women and Black men, the motivations for each may be qualitatively different (Hurtado, 1989; Sidanius & Veniegas, 2000). For example, Essed (1991) noted that the gender and race of Black women are not easily separated, and therefore lead to a sort of “gendered racism,” that makes it nearly impossible to pinpoint the specific impacts of gender and race alone. In other words, simple additive or multiplicative effects may not be enough to capture the unique experiences of Black women. With this in mind the current research also argues that *perceptions* of Black women may not be simply additive or multiplicative evaluations due to race and gender. Rather, Whites’ perceptions of Black women, and particularly their categorizations of them, may reflect the relative non-prototypicality, or *invisibility* of Black women.

### *History and Stereotypes of Black Women*

There is considerable evidence suggesting the experiences and perceptions of Black women are very different from those of their White female and Black male counterparts. Many of these differential experiences may be rooted in slavery. For

example, historically, Black women have held very different roles than White women (Bell, 1992; Brown-Collins & Sussewell, 1986; Davis, 1981; Sanchez-Hucles, 1997). They were more likely to be forced into both public and private spheres (working outside the home and rearing children in the home) and during times of slavery were assumed to have the favored position (rearing children of White families) over Black men (Brown-Collins & Sussewell, 1986; Davis, 1981). Black women were and are more likely than White women to be head of the household (Brown-Collins & Sussewell, 1986; Davis, 1981).

These roles are very different from the 19<sup>th</sup> century ideology of femininity, and sharply contrast with perceptions of White women as passive and submissive. There is some evidence to suggest that Black women are associated with “masculine” traits (self-reliance, independence, assertiveness, and strength), more so than White women (Binion, 1990; Robinson, 1983; West, 1995), and less associated with feminine traits such as emotional, passive, and dependent (Landrine, 1985). While they are perceived as more masculine and less feminine, Black women themselves tend to hold more traditional beliefs about female roles in family than White women (Binion, 1990).

Additionally, scholars have described three subtypes of Black female stereotypes -- the mammy, sapphire, and jezebel -- clearly illustrating the diverse and unique perceptions associated with Black women (Thomas, Witherspoon, & Speight, 2004 ). The mammy is characterized as obese and dark-skinned. She is the housekeeper, cook, nanny and is assumed to be non-threatening, strong, selfless and

supportive (Thomas, et al., 2004). The sapphire is nagging, emasculating, loud, and argumentative. The jezebel was derived from the sexual exploitation of Black women during slavery and in many ways was used as a justification for rape (Thomas et al., 2004). She is seductive, manipulative, animalistic, hypersexed, and unable to control sexual desires.

The traits associated with the mammy, sapphire, and jezebel also map on to contemporary stereotypes of Black women (Thomas et. al., 2004; Jones, & Shorter-Gooden, 2003). For example, Jones and Shorter-Gooden (2003) asked Black women to list negative stereotypes held of their group. In line with the sapphire and jezebel in particular, commonly listed attributes were promiscuous, immoral, whore, hostile, rude, incompetent, domineering, emasculating, aggressive, on welfare, lots of babies, loud, money-hungry, and lazy. Further, 80% felt that they being were negatively affected by these stereotypes. Thomas et al. (2004) also asked Black women to rate the existence of stereotypic traits associated with the mammy, sapphire and jezebel subtypes and found agreement that these traits reflected perceptions of their group as a whole. These stereotypes are unlike the passive and emotional stereotypes associated with White women. Further, while many of these stereotypic traits are consistent with the predominant stereotypes of Black men in the United States (e.g., lazy, hostile, and uneducated; see Devine & Elliot, 1995), others (e.g., non-threatening and selfless) are in direct opposition to the stereotypes of Black men.

While there are clear historical and contemporary differences in the roles and stereotypes of Black women, there is little research directly comparing the stereotypes

of Black women to Black men and White women. In one exception, Niemann, Jennings, Rozelle, Baxter, and Sullivan (1994) had participants list 10 adjectives associated with different race and gender groups. Using cluster analysis they found that while some adjectives/traits associated with Black women overlapped with Black men (e.g., unmannerly, athletic) and White women (e.g., pleasant, sociable), the frequencies at which they were mentioned differed. In other words, the strength of the association between the adjectives and groups differed, indicating there is something particular about each group as an entity rather than just a mere breakdown by race and gender.

### *Black Women and Life Outcomes*

Beyond the history and stereotyping of Black women, the most common way researchers have attempted to test the double jeopardy hypothesis is to examine the self-esteem/well-being of Black women and discrimination in economic/work-related domains. For example, double jeopardy would predict that Black women are more likely than Black men or White women to have lower self-esteem, face workplace discrimination, and wage disparities.

### *Self-Esteem*

While the double jeopardy hypothesis suggests that being Black and female leads to worse outcomes (and particularly lower self-esteem), some have argued that Black women's unique experiences have fostered resilience and strength (Brown-Collins & Susswell, 1986; Myers, 1980). They have learned how to cope under many stressors (Myers, 1980; Shorter-Gooden, 2004), have acted as activists and upheld

strong community bonds (Brown-Collins & Susswell, 1986). Others argue, however, that Black women have been socialized to appear strong, independent, and resilient, and in many respects this “superwoman” image is only a façade that does not protect them from the many injustices associated with being Black and women (Bell & Nkomo, 1998; Shorter-Gooden, 2004). In other words, while they remain strong on the outside, discrimination due to racism and sexism still affects them internally.

Empirical evidence on self-esteem however, seems to support resilience rather than disadvantage. In a meta-analysis, Major, Barr, Zubek, and Babey (1999) found that while White males have slightly higher self-esteem than White women ( $d = -.20$ ), Black women and men do not differ ( $d = .03$ ). Further, it was White women who experienced the lowest levels of self-esteem. In addition, research on well-being in adolescents suggests Black and Hispanic girls report higher levels of scholastic competence and lower levels of depression than Black and Hispanic boys and White girls (McCleod & Owens, 2004). Further, Black and Hispanic girls reported higher levels of well-being than minority boys. Therefore, regardless of the injustices Black women may face, it appears Black women are not worse off with respect to self-esteem.

### *Workplace Outcomes*

There is a considerable amount of literature addressing workplace discrimination and economic/wage-related outcomes for Black women. A good deal of this research contradicts the double jeopardy hypothesis. For example, Bergman and Drasgow (2003) found that in a study of women in the U.S. military, race was not

a factor in the rate of sexual harassment reported (White, Asian, Hispanic, Black and Native American women reported equal levels of harassment). In other words, belonging to a racial minority group did not increase gender related harassment.

In addition, Sidanius and Pratto (1999) analyzed U.S. census data starting in 1949 and found that controlling for education, the income ratio between Blacks and Whites has improved. But improvement has been greater for women than men – that is, the difference in income between Black and White women has been smaller since the 1940's than the difference between Black and White men. Other results show that while Black women's advancement in some professions is hindered by being a woman, they are not worse off than White women (Sidanius & Pratto, 1999).

One study that does support the double jeopardy hypothesis involved an analysis of National Employee Survey data (from 1993 and 1997) to assess the joint effects of race and gender on workplace autonomy (Petrie & Roman, 2004). Black women perceived the lowest level of workplace autonomy, followed by Black men, White women and then White men. Further, Black women (28%) were least likely to hold supervisory positions, followed by Black men and White women (35%). Not surprisingly, White men were most likely to be in supervisory roles (44%).

In conclusion, the bulk of the data on workplace experiences do not suggest that Black women are worse off than White women and Black men in many respects, though they may experience worse work-related outcomes related to well-being (i.e., autonomy).

*Other Life Outcomes*

Beyond well-being and work related outcomes, other researchers have looked for evidence of double jeopardy in other domains. For example, Black and White male and females differ in the amount car dealers demand them to pay for cars (Ayres, 1995; as cited in Sidanius & Pratto, 1999). For example, final prices offered to White women were \$215.70 higher than final offers made to White men, and Black women were asked to pay \$446.30 more than White men. But it was Black men who received the worst treatment, being asked to pay \$1,133.60 more for a car than White men.

In addition, within the criminal justice system, there is some data to suggest that while Blacks are more likely to be arrested than Whites; when controlling for variables associated with legal relevance of the crime (e.g., seriousness of the crime), racial discrimination can be found only among Black males (Hood & Cordovil, 1992, as cited in Sidanius & Veniegas, 2000). In other words, Black females are less likely to be discriminated against because of their race within the criminal justice system.

It seems then, that Black women are not worse off as double jeopardy would suggest, in fact it seems when compared to Black men quite the opposite occurs. If double jeopardy were correct, Black women would experience lower self-esteem, less pay, and overall more discrimination, than Black men and White women. However, this does not seem to be the case. Given the inconsistencies in the literature between what double jeopardy predicts and actual realities, some researchers have proposed alternatives.

*An Alternative to Double Jeopardy View*

One alternative hypothesis to double jeopardy is the *ethnic-prominence hypothesis* (Levin, et al., 2002). This perspective suggests that race trumps gender; such that women of color are more likely to expect to experience discrimination because of their race than because of their gender. Levin et al. (2002) reason that 1) people more likely to think of themselves in terms of numerical minorities rather than numerical majorities, and Black women only experience minority status on the race dimension, and 2) relations between the races are more conflicted and based on less contact than relations between men and women. Given this, discrimination based on race should be of greater personal relevance to Black women (or women of color in general). Indeed, Levin et al. found that Black and Latina women perceived equal amounts of discrimination as their male counterparts, while White women perceived more discrimination than White males. Of particular interest, Black and Latina women perceived ethnicity to contribute to overall discrimination, but not gender, while White women perceived gender to effect their overall discrimination and not ethnicity. In other words, both gender and race are not perceived as having contributed to discrimination equally among both Black and Latina women. Jones and Shorter-Gooden (2003) mirrored these effects. When asking Black women if they faced discrimination because they are Black, 90% answered yes, but when they were asked if they faced discrimination because they are women, only 69% answered yes. It seems then that the ways in which Black women categorize themselves (i.e., by



their race or gender) has effects on their perceptions of racial and gender discrimination that is quite inconsistent with the double jeopardy hypothesis.

Given that the research literature generally is unsupportive of double jeopardy, the ethnic prominence hypothesis is one alternative to begin to understand Black women's experiences of discrimination. However, as Levin et al. note, there still may be situations in which either gender or race or both are salient. This suggestion leads to a more basic question about Black women – how are they categorized and perceived by others?

### *Perceptions and Categorizations of Black Women*

The previous sections highlighted life outcomes for Black and White women and men as well as Black women's perceptions of how they are viewed by others. The focus in the present research is on how Black women are *perceived* by others, specifically by Whites, in an attempt to offer an alternative to double jeopardy. One way to address perceptions is to investigate a very basic level of social information processing, categorization. More specifically, are Black women categorized by race, gender, or both? Research has suggested that the way in which we categorize others has effects on our perceptions of others (Zárate & Smith, 1990), and further affects both evaluative and discriminatory reactions (Biernat & Vescio, 1993; Taylor, Falcone, 1982).

Zárate and Smith (1990) asked participants to view photos of Black or White male or female targets and answer questions about category membership. For example, a photo of a Black woman might be presented, and the word “Black” would

appear on the computer screen. The participants' task was to answer "yes" or "no" to indicate whether the face fit the category. Reaction times and errors were measured as indicators of categorization, such that the quicker and more accurate participants were, the more salient the category. Zárate and Smith (1990) found that male targets were more quickly categorized by race than gender, but female targets were more quickly categorized by gender than race (studies 1 and 2). Of particular interest, White (and Black participants in study 1) participants categorized Black targets more slowly than White targets, but especially when the targets were Black women (Study 2). While Black men were categorized 30-40 ms slower than White men, Black women were categorized 90-130 ms slower than White women. These researchers suggested Black women's membership as both women and Black made it more difficult for White participants to correctly categorize them. Further, the data do not suggest that race and gender were both used in the categorization process, as double jeopardy would predict. Instead, race was prominent in the categorization of men, but gender was prominent in the categorization of women. Additionally, there was something "special" about Black women's dual membership that led to slower categorization.

At a broader level, a number of researchers have considered categorization and the effects of crossed or multiple categorizations (Arcuri, 1982; Biernat & Vescio, 1993; Brewer, Dull, & Lui, 1981; Crisp, Walsh, & Hewstone, 2006; Deschamps, 1977; Deschamps & Doise, 1978; Klauer, Ehrenberg, Wegener, 2003; Stangor, Lynch, Duan, & Glass, 1992; van Twuyver & van Knippenberg, 1998; van Twuyver & van Knippenberg, 1999; Vescio, Judd, & Kwan, 2003). Generally, these

studies suggest that people tend to use “superordinate” categories such as race and gender or a combination of categories or “subtypes,” to organize information about others (Biernat & Vescio, 1993; Stangor, et al., 1992; Taylor, Fiske, Etcoff, & Ruderman, 1978). Of these combinations, three patterns relevant to the present study have been proposed (Brewer, Ho, Lee, & Miller, 1987; Hewstone, Islam, & Judd, 1993): Dominance, additivity, and category conjunction.

According to the *category dominance hypothesis*, one category dominates and the subordinate category is ignored (Brewer et al., 1987; Hewstone et al., 1993). Typically category dominance occurs in contexts in which one category is more relevant or informative to the situation and/or the perceiver. For example, as the ethnic prominence hypothesis suggests, Black women’s race is salient over gender and has consequences for Black women’s perceptions of discrimination. Further, Macrae, Bodenhausen, and Milne (1995) argue that context can activate a dominant category while inhibiting the use of the non-dominant category. They had participants watch a video of a Chinese woman either eating noodles from a bowl using chopsticks or putting on lipstick in front of a mirror. Then participants completed a word-identification task, in which they categorized words as either Chinese-related or female-related. When the category “Chinese” was made salient (by the woman eating noodles with chopsticks), participants responded more quickly to traits associated with Chinese than women. The reverse was true when the category “women” (putting on lipstick) was salient. More importantly, Macrae et al. found that words associated

with the non-dominant category (women in the case of the noodle video; Chinese in the case of the lipstick video) were inhibited relative to baseline.

The *additivity* view of dual categorization assumes that both categories are attended to independently and subsequently used when making judgments based on category membership (Brewer et al., 1987; Hewstone, et al., 1993). Statistically speaking two main effects are expected from the additivity model, and typically this model has gained the most support in the cross-categorization literature (for reviews see: Crisp & Hewstone, 2000; Hewstone et al., 1993; Migdal, Hewstone, & Mullen, 1998; Urban & Miller, 1998). For example, Hewstone et al. (1993) found that both religion and nationality had independent effects on evaluations of Muslim and Hindu targets.

A third possible outcome in the cross-categorization literature is *category conjunction*, which is relevant to judgment situations in which the targets may be in- or out-group members on none, one, or both category dimensions. According to this perspective, if a perceiver classifies the target as an ingroup member on both category distinctions, this unique conjunction stands in contrast to all other combinations, which are considered outgroup targets (Brewer et al., 1987; Hewstone, et al., 1993). This model has gained most of its support from studies done in desegregated schools in the United States, where children almost exclusively interact with other children of the same gender and race (Rogers, Miller & Hennigan, 1981; Schofield & Sugar, 1977). In other words, same gender and race ingroups are evaluated positively and all

other categorizations are not considered ingroups even if they are matched on one category (e.g., race or gender).<sup>1</sup>

Typically, cross-categorization research has looked at the *evaluative consequences* of categorizing based on two group memberships (for reviews see: Crisp & Hewstone, 2000; Hewstone et al., 1993; Migdal, Hewstone, & Mullen, 1998; Urban & Miller, 1998). These alternative models of category combination have been less well-investigated when it comes to consideration of a more basic level of perception, categorization--the focus of the present research. Using the cross-categorization work as a framework, however, one can hypothesize that race and gender may be used in categorization in several possible ways: 1) race or gender may dominate, 2) race and gender may exert independent influences on categorization (additivity), 3) race and gender may interact such that Black women are categorized differentially (in terms of reaction time and categorization errors, as described below) than the other category combinations (referred to as *subtyping*), and 4) the perceivers' own group membership may affect which categories are used (category conjunction). The present research uses the "who said what" paradigm to address these possibilities (Taylor, et. al., 1978).

---

<sup>1</sup> A fourth pattern, "hierarchical ordering" assumes that both categories interact and that the effects of one category are dependent on prior categorization of the other category (Brewer et al., 1987; Hewstone, et al., 1993). For example, Brewer et al. (1987) found that Hong Kong children rated other Hong Kong children more positively only when they were of the same sex as the target. Further, race only had an effect on evaluations only when the target was of the same-gender ingroup. In other words, gender was most salient for the children and the effects of race were contingent on first categorizing the target as a gender-ingroup. While this pattern is interesting, the nature of the analysis used in the present study does not allow for a direct test of this model.

### *The “Who Said What” Paradigm*

In addition to the Zárate and Smith (1990) paradigm (reaction time) and that of the cross-categorization researchers (evaluative dependent variables), another common way of studying categorization processes is through the “who said what” paradigm (Taylor et al., 1978). In this task, participants hear a conversation among a group of people in which each statement in the conversation is simultaneously paired with a picture of the person supposedly saying the statement. After the conversation is over, participants are presented with each statement along with a set of foils (Klauer & Wegener, 1998). The participant’s task is to indicate if the statement was part of the earlier conversation or not. If yes, the participant is asked to indicate who said the statement by selecting the appropriate photo. Error patterns are indicative of categorization processes. Typically, researchers are interested in within- versus between-category errors, with the former being more common than the latter. For example, participants are more likely to confuse a statement said by one woman with another woman than they would be to confuse her with a man. The “who said what” task has shown to be relatively noninvasive in that participants are unaware that categorization is being assessed (Klauer et al., 2003). Further, it is assumed to tap into relatively automatic aspects of schematic processing.

Of particular relevance to the present study, Stangor et al., (1992) used the “who said what” task and found that participants categorized using subtypes (an interaction between race and gender) comprised of race and gender (e.g., Black women, White women). In addition, participants used gender as an independent

category, while they did not use race as an independent category (study 1-5), suggesting for those participants gender offered information that was more beneficial than race in that particular context. Within the framework of patterns suggested in cross-categorization research, there was a subtyping effect of gender and race, but gender also acted as the dominant category.

When studying categorization, however, researchers typically study broad within- category and between-category effects. For example, Stangor et al. (1992) found larger within race/within gender errors than any other type indicating participant were most likely to use subtypes (i.e., Black woman vs. Black or woman). Evidence that gender was more likely to be used than race as an independent category was derived from the finding that participants were more likely to make within-gender/between race errors than within-race/between-gender errors. From this analysis however, it is impossible to tell which subtype is most likely to be used. In other words, it may be the case that the subtype “Black woman,” is more likely to be used than the subtype, “White woman.” Further, because only broad category errors are used and not broken down by the race and gender of each target, it is impossible to tell if there are independent effects of race and/or gender that differ by target. For example, it is possible that for Blacks (or Black women or men), race is used as an independent category equally or more so than gender (compared to White targets).

By breaking down the broad types of within- and between-category errors, however, it is possible to analyze which specific types of targets are implicated in each type of error (see Table 1). The current research takes a new approach (adapted

from Klauer et al., 2003) to analyzing “who said what” data in order to understand precisely with whom each target is confused. I can therefore use the who said what task to test if Black women, relative to Black men and White men and women, are more likely to be categorized by their gender alone, race alone, or both gender and race.

### *The Invisibility Hypothesis*

The present research also follows up on Zárate and Smith’s (1990) categorization approach by more fully investigating how Black women are categorized relative to other groups. Looking at reaction time only, Zárate and Smith found that that Black women were categorized more slowly than other targets which may indicate something “special” about the categorization process as it pertains to Black women. It may be that it simply takes more time to categorize two categories rather than one, but all of the targets in this case belonged to two (gender and race) categories. Instead, it could be that Black women as a subtype deviate from the prototypes of both Black (Black men) and women (White women), leading to slower categorization processes. In other words, people may not be as “good” at categorizing Black women as either Black or women as they at categorizing Black men as Black, and White women as White. This prediction might be conceptualized as an *invisibility* approach. The invisibility hypothesis assumes categorization errors should be higher for Black women than White women or Black or White men because they do not fit the “prototype” of “Black” or “woman.”



Invisibility of groups has been discussed in particular relevance to Native Americans (Fryberg & Townsend, in press). Typically it is defined as an absence of representations of oppressed groups and/or individuals. Invisible groups or individuals may be misrepresented or not represented at all (Fryberg & Townsend, in press). During instances of misrepresentation relative invisibility occurs such that the group representation is erroneous or outdated. During instances of absolute invisibility, there are no representations of that particular group in the context being discussed. If Black women are erroneously included under the umbrella of “women” and/or “Black” without reference to their differences compared to White women or Black men, this could represent misrepresentation, or relative invisibility. If Black women are not represented at all in any given context, then this could constitute absolute invisibility. For example, in many instances discourse on women or Blacks in the United States does not include representations or even mention of Black women.

The idea that Black women are “invisible” is not a new one. Feminists have argued Black women’s voices have gone unheard and their rights have been unrecognized within discrimination policies and activism (Bell, 1992; Davis, 1981; King, 1988). Interestingly, this idea has never, to my knowledge, been empirically tested.

The invisibility hypothesis can be characterized as subsidiary to the double jeopardy approach in that it does suggest there is something unique about joint categorization as “female” and “Black”. However, the invisibility approach differs

from double jeopardy in one important way. It does not assume that the double stigmatization of (female) gender and (Black) race is what leads to worse outcomes; instead it assumes that the non-prototypical nature of Black women's dual membership as both women and Black leads to decreased visibility in comparison to White women and Black men.

#### *Overview of the Present Research*

Participants in the present study completed a "who said what" task such that they heard a conversation among eight targets (2 Black females, 2 Black males, 2 White females, 2 White males). They were then asked to match each statement with the person who said it. Typically, researchers using this paradigm measure within- and between-category error totals. Four error types are possible: Within-race/within-gender, within-race/between gender, between-race/within-gender, and between-race/between-gender. For example, if a Black woman said a statement and the participant incorrectly guessed that a White woman said it, this is a between-race/within-gender error. The expectation derived from the literature is that within-category errors should be most common, and between-category errors should be least common (e.g., Taylor et al., 1978). In addition to this global prediction, I was particularly interested in errors involving Black women: Are they implicated in more errors than members of other groups? When involved in errors, with whom are they confused? Thus, for each of the four types of errors, I will delve more specifically into comparisons between Black women and the three other groups. Reaction times

during the “who said what” selection were also measured in attempts to replicate Zárate and Smith (1990).

Additionally, participants in this study were asked to evaluate the targets on traits related to warmth and competence (Fiske, Cuddy, Glick, & Xu, 2002). These two trait dimensions capture much of the space into which judgments of individuals and groups fall, and Fiske et al. have demonstrated that many group stereotypes can be mapped onto these dimensions. For example, women tend to be judged high in warmth but low in competence and men the converse. Black professionals tend to be rated moderate on warmth and high on competence, while poor Blacks are rated low on both competence and warmth. Interestingly, Fiske et al., (2002) did not assess whether these racial stereotypes apply equally well to men and women. In this study, after guessing who said what, participants were asked to rate individual targets on attributes related to both dimensions.

Mirroring typical “who said what” data, I expected that participants would be more likely to make within-category errors than between-category, regardless of the race and gender of the target (e.g., Fiske et al., 2002). But for each type of error, the invisibility hypothesis leads to the prediction that Black women’s statements will be more likely to be confused with other targets than any other group. More specifically, Black women should be *least* likely to be given credit for the statements they said and instead their statements should be incorrectly attributed to other targets. In other words, in line with the literature on invisibility, Black women should be misrepresented the most in the conversation representing relative invisibility (Fryberg

& Townshed, in press). Absolute invisibility, or non-representation, is not predicted in that Black women are in the conversation and therefore are to some extent represented in this context.

To summarize the findings typically discussed in cross-categorization research, four patterns may emerge: 1) race or gender may dominate (category dominance), 2) race and gender may be categorized independently (additivity), 3) race and gender may interact (subtyping), and/or 4) the perceivers' own group membership may affect which categories are used (category conjunction). Of these patterns, it is predicted that whatever overall patterns emerge, Black women should be the most implicated in the specific types of errors. In other words, it is not predicted that one particular pattern should emerge more than others in relation to Black women in particular but instead they should most likely be confused with all other targets compared to White women, Black men, and White men.

Thus, for the within-race/within-gender errors, I expected that more confusions would occur between the two Black women than any other group (see Table 1 for clarification of each possible error type). For the within race/between-gender errors, I expected Black men and women to be confused more than White men and women. Of the between-race/within gender errors, I expected that Black and White women would be more likely to be confused than Black and White men. Finally, of the between-race/between-gender errors, Black women should be more likely to be confused with White men, than White women will be confused with Black men. Participants' gender was examined as a moderator of each error type

## Method

### *Participants and Design*

Participants were 65 predominately White (2 Black, 2 Asian, 3 Hispanic, 3 Multi-racial, and 2 Indian) University of Kansas undergraduate students (40 females; age:  $M = 19.17$   $SD = 1.57$ ). Participation was voluntary and compensated with course credit.

The research used a 2 (gender of participant: male vs. female)  $\times$  2 (race of target: Black and White)  $\times$  2 (gender of target: male and female) mixed design, with participant gender as the only between-subjects variable. Participants were exposed to eight targets depicted having a group discussion: Two Black females, 2 Black males, 2 White females, and 2 White males. These targets were represented via photographs and audiotaped statements (produced by Black and White women and men). Each participant heard the group discussion in the same order, but which target said which statement was randomly varied for each participant. To measure categorization, both within-category (within-race, within-gender) and between-category errors (between-race, between-gender) were assessed. In addition, participants rated each target on attributes relevant to warmth and competence, two dimensions relevant to stereotype content.

### *Materials*

To develop the group discussion, we began by having volunteers role-play conversations about obtaining a job after graduating from college. These conversations were used as a template for the group discussion created for the main

study. Sixteen gender-neutral statements about job-seeking were selected for use in the main study, in which the discussants were described as college senior business students trying to get a job in the corporate world. Example statements included, “So I’ve been going to a lot of interviews lately; they always ask the same questions,” and “Yeah they do, ‘tell me a little about yourself,’ is always the first question” (see Appendix A).

For the group discussion, the 16 statements were distributed such that each target said two statements and each statement was paired with a stimulus face. Headshots of Black and White women and men (2 of each race and gender combination) were taken from NimStim Face Stimulus Set (The MacArthur Foundation Research Network on Early Experience and Brain Development).<sup>2</sup> All faces were categorized as “open-mouth happy faces,” skin color was consistent within races, and all faces were rated as equal in attractiveness and facial expression.

Voices were recorded using volunteers who were of the same race and gender as the photographed targets. The conversation was always presented in the same order; however, the face and voice order was randomly assigned for each participant. Further, all audio voice recordings were always paired with the same target photo. All voice recordings were taken from the appropriate gender and race. In other words, a Black woman photo was always paired with a Black woman’s voice. All experimental procedures were presented using MediaLab programming.

---

<sup>2</sup> Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development. Please contact Nim Tottenham at [tott0006@tc.umn.edu](mailto:tott0006@tc.umn.edu) for more information concerning the stimulus set.

*Procedure*

Participants first signed a consent form, and then were informed they would be participating in a study on group dynamics in which they will be hearing a conversation among eight college business seniors discussing their current and upcoming interviews in the corporate world. They were lead to believe we were interested in “different kinds of observations about the groups.” To cover the true purpose of the experiment and to make the presentation of the faces more believable, participants were told the conversation was taped by audio only, but that we wanted them to feel as engaged in the conversation as possible. Therefore, we chose to show a picture of the person who said each statement simultaneously with the audio of the statement.

Participants then followed the instructions given on the computer and began the “who said what” task. First they heard the conversation with a face simultaneously paired with each statement. Each statement was roughly 7 seconds long and each face was shown for exactly 7 seconds. After the conversation, participants were shown 32 statements, one at a time (the 16 statements that were actually depicted in the conversation plus 16 foils; see Klauer & Wegener, 1998). Participants were first asked if they had heard this statement in the previous conversation (yes or no). If they answered no, a new statement was presented. However, if they answered yes, a screen popped up with the eight pictures of the targets they had just seen with an assigned number under each picture. Participants

were asked to select who said the statement by entering the corresponding number of the target person.

After answering the “who said what” questions, participants were shown each picture, one at a time, and asked to rate the depicted target on a series of seven questions relevant to perceived warmth and competence (Fiske et al. 2002; see Appendix B). For example, to assess competence, participants were asked “How competent is this person” (not at all competent – very competent) and “how successful do you think this person will be in the future” (not at all successful – very successful). To assess warmth, participants were asked “How warm do think this person is” (not at all warm – very warm), and “from the brief conversation you just heard, how much do you like this person?” (not at all – very much). Further, to represent warmth and competence in the workplace they were asked such questions as, “If you had the possibility of working with this person in a job setting, how much do you think you would like working with him or her?” (not at all – very much). Additionally, participants were asked, “how much do you think the race (gender) of this person will affect if he or she will be hired for a job for which he or she is qualified?”(not at all – very much). All nine questions were answered on 7-point rating scales; questions were presented in a randomized order for each participant.

Finally, participants answered a series of demographic questions and were asked to state what they felt the true purpose of the experiment was on the computer. Finally they were debriefed by a research assistant and thanked for their time.



## Results

### *Error Analyses*

Following past “who said what” literature, we hypothesized participants would be most likely to make within-race/within-gender errors regardless of the race and gender of the target. To directly test our main hypotheses, it was predicted that within each specific type of error, Black women would be most implicated. In other words, they would be least likely to be correctly matched with a statement they said. We also tested to see if participant gender moderated these effects. In most cases, participant gender had no effect, and therefore was not included in many of the final analyses. Participant gender will only be discussed when it effected errors or judgments. Analyses were computed using repeated measures Analysis of Variance (ANOVA), followed by paired-sample t-tests for tests of simple effects.

### *Error Totals by Error Type*

We first conducted an analysis to examine the overall rates of within- and between- gender/race errors, collapsing across target race and gender. For within-race/within-gender errors (Black woman confused with Black woman, White man confused with White man, etc.) there is only one possible incorrect target, but for every other error type there are two possible incorrect targets. For example, for between-race/within-gender errors, a White woman could be confused with either of the two Black women, or a Black man could be confused with either of the two White men, for within-race/between-gender errors, a Black woman could be confused with either of the two Black men, or a White man could be confused with either of the two

White women, and for between-gender/between-race errors, a Black woman could be confused with either of the two White men, or a Black man could be confused with either of the two White women. To correct for these unequal possibilities, within-race/within-gender errors were multiplied by two (see Taylor, et. al., 1978).

These four error types were submitted to a race-error (within-between)  $\times$  gender-error (within-between) repeated measures ANOVA. Participants were more likely to make errors within-race ( $M = 2.30$ ,  $SD = 1.59$ ), than between-race ( $M = 1.52$ ,  $SD = 1.59$ ),  $F(1, 64) = 16.91$ ,  $p < .001$ . Further, participants were more likely to make within-gender errors ( $M = 2.39$ ,  $SD = 1.64$ ) than between-gender errors ( $M = 1.44$ ,  $SD = 1.20$ ),  $F(1, 64) = 31.68$ ,  $p < .001$ . Replicating past research, there was an interaction (or subtyping effect) between race and gender errors, such that participants were more likely to make within-race/within-gender errors ( $M = 3.17$ ,  $SD = 1.97$ ) than they were to make within-race/between-gender ( $M = 1.43$ ,  $SD = 1.21$ ), between-race/within-gender ( $M = 1.60$ ,  $SD = 1.32$ ), and between-race/between-gender errors ( $M = 1.45$ ,  $SD = 1.19$ ),  $F(1, 64) = 21.59$ ,  $p < .001$  (see Figure 1). Within-race/within-gender errors differed from all other error types (all  $p$ 's  $< .0001$ ) and there were no differences among within-race/within-gender, between-race/within-gender, and between-race/between-gender errors (all  $ps > .05$ ). This pattern of errors is generally interpreted as evidence of subtyping, in that perceivers categorized (and therefore made more errors) within race/gender category combinations.

*Assessing the Roles of Race and Gender within Each Error Type*

*A note on issues of dependency.* Typically, researchers using the “who said what” paradigm to analyze only error type totals (as we did above) collapsing across the targets’ differentiating factors (e.g., race and gender; see Klauer et. al., 2003). However, we were particularly interested in the effects of race and gender within each error type. While each “who said what” answer is independent (in that each statement could be attributed to any of the eight targets), it is likely participants remember that each target contributed equally to the conversations, and roughly track whom they have previously indicated said a given statement. For example, if a participant indicated that Black woman A said statement X, regardless of if whether this choice was right or wrong; he or she may be less likely to attribute a future statement to Black woman A. The issue is further complicated within each error type. For example, if the participant selects a Black woman but the statement was actually said by a Black man, then when a statement the Black woman actually said appears, participants may be more likely to attribute it to the Black man, effectively equalizing the distribution of their answers to each available target. When we consider within-race/between-gender errors, this dependency would inflate the means of errors from Black woman (original speaker) to Black man (incorrect speaker), and Black man to Black woman.

To deal with this potential dependency, in the analyses of within-race/between-gender, between-race/within-gender, and between-race/between-gender errors, each dependent race-gender dyad were collapsed and independent t-tests were preformed. This was not necessary for the within-race/within-gender errors because

race and gender are contained within each type of error (original to incorrect Black woman speaker) and not compared across the analysis.

*Within-race/within-gender errors.* Of the within-race/within-gender errors it was predicted that more confusions would occur between the two Black women than any other group (i.e., White women, Black men, and White men). As predicted, there was a (marginally) significant target race by target gender interaction,  $F(1, 64) = 3.21, p = .078$  (see Figure 2). Driving this pattern, the two Black female faces ( $M = .52, SD = .64$ ) were marginally more likely to be confused than the two Black male faces ( $M = .32, SD = .50$ ),  $t(64) = 1.86, p = .068$ . White female ( $M = .34, SD = .57$ ) and White male faces ( $M = .40, SD = .52$ ) were equally likely to be confused compared with Black female and male targets.

*Within-race/between-gender errors.* These errors include cases in which statements by a Black woman (BW) are confused with those by a Black man (BM; and vice versa), and statements by a White woman (WW) are confused with those by a White man (WM; and vice versa; see Table 1). We predicted that of the within-race/between-gender errors, Black women and men would be more likely to be confused than White women and men. Collapsing the BW → BM and BM → BW errors as well as the WW → WM and WM → WW errors (see Figure 3) to account for any possible dependency, Black women and men ( $M = .85, SD = .87$ ) were marginally more likely to be confused with each other than White women and men ( $M = .58, SD = .75$ ),  $t(64) = 1.95, p = .055$ .

*Between-race/within-gender errors.* These errors included cases in which statements said by a Black woman are incorrectly attributed to a White woman (and vice versa), and statements said by a Black man are incorrectly matched to a Black man (and vice versa; see Table 1). Of the between-race/within-gender errors, we expected that Black and White women would be more likely to be confused than Black and White men. Again the BW → WW and WW → BW errors as well as the BM → WM and WM → BM errors (see Figure 4) were collapsed to control for any possible dependency. In line with predictions, Black and White women ( $M = .92$ ,  $SD = .91$ ) were marginally more likely to be confused with each other than Black and White men ( $M = .68$ ,  $SD = .85$ ),  $t(64) = 1.71$   $p = .092$ .

*Between-race/between-gender errors.* Finally, the between-race/between-gender errors included confusing statements said by a Black woman with a White man (and vice versa) and confusing statements a White woman with a Black man (and vice versa; see Table 1). To account for any possible dependency, BW → WM and WM → BW errors as well as WW → BM and BM → WW errors were collapsed. Of these errors, given the invisibility hypothesis we predicted that Black women and White men would be more likely to be confused than White women and Black men. Consistent with this prediction, Black women and White men ( $M = .89$ ,  $SD = .77$ ) were more likely to be confused than White women and Black men ( $M = .55$ ,  $SD = .69$ ),  $F(1, 63) = 6.81$ ,  $p = .011$ . In addition, participant gender moderated this effect,  $F(1,63) = 9.00$ ,  $p = .004$ . As depicted in Figure 5, the predicted pattern emerged only among female participants, who were more likely to confuse Black women and White

men ( $M = .98$ ,  $SD = .73$ ) than Black men and White women ( $M = .40$ ,  $SD = .59$ ),  $t(39) = 4.16$ ,  $p < .0001$ . But among male participants, these means did not differ ( $M = .76$ ,  $SD = .83$  for Black women/White men confusions;  $M = .80$ ,  $SD = .76$  for White women/Black men),  $t(39) = -2.96$ ,  $p = .77$ .

### *Accuracy Analysis*

Although the analyses above focused on errors, it is worth noting that of the 16 possible statements, participants correctly attributed 7.11 of them ( $SD = 2.41$ ), for an accuracy rate of 44 %. While participants were correct less than half of the time, they did indicate the statement was said before quite accurately. Out of the 16 possible statements they correctly indicated the statement was said in the conversation 81.5 % of the time ( $M = 13.04$ ,  $SD = 2.05$ ), regardless of if they matched it to the correct speaker.

The complement to the predictions above is that statements made by Black women should be *least* likely to be correctly attributed than those made by any other target. The number of statements (out of 4 possible) correctly assigned to each race/gender combination were submitted to a participant gender  $\times$  target race  $\times$  target gender ANOVA. The predicted target race  $\times$  target gender interaction was not significant,  $F(1, 63) = 1.89$ ,  $p = .174$ , but it was the case that accuracy was lowest for Black women ( $M = 1.60$ ,  $SD = 1.06$ ). Accuracy was highest for Black men ( $M = 1.91$ ,  $SD = .95$ ), with White women ( $M = 1.86$ ,  $SD = 1.04$ ) and White men ( $M = 1.74$ ,  $SD = 1.00$ ) falling between these two groups. Reaching conventional significance levels was the target gender  $\times$  participant gender interaction,  $F(1, 63) = 6.41$ ,  $p = .014$  (see

Figure 6). Female participants were equally likely to be accurate for female ( $M = 1.95$ ,  $SD = .73$ ) and male targets ( $M = 1.81$ ,  $SD = .73$ ;  $t(39) = 1.01$ ,  $p = .318$ ), but male participants were more likely to be accurate when the target was male ( $M = 1.84$ ,  $SD = .79$ ) than when the target was female ( $M = 1.38$ ,  $SD = .78$ ;  $t(24) = 2.26$ ,  $p = .034$ ).

### *Foil Statements*

In general, we predicted that participants would be reasonably accurate in recognizing that a foil was, in fact, a foil (not a statement presented in the earlier conversation). Out of the 16 presented foils, participants correctly labeled them new 91.25% of the time ( $M = 14.60$ ,  $SD = 1.60$ ).

Of the few foil statements errors, it was predicted that participants would be more likely to incorrectly attribute a foil statement to a Black woman than to any other target. To clarify, the invisibility hypothesis predicts Black women would be less likely to be given credit for statements they actually said. While it seems that invisibility would also suggest they would not be given credit for saying the foil statements, such that they should not be visible here either, we actually predict the opposite. In other words, not only should they not be given credit for statements they did say but they should be incorrectly attributed to new statements they never said. At a contextual level, this suggests their voices go unheard (i.e., errors) and are essentially misremembered (i.e., foil statements). We submitted the number of “false alarm” claims attributed to each type of target to a participant gender  $\times$  target race  $\times$  target gender repeated measures ANOVA. As there was no effect of participant

gender, it was dropped from the analyses and a target race  $\times$  target gender repeated measures ANOVA was performed. There was a marginal target gender main effect,  $F(1, 64) = 3.72, p = .058$ , such that participants were more likely to incorrectly guess that a woman said the foil statement ( $M = .41, SD = .67$ ), than a man ( $M = .29, SD = .58$ ). This was qualified by a target race  $\times$  target gender interaction,  $F(1, 64) = 5.43, p = .023$ . As depicted in Figure 7, of the incorrect attributions to female targets, White ( $M = .37, SD = .63$ ) and Black ( $M = .45, SD = .71$ ) females were equally selected. But for incorrect attributions to male targets, White men ( $M = .40, SD = .66$ ) were more likely to be given credit for foil statements than Black men ( $M = .18, SD = .50$ ). Further simple effects tests showed that Black women were more likely to be selected for saying the foil statement than Black men ( $p < .05$ ), but that White women and men were equally likely to be given credit for the foils. Therefore, it seems Black women and White men were generally most likely to (incorrectly) be given credit for foil statements.

### *Reaction Times*

In addition to the categorization errors, we were also able to assess the speed with which “who said what” answers were made. Following past literature on speed of categorization (Zárate & Smith, 1990) we predicted that statements attributed to Black women would be made more slowly than those attributed to White women. It also was predicted that participants would categorize fastest by their own gender ingroup status (i.e., men would categorize male targets more quickly than female targets; women would categorize female targets more quickly than male targets).



Also, given Zarate and Smith found that Whites (and Blacks) categorized Whites faster than Blacks; participants here should also categorize Whites faster than Blacks. It should be noted that Zárate and Smith's task involves matching a gender or race label to a photo, whereas the current task involves matching a statement to a person. In other words, the tasks are very different and require different levels of information processing. Nonetheless, Zárate and Smith's reaction times findings guided the predictions for this judgment task.

When computing the reaction times by gender and race, sparse data becomes a problem. If one wants to assess RTs for judgment errors, the low rate of errors overall makes it likely that there is missing RT data for a given type of error, and the repeated measures analysis is not possible. To deal with this, we substituted the overall mean reaction time ( $M = 6398.90$ ,  $SD = 6725.89$ ) for every missing data point. Reaction times are in milliseconds.

Reaction time data were submitted to a participant gender  $\times$  target race  $\times$  target gender repeated measures ANOVA. Participant gender is only discussed when it affected reaction times.

*Reaction times for incorrect judgments.* There were no differences in reaction times for incorrect attribution of statements made by the four target groups, all  $ps > .151$ .

*Reaction times for correct judgments.* An analysis of correct judgments (e.g., correctly indicating that a statement by a particular Black woman was actually made by that woman) revealed a significant target race main effect,  $F(1, 64) = 7.88$ ,  $p =$

.007. Participants took longer to correctly attribute statements to Black targets ( $M = 6321.69$ ,  $SD = 864.57$ ) than White targets ( $M = 6197.19$ ,  $SD = 675.43$ ). The interaction between race and gender was not significant,  $F(1, 64) = .019$ ,  $p = .891$ . Nonetheless, the four relevant means are depicted in Figure 8. Of particular interest, somewhat in line with Zárate and Smith (1990), simple effects analysis indicated that statements made by Black women ( $M = 6335.66$ ,  $SD = 432.65$ ) were attributed more slowly than those made by White men ( $M = 6176.91$ ,  $SD = 283.33$ ;  $t(64) = 2.48$ ,  $p = .016$ ), and marginally more slowly than those made by White women ( $M = 6217.47$ ,  $SD = 394.21$ ;  $t(64) = -1.77$ ,  $p = .081$ ). Statements made by Black males ( $M = 6307.71$ ,  $SD = 433.91$ ) were attributed more slowly than statements made by White males ( $t(64) = -2.16$ ,  $p = .034$ ). Thus, while Black women's statements were not attributed more slowly than Black male statements ( $p = .716$ ), the general pattern suggests that participants reacted slowest when categorizing Black women's statements.

### *Evaluative Judgments*

To analyze the nine follow-up evaluative questions, all questions except the two questions asking, “how much do you think the race (gender) of this person will affect if he or she will be hired for a job for which he or she is qualified?” were submitted in a principal components analysis, run separately by target race and gender. The questions assessing race and gender were not included in this analysis because they were less about evaluations of the person rather, specific questions regarding the race and gender of the target.

Of the remaining seven follow-up evaluative questions (e.g., judgments of warmth and competence) similar results emerged regardless of the target. In each of these runs, only one factor emerged with an eigenvalue  $> 1.0$  (e.g., 5.24) and factor loadings on this factor were always above .70. Therefore, we assumed all seven questions assessed a single evaluative factor, and therefore we computed an index for each target by averaging the ratings across all 7 questions.

The evaluative index was then submitted to a participant gender  $\times$  target race  $\times$  target gender ANOVA. There was a target race main effect, such that Blacks were overall evaluated more positively ( $M = 5.00$ ,  $SD = .76$ ) than Whites ( $M = 4.79$ ,  $SD = .75$ ),  $F(1, 64) = 14.14$ ,  $p < .0001$ . Further, there was a target gender main effect such that women ( $M = 5.04$ ,  $SD = .80$ ) were evaluated more positively than men ( $M = 4.74$ ,  $SD = .71$ ),  $F(1, 64) = 27.58$ ,  $p < .0001$ . These results were further qualified by a target race by target gender interaction,  $F(1, 64) = 6.34$ ,  $p = .014$ , which is depicted in Figure 9. There was no difference in how positively Black ( $M = 5.08$ ,  $SD = .81$ ) and White women ( $M = 5.00$ ,  $SD = .79$ ) were evaluated,  $t(64) = .837$ ,  $p = .406$ . However, Black men ( $M = 4.91$ ,  $SD = .71$ ) were evaluated more positively than White men ( $M = 4.57$ ,  $SD = .71$ ),  $t(64) = 5.96$ ,  $p < .0001$ . In addition, Black ( $t(64) = 5.98$ ,  $p < .0001$ ) and White ( $t(64) = 5.94$ ,  $p < .0001$ ) women were evaluated more positively than White men, and Black women were rated marginally more positively than Black men ( $t(64) = 1.94$ ,  $p = .056$ ). In other words, White men were evaluated particularly negatively.

Ratings assessing how much the target's race would affect hiring decisions (single item), were submitted to a target race  $\times$  target gender repeated measures

ANOVA (participant gender was not significant). From an ethnic prominence hypothesis standpoint, it could be suggested that Black women and men should be equally affected by race. As anticipated, there was a main effect for target race,  $F(1, 64) = 14.65, p < .0001$ , such that participants indicated Black targets ( $M = 3.72, SD = 1.59$ ) would be more effected by race than White targets ( $M = 3.14, SD = 1.74$ ). Interestingly, while there was an effect of race, these means are below the midpoint on a 7-pt scale, indicating, overall participants did not think race would be an issue in hiring decisions.

Ratings assessing how much the target's gender would effect hiring decisions (single item) were submitted to a target race  $\times$  target gender repeated measures ANOVA (participant gender was not significant). There was a target race main effect,  $F(1, 64) = 5.72, p = .02$ , such that participants thought Black targets ( $M = 3.72, SD = 1.55$ ) would be more affected by gender when being considered for a job they are qualified for than White targets ( $M = 3.42, SD = 1.67$ ). Further, there was a target gender main effect,  $F(1, 64) = 16.63, p < .0001$ , such that participants indicated female targets ( $M = 3.81, SD = 1.62$ ), would be more affected by their gender during hiring situations than male targets ( $M = 3.32, SD = 1.60$ ; see Figure 10). Again these means are relatively low. The target race  $\times$  target gender interaction was not significant, ( $F(1, 64) = .174, p = .678$ ), but gender was rated as playing a bigger role in the hiring of Black women ( $M = 3.94, SD = 1.59$ ), than in the hiring of any other

group (White women  $M = 3.69$ ,  $SD = 1.68$ , Black men  $M = 3.49$ ,  $SD = 1.52$ , White men  $M = 3.16$ ,  $SD = 1.68$ , all  $ps < .065$ ).<sup>3</sup>

### Discussion

This research began with an interest in how Black women are categorized by Whites. Given the lack of empirical support for the double jeopardy hypothesis, the present work introduced another viewpoint, the invisibility hypothesis. Invisibility suggests that Black women represent a special case of non-prototypicality, in that they do not fit the prototype of either “female” or “Black.” As a result, they are less noticed and less accurately categorized; their contributions are confused with those of others. In general, the data were supportive of this hypothesis.

#### *Category Errors*

Replicating past “who said what” effects (e.g., Talyor et al, 1978; Klauer et al., 1998), participants were most likely to use subtyping, confusing targets within race and gender. Some research has suggested that people are also more likely to use gender than any other category, such that for example, between-race/within-gender errors are higher than within-race/between-gender errors (Stangor et al., 1992). While we did not replicate these effects statistically, within-gender/between-race errors ( $M = 1.60$ ) were slightly higher than between-race/within-gender errors ( $M = 1.43$ ). Other researchers have shown that the context in which the “who said what” task is presented can diminish the effects of gender dominance (Klauer et al., 2003). Given

---

<sup>3</sup> I also examined the correlations between errors and evaluative ratings as well as judgments of the extent to which race and gender mattered for hiring each target. Of 240 correlations (4 overall types of errors [WRWG, BRBG, WRBG, BGWR], and 4 forms of each type of error, for 4 types of targets and

that our context was job-related, race may have been equally salient as gender, such that Blacks and women in general face discrimination in the workplace compared to their White male counterparts.

Most important from my perspective, I found the predicted pattern of effects within error type: Black women were most implicated across all error types, suggesting that they were least likely to be given credit for statements they said. Specifically, of the within-race/within-gender errors, Black women were (marginally) more likely to be confused with each other than Black men were to be confused with each other. This suggests participants were more likely to use subtyping (both race and gender) when the target was a Black female compared to when the target was a Black male.

Of the within-race/between-gender errors, Black women and men were marginally more likely to be confused with each other than White women and men. This suggests that race was more important for the categorization of Blacks than Whites. Additionally, given that the participant sample was predominantly White, this pattern of effects could indicate a kind of outgroup homogeneity effect such that outgroup members are perceived as more similar to each other than ingroup members (Ostrom & Sedikides, 1992). However, because I do not have a full ingroup-outgroup design, this hypothesis cannot be directly tested. However, again, Black women were highly implicated in this type of error.

---

3 judgment dimensions), 12 were statistically significant. Clearly there was virtually no relationship between error rates and these judgments.

For the between-race/within-gender errors, Black and White women were marginally more likely to be confused than Black and White men. In other words, gender was more salient when categorizing women regardless of race than when categorizing men, regardless of race. This finding is consistent with Zárate and Smith's (1990) reaction time finding that while women are more likely to be categorized by gender. However, Zárate and Smith also found men were more likely to be categorized by their race and we did not find this effect. It may be that reaction times are very different than categorizing statements, and therefore, Zárate and Smith's findings are not relevant to the present effects. Regardless, gender was more salient, suggesting it was equally informative (Stangor et al., 1992) for Black and White women compared to males.

Finally, of the between-race/between-gender errors, Black women were more likely to be confused with White men than White women were to be confused with Black men, but only among female participants. This may be one of the most perplexing findings. From an ingroup/outgroup perspective, it makes sense that White women would be unlikely to confuse their own race- and gender-ingroup with Black men. However, men were just as likely to confuse Black women with their own ingroup as they were to confuse white women with Black men. Further, White women confused Black women, their own gender-ingroup, with White men, their own race-ingroup. One explanation for this pattern may be that White women do not perceive Black women as part of their gender-ingroup (i.e., Black women are non-prototypical women). At more contextual level, this speaks to the history of the

feminist movement in which, White women ignored Black women (Davis, 1981). This finding is compelling and speaks directly to the invisibility hypothesis. Thus it is supportive of the non-prototypical nature of Black women, such that both White women and men confused Black women with White men who clearly do not represent either of Black women's race or gender categories.

In conclusion, of the various types of category errors, Black women were most implicated. Statements they made were marginally more likely to be attributed to the other Black woman, and to either of the two Black men, and the two White women. Black women were also significantly more likely to be confused with White men than White women were to be confused with Black men, at least by female participants. In sum the data support the invisibility hypothesis, such that Black women were less likely to be given credit for statements they said than any other group. There seems to be something "special" about the categorization of Black women, and it warrants further investigation.

*Category errors: Ruling out an alternative explanation.* One possible criticism of using faces in research on categorization is that similarity between any two faces can lead to categorization errors. For example, if the two Black women look particularly similar they should be more likely to be confused with each other than two faces that do not look like each other. To test this, 48 undergraduate participants (48 female; age,  $M = 19.21$ ,  $SD = 1.95$ ) rated each possible pair of faces on a Likert type scale from 1 (not at all similar) to 7 (very similar). The judged similarity between pairs of faces was low overall, with similarity judged highest among the two



Black men ( $M = 3.69$ ;  $SD = 1.45$ ), and lowest among the two White men ( $M = 1.60$ ;  $SD = .79$ ). The two Black women were judged quite low in similarity ( $M = 2.79$ ;  $SD = 1.53$ ), and the White women as more similar to each other ( $M = 3.38$ ;  $SD = 1.42$ ). The White male mean similarity was significantly lower than that of each of the other groups,  $ts(47) > 6.67$ ,  $ps < .001$ , and Black women's similarity was significantly lower than that of White women or Black men,  $ts(47) < 2.54$ ,  $ps < .02$ . Though there were some differences in within-group similarity, it is clear that the Black female faces were *not* more similar to each other – in fact, were *less* similar to each other – than the Black male or White female faces. More specifically, the similarity of Black women's faces can not explain the effects obtained.

#### *Accuracy and Foil Statements*

It was predicted that of the correct “who said what” matches, participants would be least accurate in correctly attributing Black women's statements. While the predicted interaction did not reach statistical significance, participants were, in general, the least accurate in attributing statements made by Black women. While it seems logical that if Black women accumulated the most errors, they would also be least likely to be correctly matched, it is also the case that these data are not exactly symmetrical. The analysis used to investigate errors was more fine-tuned in that it looked specifically at with whom each target was confused, while the accuracy analysis is an overall score. Also, the error data differ due to dependency in the data, such that an error implicating a Black woman could also implicate another target, potentially reducing the accuracy rate for both groups.

Instead of the predicted interaction between race and gender, I found that female participants were equally accurate in identifying who said what for female and male targets, but male participants were more accurate when the target was male than when the target was female. The race of the target had no effect on accuracy rates. Given that errors are a “better” measure of category perception, these results do not stand in contrast to the invisibility hypothesis but simply indicate that White males are less likely to remember statements said by women in general, which may have implications for gender discrimination.

Of the foil statements participants incorrectly stated were actually said in the conversation, it was predicted that Black women would be most implicated. In other words, it seems plausible that if participants are unlikely to correctly match statements they said, they should conversely be likely to match them with statements they did not actually say. Results indicated that among female targets, participants were equally likely to match foil statements to Black and White women, however, among male targets, participants were more likely to incorrectly match foil statements to White males than they were to match such statements to Black males. Black males were the least likely to be given credit for foil statements. Of particular interest, Black women and White men were generally most likely to incorrectly be given credit for statements they did not even say. At face value this seems to indicate Black women are not invisible but rather as visible as White men such that participants thought they were part of the conversation even if the statements were not actually said. However, taken together with the error data, not only were Black women more likely to be

incorrectly matched with new foil statements, but they were also more likely than any other group to not be given credit for statements they did say. Their “voices” were not only ignored, but they were misremembered, lending more support to the invisibility hypothesis. On the other hand White men were more likely both to be incorrectly matched with new foil statements, as well as correctly matched with statements they did say (compared to Black women). White men were given credit for not only the statements they did say but new statements.

#### *Reaction Time Data*

Based on Zárate and Smith (1990) and the invisibility hypothesis, I expected that statements made by Black women would be attributed more slowly than those made by any other target. I also predicted that attribution would be faster for Whites than Blacks, and that female participants would be quicker in attributing female than male statements (and vice versa for male participants).

There were no differences in reaction times for making errors, but of the correctly matched statements, participants were, as predicted, slower at categorizing statements made by Black than White targets. Of particular interest, of the correct matches, statements made by Black women were attributed more slowly than those made by White males, and marginally more slowly than those made by White females. This suggests that in general, participants had the most difficulty with assigning Black women’s statements. In combination with the error data, these findings are suggestive of Black women’s invisibility, such that participants were slower at remembering which statements they actually said.

*Evaluative Ratings*

With regard to evaluations, I was interested in how Black women would be evaluated on dimensions of warmth and competence. But all of the seven ratings loaded on one factor, and results indicated that females were evaluated more positively than males and Blacks more positively than Whites. Black women were evaluated more positively than Black and White males. However, the group driving the effects was White males, who were evaluated particularly negatively. These results are puzzling. At face value, they suggest White men are victims of double jeopardy – negatively evaluated because of their race and gender! We doubt this is the case. Instead, it is possible that shifting standards were applied, such that the standards for White men were higher than for any other group and therefore, they were rated more negatively (Biernat, Manis, & Nelson, 1991). Another possibility is that participants were trying to be politically correct – favoring women and Blacks (Crosby, Bromley, & Saxe, 1980; Dovidio & Gaertner, 1986). At any rate, Black women were not worse off than any other group in respects to their evaluations; in actuality they were generally rated the most positively (along with White women) suggesting again that their dual minority status does not lead to worse outcomes as double jeopardy would suggest.

Participants were also asked to rate how much they thought the gender and race would affect each target's likelihood of getting a job for which they were qualified. For target's race, participants indicated Black women and men would be most affected by their race in hiring decisions. As mentioned before, this lends

support to the ethnic prominence hypothesis (Levin et al., 2002), such that race is equally important for Black women and men in reference to discrimination. Some proponents of double jeopardy would say that Black women's minority gender status should also influence effects due to race, but this was not supported here.

On the question assessing how much the gender of the target would effect getting a job for which he or she was qualified, female targets and Black targets were rated as most likely to be affected by gender during hiring decisions. While the finding that female targets were rated as more likely to be affected by gender is not surprising, it is somewhat surprising that Black women and men did not differ on this rating. One problem with the framing of the question is that it did not distinguish between gender harming or helping an individual – it could be that gender was perceived to harm Black women but help Black men. Overall, the means on these items were relatively low, such that overall participants did not think race or gender would affect hiring decisions.

#### *Limitations and Future Directions*

The present work offers an alternative to double jeopardy. While there is support for invisibility rather than double jeopardy, the scope of the analyses and interpretation can not go beyond basic social information processing. We only can conclude that Black women's race and gender are not equally salient or comparable to White women's gender and Black men's race, rather Black women represent a special case and may be seen non-prototypical for either gender and race category. Given these findings, future research should both continue addressing the basic level

questions (e.g., are Black women seen as non-prototypical exemplars of “Black” and “women”?), and move into questions beyond basic categorization. For example, currently we are looking at if Black women’s voices are less represented (credit for leadership, work well done, etc.) in the work place than White women and men and Black men.

Further, the present work by no means addresses all the contexts in which Black women may be perceived. It is possible in some context Black women’s race and gender take on very different categorization patterns than found here. Further, the present work does not look at experiences of Black women, but rather is an attempt to see how Black women are perceived with the goal of identifying a new sort of discrimination (i.e., invisibility). It is possible that Black women not only experience worse outcomes relative to White women and Black men in some contexts than others, but this may also map onto how they are perceived by Whites, such that their gender and race categorizations are salient and prototypical.

In addition, we were interested in if participant gender would moderate any of the effects. In some instances gender did interact with the race or gender of the target, however overall, gender did not seem to play much of a role throughout the analyses. One reason for the lack of a gender effect may have been due to lack of power given unequal sample sizes between female and males. While there were 40 female participants there were only 25 male participants. While gender effects were not the sole interest of the present work they are important for understanding processes of categorization and future research should insure equal and adequate sample sizes.

*Implications and Conclusions*

The present paper began by introducing Sojourner Truth's "Ain't I a woman?" speech. Black women have historically been ignored. They were left out of women's rights by White women, and now while racial and gender discrimination are widely studied, they are not included. The present work emphasizes this lack of attention suggesting invisibility may be one way in which Black women are discriminated against, and it may be representative of how they diverge from the prototype of woman and Black. At a basic level, invisibility is supported, and it seems imperative that research on race and gender look at Black women as "Black women" that not only experience different things than their White women and Black male counterparts, but are perceived differently.

## References

- Arcuri, L. (1982). Three patterns of social categorization in attribution memory. *European Journal of Social Psychology, 12*, 271–282.
- Beale, F (1970). Double jeopardy: to be Black and female. In T. Cade (Ed). *The Black woman: An anthology* (pp. 90-100). New York: New American Library.
- Bell, E.L.(1992). Myths, stereotypes, and realities of Black women: A personal reflection. *Journal of Applied Behavioral Science, 28*, 363-376.
- Bell, E. L. J. E., & Nkomo, S. M. (1998). Armoring: Learning to withstand racial oppression. *Journal of Comparative Family Studies, 29*(2), 285-295.
- Bergman, M.E., & Drasgow, F. (2003). Race as a moderator in a model of sexual harassment: An empirical test. *Journal of Occupational Health Psychology, 8*, 131-145.
- Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60*, 485-499.
- Biernat, M., & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology, 29*, 166– 202.
- Binion, V.J. (1990). Psychological androgyn: A black female perspective. *Sex Roles, 22*, 487-507.
- Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology, 41*, 656–670.



- Brewer, M. B., Ho, H.K., Lee, J.Y., & Miller, N. (1987). Social identity and social distance among Hong Kong school children. *Personality and Social Psychology Bulletin*, *13*, 156–165.
- Brown-Collins, A.R., & Sussewell D.R. (1986). The afro-american woman's emerging selves. *The Journal of Black Psychology*, *13*, 1-11.
- Crisp, R.J., Hewstone, M. (2000). Multiple categorization and social identity. In *Social Identity Theory: Trends in theory and research*, Capozza D, Brown R (eds). Sage: Beverly Hills, CA; 149-166.
- Crisp, R.J., Walsh, J., & Hewstone, M. (2006). Crossed categorization in common ingroup contexts, *Personality and Social Psychology Bulletin*, *32*, 1204-1218.
- Crosby, F, Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, *87*, 546-563.
- Davis, A. (1981). *Women, race, and class*. New York: Random House.
- Deschamps, J.C. (1977). Effect of crossing category memberships on quantitative judgment. *European Journal of Social Psychology*, *7*, 517–521.
- Deschamps, J.C., & Doise, W. (1978). Crossed category membership in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 141–158). Cambridge: Cambridge University Press.
- Devine, P.G. & Elliot, A.J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139-1150.

- Dovidio, J.F., & Gaertner, S.L. (1986). Prejudice, discrimination, and racism: Historical trends and contemporary approaches. In J.F. Dovidio & S.L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 1-34). San Diego, CA: Academic Press.
- Essed, P. (1991). *Understanding everyday racism: An interdisciplinary theory*. Newbury Park, CA: Sage.
- Fiske, S.T., Cuddy, A.J.C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878-902.
- Fryberg & Townsend (in press). The Psychology of Invisibility. In Adams, G., Biernat, M., Branscombe, N., Crandall, C.S., & Wrightsman, L.S. (Eds.), *Commemorating Brown: The social psychology of prejudice and discrimination*. Washington, DC: American Psychological Association.
- Hewstone, M., Islam, M. R., & Judd, C. M. (1993). Models of crossed categorization and intergroup relations. *Journal of Personality and Social Psychology*, 64, 779–793.
- Hurtado, A. (1989). Relating to privilege: Seduction and rejection in the subordination of white women and women of color. *Signs: Journal of Women and Culture in Society*, 14, 833-855.
- Jones, C., & Shorter-Gooden, K. (2003). *Shifting: The double lives of Black women in America*. New York: HarperCollins.

- Klauer, K.C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effect patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology, 39*, 332-354.
- Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the “Who said what?” paradigm. *Journal of Personality and Social Psychology, 75*, 1155–1178.
- King, D. K. (1988). Multiple jeopardy, multiple consciousness: The context of a black feminist ideology. *Signs, 14*, 42-72.
- Landrine, H. (1985). Race x class stereotypes of women. *Sex Roles, 13*, 65-75.
- Levin, S., Sinclair, S., Veniegas, R.C., & Taylor, P.L., (2002). Perceived discrimination in the context of multiple group memberships. *Psychological Science, 13*, 557-560.
- Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69*, 397–407.
- Major, B., Barr, L., Zubek, J., and Babey, S.H. (1999). Gender and self-esteem: A meta-analysis. In *Sexism and Stereotypes in Modern Society*, Swann W.B., Langlois, J.H., & Gilbert, L.A. (eds). American Psychological Association: Washington, DC.
- McCleod, J.D., & Owens, T.J. (2004). Psychological well-being in the early life course: Variations by socioeconomic status, gender, and race/ethnicity. *Social Psychology Quarterly, 67*, 257-278.

- Migdal, M. J., Hewstone, M., & Mullen, B. (1998). The effects of crosses categorization on intergroup evaluations: A Meta-analysis. *British Journal of Social Psychology, 37*, 303–324.
- Myers, L. W. (1980). *Black women: Do they cope better?* Englewood Cliffs, NJ: Prentice Hall.
- Niemann, Y. F., Jennings, L. Rozelle, R., Baxter, J.C., & Sullivan, E. (1994). Use of free responses and cluster analysis to determine stereotypes of eight groups. *Personality and Social Psychology Bulletin, 20*, 379-390.
- NimStim Face Stimulus Set. Research Network on Early Experience and Brain Development. Retrieved October 14, 2005 from <http://www.macbrain.org/faces/>.
- Ostrom, T., M., & Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychological Bulletin, 112*, 536-552.
- Petrie, M., & Roman, P.M. (2004). Race and gender differences in workplace autonomy: A research note. *Sociological Inquiry, 74*, 590-603.
- Rodgers, M., Miller, N., & Henningan, K (1981). Cooperative games as an intervention to promote cross-racial acceptance. *American Educational Research Journal, 18*, 513-516.
- Robinson C.R. (1983). Black women: A tradition of self-reliant strength. *Women and Therapy, 2*, 135-144.

- Sanchez-Hucles. J.V. (1997) Jeopardy not bonus status for African American women in the work force: Why does the myth of advantage persist? *American Journal of Community Psychology*, 25, 565-580.
- Scholfield, J., & Sager, H. (1977). Peer interaction patterns in an integrated middle school. *Sociometry*, 40, 130-13.
- Shorter-Gooden. K. (2004). Multiple resistance strategies: How African American women cope with racism and sexism. *Journal of Black Psychology*, 30, 406-425.
- Sidanius. J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Sidanius, J., & Veniegas, R.C. (2000). Gender and race discrimination: The interactive nature of disadvantage. In S. Oskam (Ed.), *Reducing prejudice and discrimination* (pp. 47-69). Mahwah, NJ: Erlbaum.
- Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62, 207-218.
- Taylor, S.E., & Falcone, H. (1982). Cognitive bases of stereotyping: The relationship categorization and prejudice. *Personality and Social Psychology Bulletin*, 8, 426-432.
- Taylor, S.E., Fiske, S.T., Etcoff, N.L., & Ruderman, A.J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, 36, 778-793.

- Thomas, A. J., Witherspoon, K. M., & Speight, S. L. (2004). Toward the development of the stereotypic roles for Black women scale. *Journal of Black Psychology, 30*, 426–442.
- Urban, L. M., & Miller, N. (1998). A theoretical analysis of crossed categorization effects: A meta-analysis. *Journal of Personality and Social Psychology, 74*, 894–908.
- van Twuyver, M., & van Knippenberg, A.V. (1998). Effects of group membership and identification on categorization and subtyping in memory. *European Journal of Social Psychology, 28*, 531-553.
- van Twuyver, M., & van Knippenberg, A.V. (1998). Social categorization as a function of relative group size. *British Journal of Social Psychology, 38*, 135-156.
- Vescio, T.K., Judd, C.M., & Kwan, V.S.Y. (2003). The crossed-categorization hypothesis: Evidence of reductions in the strength of categorization, but not intergroup bias. *Journal of Experimental Social Psychology, 40*, 478-496
- West, C. M. (1995). Mammy, Sapphire, and Jezebel: Historical images of Black women and their implications for psychotherapy. *Psychotherapy, 32*, 458-466.
- Zárate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition, 8*, 161–185.

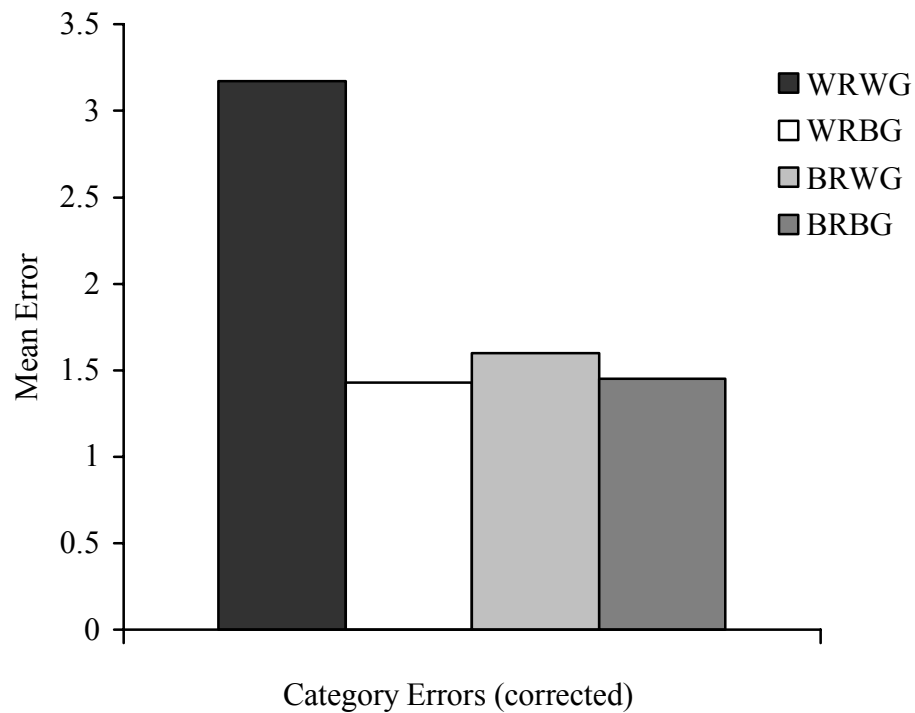
Table 1

*Possible Error Types for Each Target*

| Original Speaker |  | Assigned to         |                                       |             |           | New statements (foils) |
|------------------|--|---------------------|---------------------------------------|-------------|-----------|------------------------|
|                  |  | The correct speaker | An incorrect speaker from the set of: |             |           |                        |
|                  |  | Black Woman         | Black Man                             | White Woman | White Man |                        |
| Black Woman      |  | WRWG                | WRBG                                  | BRWG        | BRBG      |                        |
| Black Man        |  | WRBG                | WRWG                                  | BRBG        | BRWG      |                        |
| White Woman      |  | BRWG                | BRBG                                  | WRWG        | WRBG      |                        |
| White Man        |  | BRBG                | BRWG                                  | WRBG        | WRWG      |                        |

Note: Error types are represented in each cell: Within-race/within-gender (WRWG), within-race/between-gender (WRBG), between-race/within-gender (BRWG), and between-race/between-gender (BRBG). Participants can also attribute the statement to the correct speaker (accuracy), or they can incorrectly indicated a new foil statement was actually said and match that statement to one of the four possible target race and gender combinations.

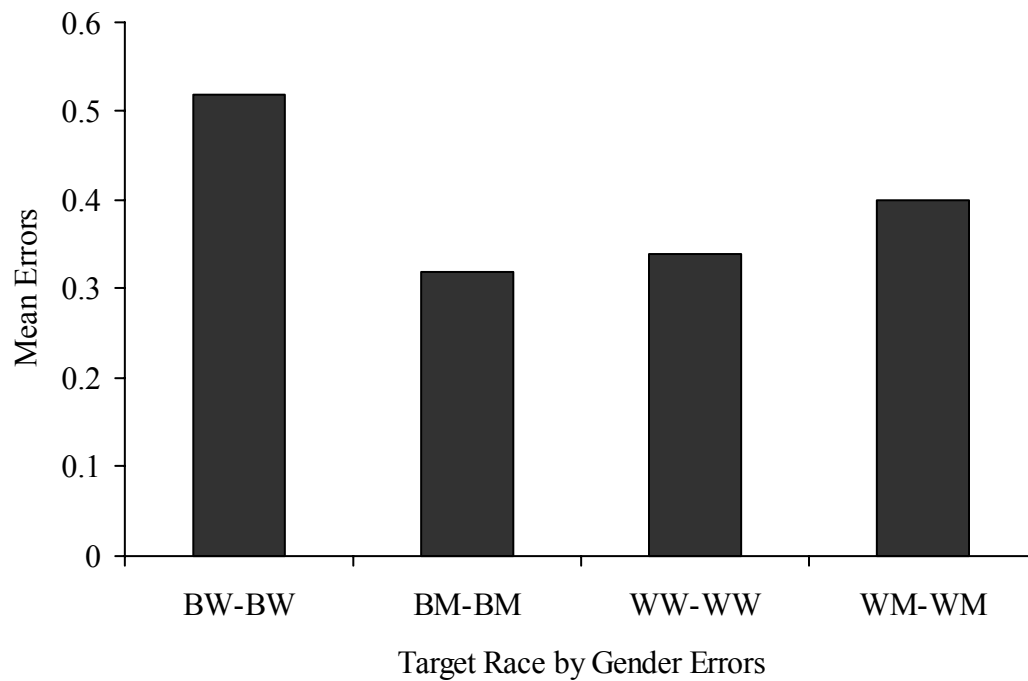
Figure 1. Overall category errors



Note: Within-race/within-gender (WRWG), within-race/between-gender (WRBG), between-race/within-gender (BRWG), and between-race/between-gender (BRBG) are indicated in the legend. To correct for these unequal possibilities, WRWG errors were multiplied by two. The total number of errors possible for each participant is 16.

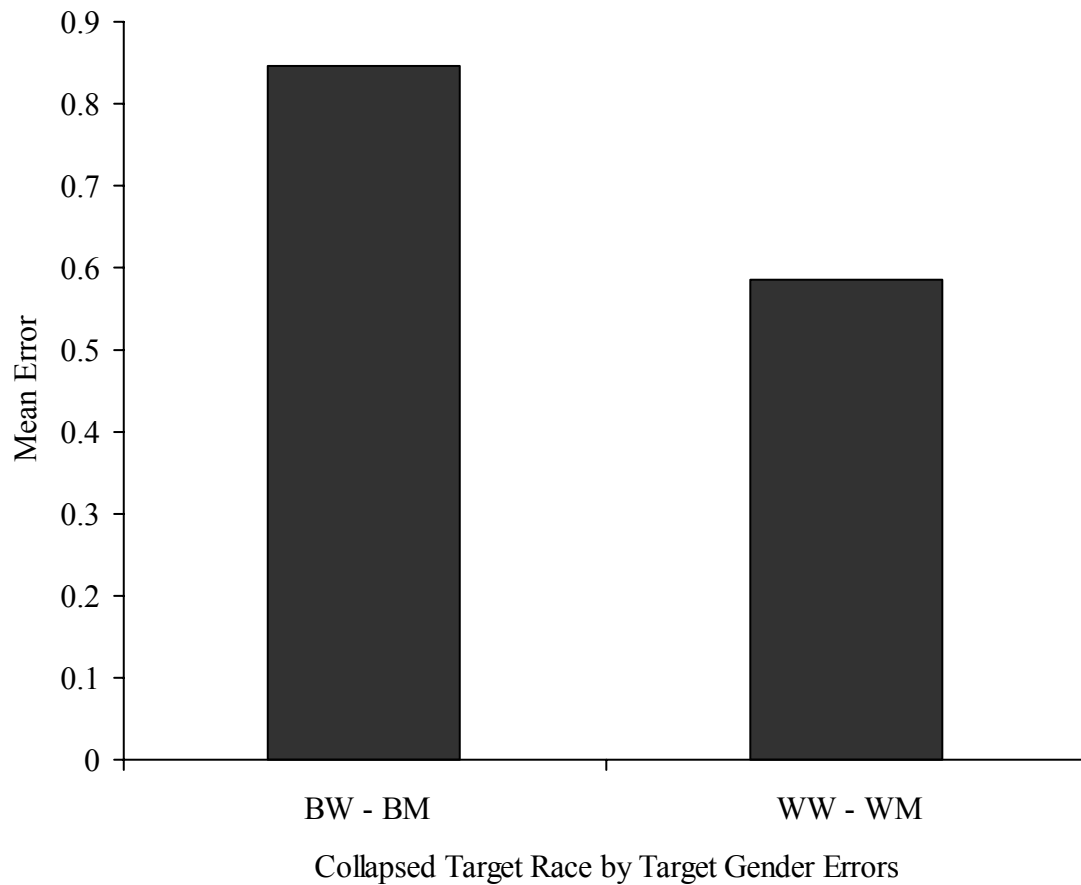


Figure 2. Within-race/within-gender errors



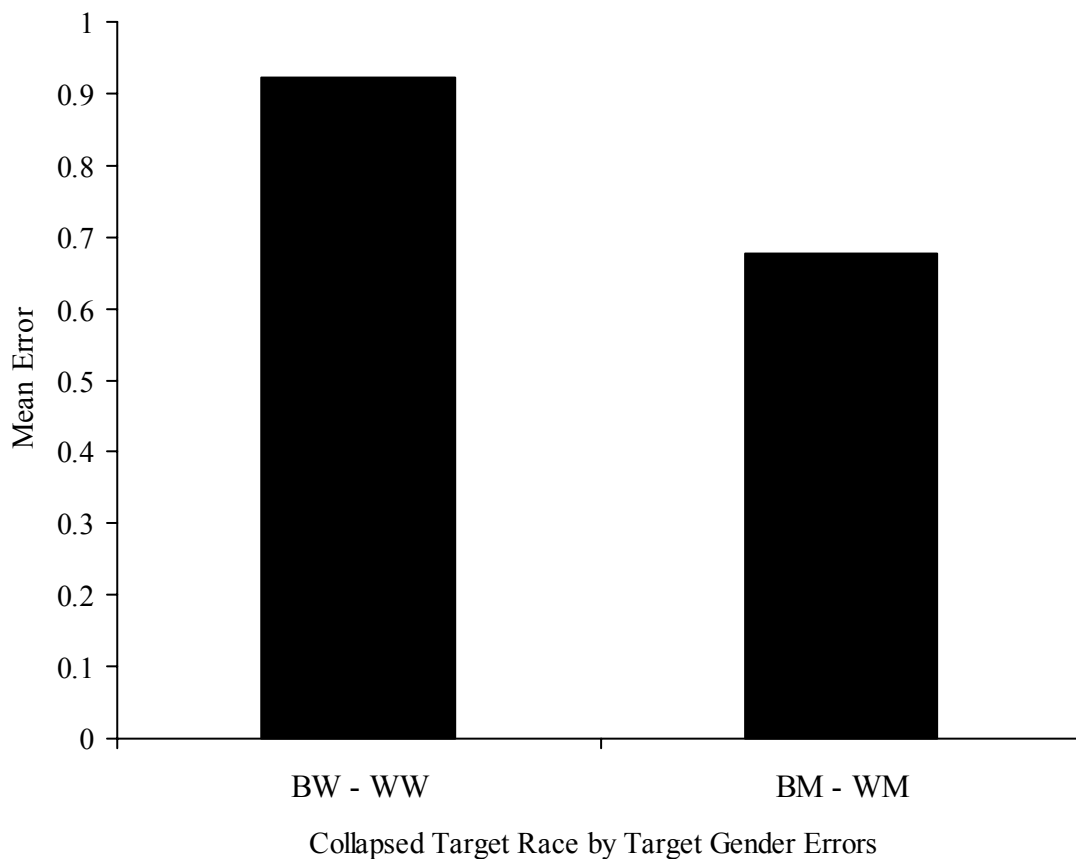
Note: Abbreviation “BW-BW” indicates Black women were confused with other Black women. BM stands for Black men, WW stands for White women, and WM stands for White men. Total number of errors possible per participant is 16. Possible errors per target type is 4.

Figure 3. Within-race/between-gender errors



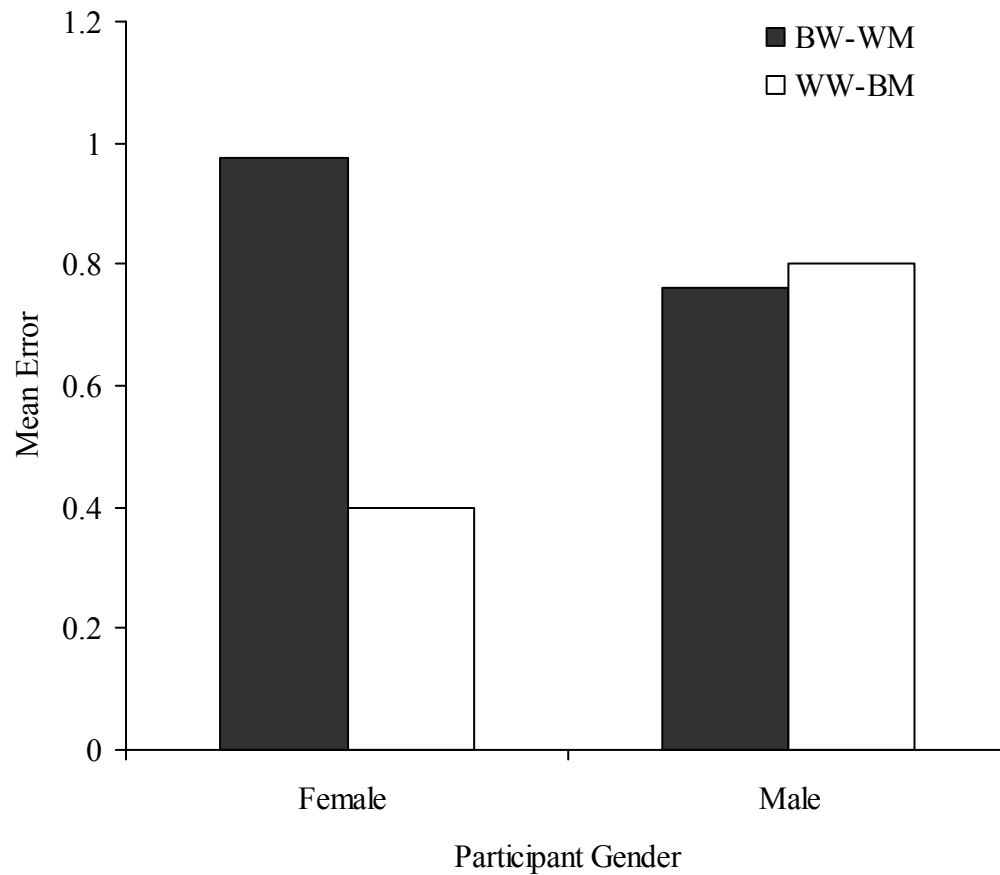
Note: Abbreviation “BW-BM” indicates Black women (BW) and Black men (BM) were incorrectly confused. “WW-WM” indicates White women (WW) and White men (WM) were incorrectly matched. Total number of possible errors per participant is 16. Possible errors per target type is 4.

Figure 4. Between-race/within-gender errors



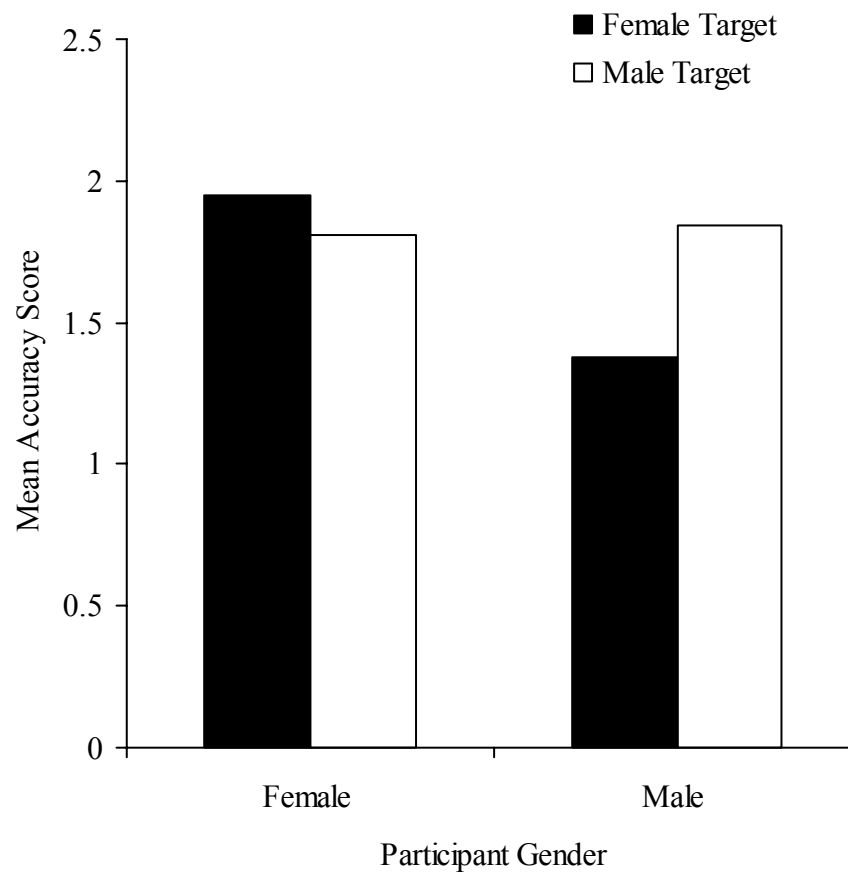
Note: Abbreviation “BW-WW” indicates Black women (BW) and White women (WW) were incorrectly confused. “BM-WM” indicates Black men (BM) and White men (WM) were incorrectly matched. Total number of errors possible per participant is 16. Possible errors per target type is 4.

Figure 5. Participant genders effects on between-race/between-gender errors



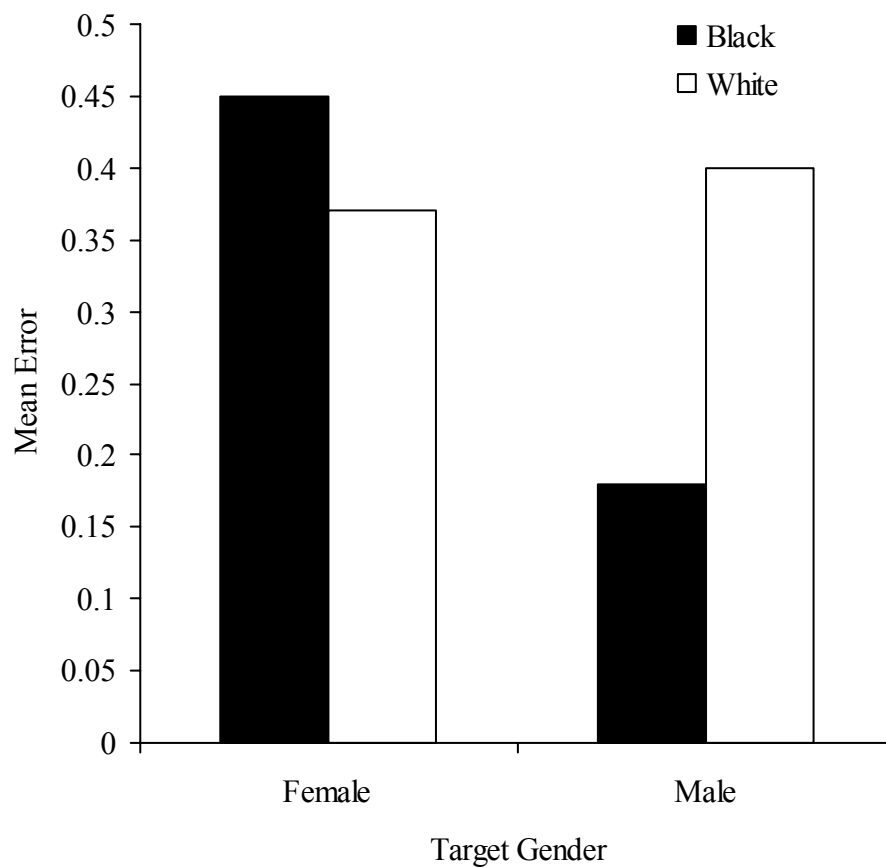
Note: Abbreviation “BW-WM” indicates Black women (BW) and White men (WM) were incorrectly confused. “WW-BM” indicates White women and Black men were incorrectly confused. Total number of errors possible is 16. Possible errors per target type is 4.

Figure 6. Overall accuracy by target and participant gender



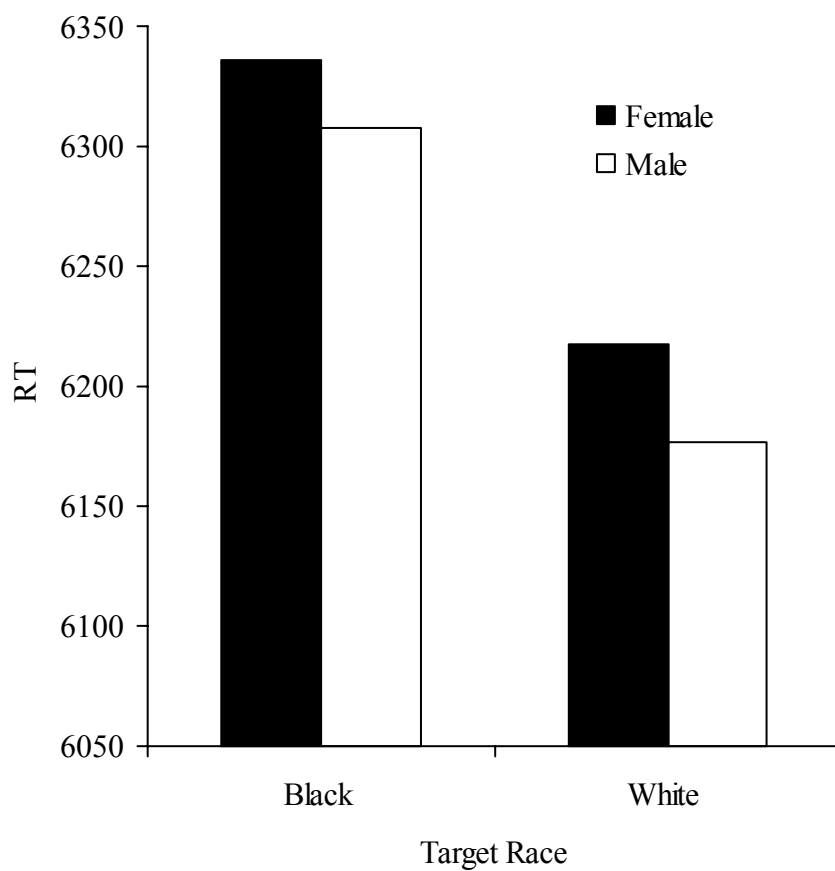
Note: Total possible accuracy score is 16 per participant. Possible errors per target type is 4.

Figure 7. Mean foil statement errors by target race and gender



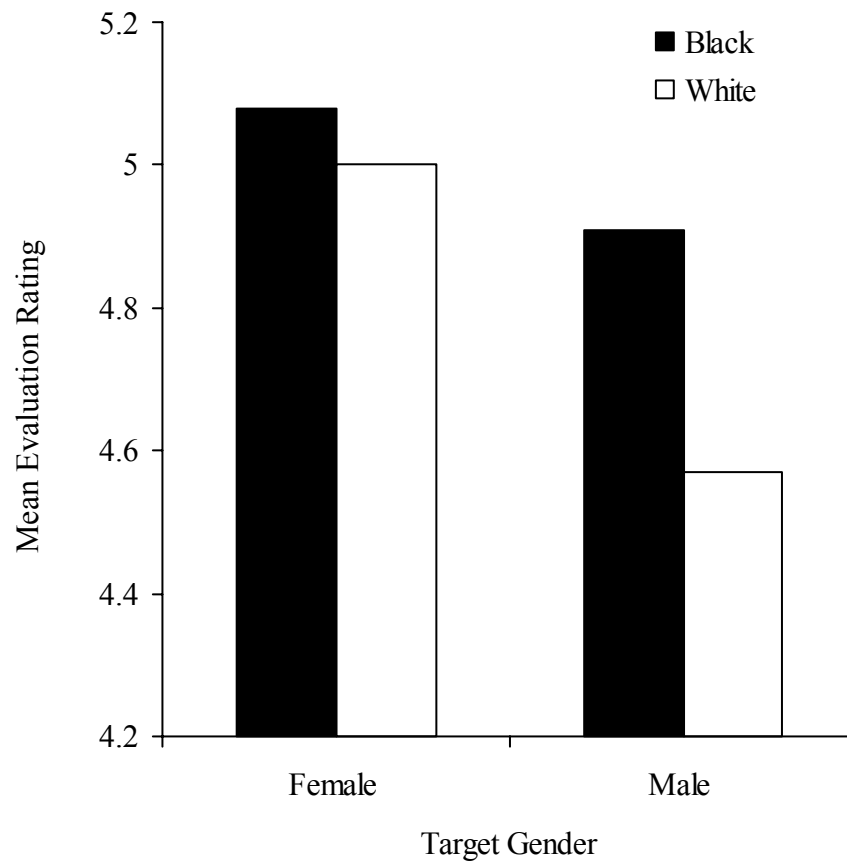
Note: Total possible foil statement errors is 16.

Figure 8. Mean reaction times for correct matches by target race and gender



Note: Higher reaction times (RT) are indicative of slower responses. RTs are in milliseconds.

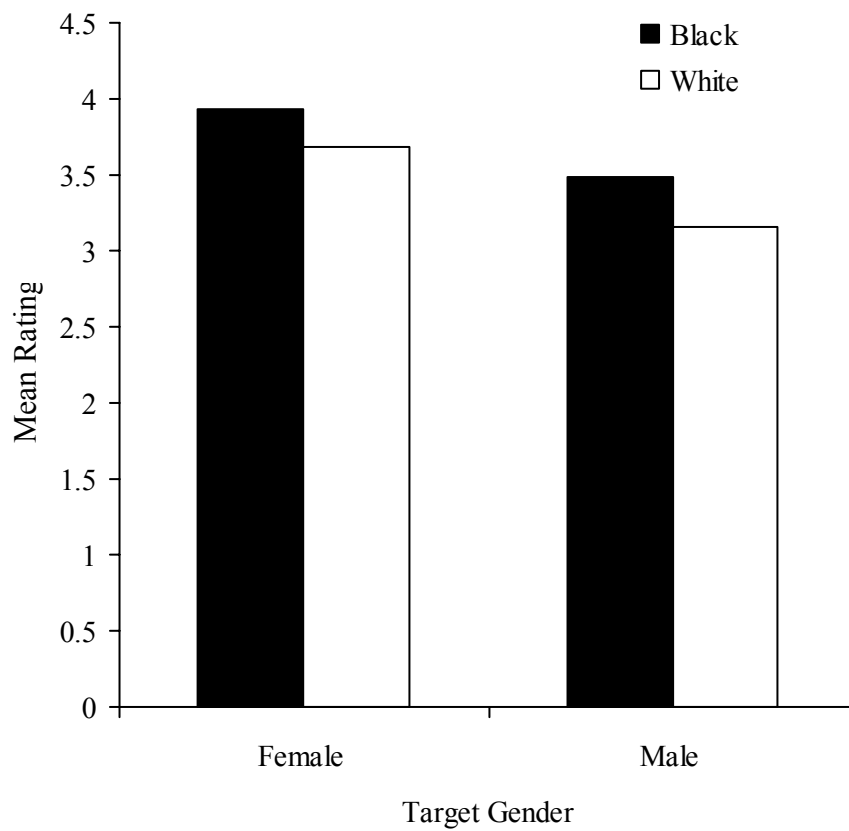
Figure 9. Evaluation ratings by target race and gender



Note: Scale ranges from 1 to 7, higher numbers indicating more positive evaluations.



Figure 10. Gender in hiring decisions by target race and gender



Note: Scale ranges from 1 to 7, higher numbers indicating more influence of gender in hiring decisions.

## Appendix A: “Who Said What” Conversation

1. So I’ve been going to a lot of interviews lately, they always ask the same questions.
2. Yeah they do, “tell me a little about yourself,” is always the first question.
3. I’ve been asked a few times what I would do if I had a problem with a co-worker.
4. Sometimes people try to tailor their answers to the job, but I was told to always answer honestly.
5. Either way, I find it is important to know answers without taking too much time to think about them.
6. That’s why I love phone interviews, you can have notes on how to answer and they never know.
7. Yeah but phone interviews are so awkward; you never know what they think of you when you are done with it.
8. At least you can hide any nervousness. I do sometimes think it is hard to make a good impression on the phone though.
9. The same with in-person interviews. You can think you are good at something but sometimes it doesn’t show in the interview.
10. I found good eye-contact to help me come across as competent and confident in my abilities...even if I’m not.
11. There is a fine line between confident and arrogant though. It’s a toss-up you don’t want to be shy either.
12. A friend of mine just told me he didn’t get a job because he liked the company too much and seemed too desperate.
13. Wouldn’t it be nice if there was a formula for how to act and what to say...it would make things so much easier.
14. I’m so ready to graduate and find a job, but the whole interview thing is so stressful.
15. My friends that have graduated though, say they wish they were still in school...so I say no rush.
16. I thought I wasn’t in a rush, but then I started interviewing and now I just want someone to hire me.

## Appendix B: Evaluative Questions

Directions: Please answer the following questions pertaining to the people in the conversation you just heard. Your answers are completely confidential, so feel free to be completely honest.

1. How competent do you think this person is? (scale 1-not at all competent, to 7-very competent)
2. How warm do you think this person is? (scale 1-not at all warm, to 7-very warm)
3. How likely to you think this person is to get hired for a job in the next 6 months? (scale 1-not at all likely, to 7-very likely)
4. From the brief conversation you just heard, how much do you like this person? (scale 1-not at all, to 7-very much)
5. If you were to interact with this person how likely would you to become friends with him or her? (scale 1-not likely, to 7-very likely)
6. If you had the possibility of working with this person in a job setting, how much do you think you would like working with him or her? (scale 1-wouldn't like it at all, to 7-would like it very much)
7. How successful do you think this person will be in the future? (scale 1-unsuccessful, to 7-very successful)
8. How much do you think the race of this person will effect if he or she will be hired for a job he or she is qualified for? (scale 1-not at all, to 7-very much)
9. How much do you think the gender of this person will effect if he or she will be hired for a job he or she is qualified for? (scale, 1-not at all, to 7-very much)