

UNDERSTANDING THE CONSEQUENCES OF INCOMPETENCE ON THE
VISIBILITY OF ETHNIC MINORITIES IN STEM

AN ABSTRACT

SUBMITTED ON THE NINETEENTH DAY OF DECEMBER 2016

TO THE DEPARTMENT OF PSYCHOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

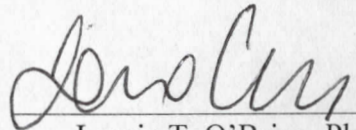
OF

MASTER OF SCIENCE

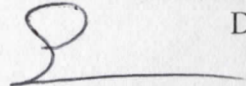
BY

Amanda R. Van Camp

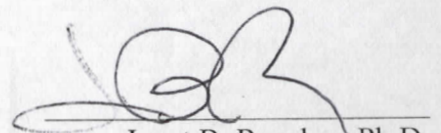
APPROVED:



Laurie T. O'Brien, Ph.D.
Director



Lisa A. Molix, Ph.D.



Janet B. Ruscher, Ph.D.

Abstract

The present research examines barriers Black women may face when entering fields related to science, technology, engineering, and mathematics (i.e., STEM). For example, recent research found that Black women are often forgotten by others because of their non-prototypical group memberships (i.e., intersectional invisibility). Furthermore, previous research using the “Who Said What” paradigm has found that people are more likely to misattribute statements made by Black women to other White men, White women, and Black men. Following this paradigm, the present research sought to determine if invisibility would no longer occur among Black women if they demonstrated incompetence through minor mistakes. It was hypothesized that after a mistake was committed, Black women would become hypervisible and that this increase in visibility would negatively affect participants’ perceptions of Black women’s success in STEM. Results from the current research determined that there was no support for invisibility or hypervisibility. Instead, results suggested that participants were more likely to misattribute statements about Black targets to White targets; however, this only occurred when targets made mistakes. Overall the results suggest a strong concern for participants to appear nonprejudiced and an unwillingness to attribute mistakes to Black targets. The findings have interesting implications for critical evaluations between ethnic groups.

Keywords: intersectional invisibility, identity, STEM

ACKNOWLEDGMENT

I would like to thank my advisor, Dr. Laurie O'Brien, as well as the members of my committee, Dr. Lisa Molix and Dr. Janet Ruscher, for their advice and guidance on this project.



100% cotton fiber

TABLE OF CONTENTS

ACKNOWLEDGEMENT	ii
LIST OF TABLES AND FIGURES	iv
INTRODUCTION	1
METHOD	9
RESULTS	15
DISCUSSION	25
APPENDICES	30
LIST OF REFERENCES	37

SOUTH WORTH

List of Tables and Figures

<u>Title</u>		<u>Page</u>
Table 1	Mean level perceptions of mistake statements	12
Table 2	Descriptive statistics for each error type	14
Table 3	Descriptive statistics for evaluation measures	15
Figure 1	Mean level within- and between- race and gender errors as a function of target race and target gender.	17
Figure 2	Mean level between-race/within-gender errors as a function of target race and target gender.	19
Figure 3	Mean level between-race/within-gender errors as a function of target race and statement type	20
Figure 4	Mean level between-race/between-gender errors as a function of target race and target gender	21
Figure 5	Mean level between-race/between-gender errors as a function of target race and statement type	22
Figure 6	Mean level overall errors as a function of target race and statement type	23
Figure 7	Mean level perception of competence as a function of target race and target gender.	24

Understanding the consequences of incompetence on the visibility of ethnic minorities in STEM

A common stereotype held for women and ethnic minorities is the perception that they are incompetent. For women, these stereotypes about incompetence arise primarily in science and other masculine domains. For example, recent research by Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012) found that women were negatively evaluated by faculty members in science positions at universities because of their gender. Specifically, female applicants for lab manager positions in chemistry, biology, and physics were perceived as less competent, less hireable, offered a lower starting salary, and overall evaluated more negatively than male applicants with the exact same qualifications. Other potential reasons for bias were ruled out because the only difference on the applications was the name of the applicant (John vs. Jennifer).

For Black people, stereotypes about incompetence and a lack of intelligence are even broader (Cuddy, Fiske, & Click, 2007; Devine, 1989). In fact, stereotypes about incompetence and unintelligence for Black people are not domain specific and can occur in any context. Recently, unintelligence was perceived to be one of the top 15 attributes listed for Blacks (Ghavami & Peplau, 2012). An analogous example to Moss-Racusin et al., (2012) is demonstrated in Bertrand and Mullainathan (2004), who submitted resumes to job openings and found that stereotypically Black names (e.g., Lakisha and Jamal) received fewer callbacks for an interview compared to stereotypically White names (e.g., Emily and Greg). Due to the fact that both applications were the same, despite the name, it is clear that implicit biases regarding the abilities of Black people played a role in the company's hiring decisions.

Stereotypically masculine domains are a context in which Black women are negatively stereotyped on the basis of both their race and their gender. Thus, masculine contexts may have unique implications for the treatment and evaluation of Black women. For example in the case of organizational leadership failure, Black female leaders were criticized more than Black male, White female, and White male leaders (Rosette & Livingston, 2012). The authors argued that Black women were criticized more than Black men and Whites in leadership positions because of their double minority status in both their race and gender group memberships.

Researchers have termed the need for some groups to demonstrate competence more so than other targets as the Prove-It-Again Bias (Williams, Phillips, & Hall, 2014). This is especially common for Black women. In a recent report, 77% of Black women in STEM fields recount having to prove their competence multiple times to colleagues and supervisors. This is thought to occur because Black women often have their successes discounted, and their mistakes magnified to the extent that they experience harsher criticism and scrutiny. Therefore, it is possible that demonstrations of incompetence can contribute to the perpetuation of negative stereotypes about the intelligence of females and Black people and this incompetence stereotype may transfer over into the domain of science, which has an expectation of competence.

Expectations about who *IS* a scientist

In recent years, the underrepresentation of women in stereotypically masculine domains, such as science, technology, engineering, and mathematics (i.e., STEM) fields has been a problem of national importance. In 2014, White women made up 13% of graduate students studying physical sciences. This percentage is even smaller for Black

women, who comprised 1% of all physical science graduate students (National Science Foundation, 2015). Because of the gender composition of STEM students, this academic climate may contribute to stereotypes of who is and isn't considered a scientist.

Research has demonstrated that women are not perceived to fit the mental representation of a scientist. One example of this is evidenced in IAT results from Project Implicit. Specifically, this research found evidence to show the difficulty of associating women and science as compared to men and science (Nosek et al., 2009). Similarly, in draw-a-scientist (Chambers, 1983) and draw-a-mathematician (Steele, 2003) paradigms, by 6 years of age, children tend to draw men. This further demonstrates how expectations of gender roles manifest themselves.

However, it is not just an issue limited to gender. To my knowledge, there isn't published research that examines whether Americans expect scientists to be White or Black. Yet, Western cultures have a cultural expectation for White ethnic identity and the male gender. Specifically, there is a tendency to expect that "people" will be White. This effect is demonstrated in Devos and Banaji (2005), which found that people often associate the category "American" with "White." It is possible that this "White male norm" hypothesis may also extend to schematic representations of agentic careers, like scientists (Zarate & Smith, 1990).

Expectations of Stereotypes and Illusory Correlations

Due to the low numbers of Black people and women in masculine domains like STEM, research on illusory correlations may help to explain people's perceptions and expectations about those groups. An illusory correlation is when two uncommon events co-occur and, as a result, the event is distinctive and memorable (Chapman, 1967).

Prior research has found that even a single instance of two rare pairings co-occurring is more memorable than a rare and non-rare pairing. In fact, a single unusual behavior performed by an individual from a rare outgroup receives more processing time, prompts more attributional thinking, and is more memorable (Risen, Gilovich, & Dunning, 2007). For this reason, one-shot illusory correlations can emerge. One-shot illusory correlations are cases in which a single instance of an unusual behavior by a member of a rare group creates an association between the group and the behavior.

Additionally, information and attributional processing can play a role in remembering stereotypically consistent material. For example, Hamilton and Rose (1980) found that stereotypic expectations about a group influenced people's judgments of how often certain qualities described group members. This research suggests that stereotypically consistent stimuli can maintain stereotypic beliefs about a group and aid in memory. Furthermore, people tend to retrieve preexisting judgments about a specific group from memory (Hastie & Park, 1986). This is evidenced in prior research by Howard and Rothbart (1980), which shows that people are more likely to associate the ingroup with favorable expectations and the outgroup with unfavorable expectations following social categorization. Therefore, information is more likely to be remembered when it is relevant to a schema than when it is not.

Previous research has also shown that people tend to infer frequency from accessibility of occurrences in memory (Tversky & Kahneman, 1973). This may help to explain why people may use stereotypes and expectations of groups to inform judgments of how frequently events occur. Therefore, if people have stereotypic expectations that Blacks or women are incompetent, they may overestimate the frequency in which

instances of incompetence have occurred. This suggests that there may be times such as instances in which a mistake (an infrequent event) occurs, may make Black women more memorable.

Although the illusory correlation research maps onto issues of memory for different targets, the illusory correlation research has only looked at people's memories for single identity groups (e.g., Blacks, or women, but not Black women) and a rare event. The illusory correlation research has yet to examine how people will remember information about intersecting identity groups. Therefore, another dominant approach in the literature has been to employ intersectional research methodology to understand people's memories for targets of diverse backgrounds.

The Role of Intersectionality

People have been studying gender in STEM for a while, but not many researchers have taken into consideration the fact that not all women are the same. Although the prototype of a female is White, not all women in STEM are White women (Schug, Alt, & Klauer, 2014). Therefore, one rationale for employing intersectional methodology is to examine how both experiences and perceptions of women differ based on their ethnicity.

Broadly defined, intersectionality is a term used to understand how different social identities (i.e., ethnicity, gender, class, etc.) interact, overlap, and are affected by each other (Purdie-Vaughns & Eibach, 2008). Intersectional methods are used to reduce androcentrism and ethnocentrism in research practices and to capture differences that may otherwise be lost if only one identity is examined and others are held constant (Purdie-Vaughns & Eibach, 2008; Cole, 2009).

Previous empirical social psychological research on intersectionality is limited. It has also not been applied to illusory correlation research. However, it is a growing methodology as more researchers implement different intersectional approaches into their research practices to better understand how multiple subordinate group identities interact.

There are four key intersectionality approaches that dominate the literature. These include the double jeopardy hypothesis (Davis, 1981), the ethnic prominence hypothesis (Levin, Sinclair, Viniegas, & Taylor, 2002), the subordinate male target hypothesis (Sidanius & Pratto, 1999), and the intersectional invisibility hypothesis (Purdie-Vauhngns & Eibach, 2008). Although there are many approaches toward understanding how ethnicity and gender interact, in recent years intersectional invisibility has been the dominant way of thinking about certain groups (e.g., Black women and Asian men).

The intersectional invisibility hypothesis posits that compared to ethnic majority women and ethnic minority men, ethnic minority women are not recognized or valued for their contributions. This occurs as a result of perceived non-prototypicality of individuals who two subordinate identities. For example, Black women are perceived as non-prototypical of both their race and their gender group memberships (Ghavami & Peplau, 2012; Purdie-Vauhngns & Eibach, 2008).

Direct evidence of the non-prototypicality of Black women was demonstrated by Thomas, Dovidio, and West (2014). In a time-sensitive categorization task, participants were slower to respond when associating Black women with “Black” relative to Black men. It was also shown that participants were slower to associate Black women with “Women” relative to White women. These results suggest that Black women are less

prototypical of both “Black” and “Women” categories as compared to Black men and White women respectively.

Further evidence of gendered race theory and the non-prototypicality of Black women was demonstrated in Schug, Alt, and Klauer (2014), which found that people tend to indicate gendered prototypes of racial categories by choosing to use the pronoun “he” when writing about an unidentified Black individual. Similarly, there is evidence for implicit difficulty pairing Black women with race and gender categories (Goff, Thomas, & Jackson, 2008). Specifically, participants were more accurate at guessing the gender of White women than Black women and were more accurate for Black men than Black women.

It is clear that there is evidence of non-prototypicality for Black women. This non-prototypicality can manifest itself in ways that hinder memory. For instance, recent research by Sesko and Biernat (2010) tested participants’ memory for old and new faces (Study 1) as well as the memory for the attributions of statements (Study 2). Taken together the results from these studies indicate that people are less likely to remember Black women’s faces and Black women’s statements as compared to Black men, White men, and White women. Because Black women were more likely to be forgotten relative to other groups, the authors posit that this is due to the non-prototypicality of Black women and demonstrates evidence for intersectional invisibility.

However, Black women may not always be invisible. I hypothesize that Black women will become hypervisible when they make a mistake. There are two possible explanations as to why this may occur. First, stereotypes and expectations about Black women may contribute to the likelihood of remembering a Black woman who made a

mistake. Specifically, because Black women are often stereotyped as unintelligent and incompetent, people may remember their mistakes more because the information is consistent with stereotypes and people tend to remember stereotype consistent information (Hamilton & Rose, 1980). Second, the illusory correlation research suggests that because there are statistically so few Black women in STEM and mistakes are rarely occurring, that the co-occurrence of these two unusual events may contribute to increased memory for Black women who make mistakes (Howard & Rothbart, 1980). The current research will test the hypothesis that Black women will be hypervisible when they make a mistake

Study Overview

The purpose of the current research is to examine the boundary effect of intersectional invisibility and instead, understand the circumstances in which Black women are more memorable or hypervisible. In addition to understanding when Black women will be remembered, the present research also seeks to understand how they are perceived in the domain of STEM. By including evaluative measures, such as competence, belonging, and future in STEM, I will also examine whether Black women are perceived more negatively than White women, White men, and Black men who have also made mistakes.

Using an adapted version of the “Who Said What” paradigm, participants were asked to remember statements made about a person by their supervisor (Taylor, Fiske, Etcoff, & Ruderman, 1978). The adapted “Who Said What” paradigm showed participants photos of people who work as research assistants in a science lab. The photos varied by both race and gender. While viewing the photos, participants also read a

statement about the person made by their lab supervisor. The statements were either a neutral statement or a statement about a mistake that the person made.

It is hypothesized that when Black women make mistakes, they will be more memorable than other targets who made mistakes. Similarly, because I expect Black women to be more memorable, I also expect them to be evaluated more negatively as compared to other targets.

Method

Participants and Design

An a priori power analysis determined that for adequate power, 34 participants would need to be collected. However, because a recent study on intersectional invisibility using a similar paradigm recruited 65 participants (Sesko & Biernat, 2010), the sample was increased to 65. Ultimately, data from 100 Tulane University students was collected in exchange for partial course credit. Three students were flagged by the experimenters as individuals who completed the study abnormally fast (less than 8 minutes). Their data was removed from analyses, leaving a total of 97 participants in the sample. This final sample included 79 women, 15 men, and 3 identified as “other” ranging from ages 18 to 23 ($M = 19.24$, $SD = 1.12$). The sample included 77 Whites, 6 Asians, 5 Blacks, 2 Hispanics, 4 Biracial identified, and 3 identified as “other.”

Upon arrival to the laboratory, a White female experimenter randomly assigned the participant to one of 16 counterbalanced conditions in the 2 (Target Race: White vs. Black) x 2 (Target Gender: Male vs. Female) x 2 (Statement Type: Mistake vs. Neutral) repeated measures design.

Procedure

The paradigm was adapted from Sesko and Biernat's (2010) study that examined participants' inability to accurately distinguish statements made by a Black woman as compared to a Black man, White man, or White woman. Sesko and Biernat (2010) adapted this paradigm from the original "Who Said What" study.

The original "Who Said What" paradigm requires the participant's audio and visual attention because it pairs images of faces with spoken sentences supposedly made by the person in the picture. However, for the current research, the audio portion was eliminated and written sentences appeared on the screen under the face of the individual. Students read that a lab supervisor who managed the student in the picture was responsible for the statement under each picture of the target.

Target Face Selection. The selection criteria for the target faces were decided using the same criteria used in prior research in "Who Said What" paradigms (Sesko & Biernat, 2010). Target faces were selected from the Chicago Face Database (Chicago Face Database Version 2.0, 2015; See appendix A). Faces were selected if they had been rated and perceived equally in terms of prototypicality and attractiveness. All faces selected were perceived to be within the ages of 19-25. This was done so that participants would believe that the targets were college-aged and could be undergraduate research assistants working in science labs. The data from the Chicago Face Database did not include ratings of perceived competence or intellectual ability for any of the targets. Therefore, this information was not included in the selection criteria for target faces.

A total of 16 target faces were selected for the research, creating 2 sets of 8. Participants saw 8 different targets (2 Black males, 2 Black females, 2 White males, and

2 White females). Each target's photo and sentence pair was displayed on the screen for 7s each (photos taken from Chicago Face Database Version 2.0, 2015; see Appendix A). Participants saw each target twice, once with a neutral statement paired with their photo and once with a mistake statement paired with their photo.

Mistake Statement Selection. Seventeen expert raters provided ratings on 12 different mistake statements. The raters provided their responses on a scale ranging from 1 (Not at all) to 5 (Very Much). The scale examined several different aspects of each mistake. These included the extent to which raters thought that the mistake was severe, reflected a potential target's level of competence and laziness, as well as the extent to which the mistake would disrupt science, and be costly in respects to time and money (See Table 1 for means).

Table 1.
Mean level Perceptions of Mistake Statements

Statement	Severity	Competence	Laziness	Disrupting	Cost time	Cost Money
This person didn't zero out the scale before weighing the compound.	3.71	2.88	2.65	3.82	3.294	2.71
This person incorrectly calculated the results and was off by several decimal points.	4.23	3.41	2.82	4.24	4.00	3.47
This person broke expensive glassware	3.71	3.29	2.18	3.65	3.59	2.76
This person misidentified an unlabeled test tube.	3.41	2.24	1.94	2.65	3.18	4.47
This person didn't attend a mandatory safety training course.	3.82	2.76	2.06	4.06	3.88	3.06
This person didn't clean up the lab station before an inspection.	3.59	3.47	4.41	2.59	3.29	2.71
This person used the wrong automatic pipette and measured out the wrong amount of solution.	3.94	3.82	2.82	3.94	3.71	2.88
This person forgot to pre-weigh weigh boats/filter paper before measuring mass.	3.41	3.88	3.70	3.76	3.48	2.71

From the 12 initial mistake statements I deleted 4 statements that elicited extreme responses well above the mean on multiple items. I used the remaining 8 items as stimuli materials for the current study and counterbalanced mistakes among targets (See Appendix B).

Although statements were randomly presented to participants, it was never the case that participants saw related statement pairings. For instance, despite the fact that there is a neutral statement that says, “This person ordered new glassware for the lab” it was never paired with a mistake statement such as, “This person broke expensive glassware.” This ensures that participants did not draw alternative conclusions from the mistake statements.

After the photo sentence pairs, participants completed the memory test. During the memory test participants were shown a total of 32 statements. The statements included 16 original statements that were presented with the photo-sentence pairs as well as 16 foil statements that were not included earlier in the photo sentence pair portion of the study (Klauer & Wegener, 1998; See Appendix B, C, and D). Consistent with the “Who Said What” procedure, participants were first asked if they remember seeing the statement before. They were then presented with the dichotomous decision to select “Yes” or “No” on the computer screen. If participants selected “No,” they were provided with the next statement. If participants selected “Yes,” they were shown a screen with the pictures of the 8 targets they had previously seen. Each picture had a letter under each image. Participants were asked to select the image of the person with which the statement was associated. Following the completion of the task, participants were asked to complete several evaluation measures for each target.

Measures

Accuracy. The key dependent variable of interest was how accurate the participant was in remembering which target was associated with which statement. This was determined by the number and type of errors made by participants. First, I calculated mistakes that were due to people forgetting a statement that had previously been seen. Next, I examined four types of possible misattribution errors that can be made. These include errors that are within-race/within-gender (e.g., mistaking a Black woman for another Black woman), within-race/between-gender (e.g., mistaking a Black woman for a Black man), between-race/within-gender (e.g., mistaking a Black woman for a White woman), and between-race/between-gender (e.g., mistaking a Black woman for a White man). See Table 2 for descriptive statistics for each error type.

Table 2.
Descriptive Statistics for error types

Variables	Mean	SD
Total Forgotten	3.09	1.97
Within-Race/Within-Gender	3.22	2.73
Within-Race/Between-Gender	1.44	1.15
Between-Race/Within-Gender.	1.66	1.39
Between-Race/Between-Gender	1.35	1.13

Note. Within-Race/Within-Gender errors have been corrected consistent with Taylor et al. (1978).

Evaluation of Target. After the memory test was complete, participants were also asked to evaluate each target on an array of items related to how they perceived the target to fit into STEM fields. Examples of these subscales formed include how competent each target was perceived to be in STEM, the extent to which the target belonged in STEM, and the extent to which the target has a future in STEM, all of which

were shown a scale of 1 (Not at all) to 5 (Very much). See Table 3 for descriptive statistics for evaluation measures.

Competence. Competence was measured with six items, which examined the extent to which participants perceived the target as capable and skilled. Example items include, “How competent is the student?” and “How skillful is the student?” This scale yielded a highly reliable index ($\alpha = .92$).

Belonging. Belonging in STEM was measured with three items to examine the extent to which the participant perceived the target as belonging in STEM. Example items include, “This student belongs in science” and “Other people in science are a lot like this student.” This scale yielded a reliable index ($\alpha = .78$).

Future in STEM. Lastly, participants’ perceptions of the target pursuing a future in science were measured with three items. Example items include, “How likely is the student to pursue a graduate degree in science?” and “How likely is the student to be successful in science?” This scale yielded an extremely reliable index ($\alpha = .95$).

Table 3.
Descriptive Statistics for Evaluation Measures

Variables	Mean	SD
Competence	3.70	.57
Belonging	4.04	.54
Future in STEM	3.37	.68

Results

Accuracy

Errors for statements that were seen but not remembered. First, I tested whether participants disproportionately forgot statements that were made about one specific race or gender group. Errors that resulted from forgetting that a statement had

been seen were submitted to a 2(Target Race: Black vs. White) x 2(Target Gender: Female vs. Male) x 2(Statement Type: Mistake vs. Neutral) repeated measures ANOVA. Results indicated a marginally significant main effect of race, $F(1, 96) = 3.09, p = .08, \eta^2_p = .03$, such that people were more likely to forget statements associated with Black targets ($M = .43, SD = .57$) than White targets ($M = .35, SD = .57$). There was not a significant main effect of gender or statement type. There were also no significant interactions ($F_s < 1$).

Within- and between-race and gender errors: demonstrating the use of race and gender categories. Consistent with Sesko and Biernat (2010), the next step was to compute four different misattribution error types. All means represent counts of number of errors. For within-race/between-gender, between-race/within-gender, and between-race/between-gender errors, there are two possible incorrect targets. However, for within-race/within-gender errors there is only one possible incorrect target. To correct for the inequity in this error type, a correction has been applied to the within-race/within-gender errors. Therefore, we multiplied this error type by two (Taylor et al., 1978).

Errors were submitted to a 2 (Race error: Within vs. Between) x 2 (Gender error: Within vs. Between) x 2 (Statement type: Mistake vs. Neutral) repeated measures ANOVA. Results indicated that there was a significant main effect of race, $F(1, 96) = 40.98, p < .001, \eta^2_p = .17$, such that people made more within-race errors ($M = 1.16, SE = .08$) than between-race errors ($M = .75, SE = .05$). There was also a significant main effect of gender, $F(1, 96) = 19.64, p < .001, \eta^2_p = .30$, such that people made more within-gender errors ($M = 1.22, SE = .08$) than between-gender errors ($M = .70, SE = .04$). There was also a trending effect of Statement Type, $F(1, 96) = 2.71, p = .10, \eta^2_p =$

.03, such that participants were more likely to make errors remembering mistake statements ($M = 1.03, SE = .06$) than neutral statements ($M = .89, SE = .06$). Furthermore, there was a significant race x gender error interaction, $F(1, 96) = 19.76, p < .001, \eta^2_p = .17$.

For within-race errors, people were more likely to make within-race/within-gender errors ($M = 1.61, SE = .14$) than within-race/between-gender errors ($M = .72, SE = .06$), $F(1, 96) = 25.94, p < .001$. For between-race errors, people were just as likely to make between-race/within-gender ($M = .83, SE = .07$) as between-race/between-gender errors ($M = .68, SE = .06$), $F(1, 96) = .29, p = .59, ns$ (see *Figure 1*).

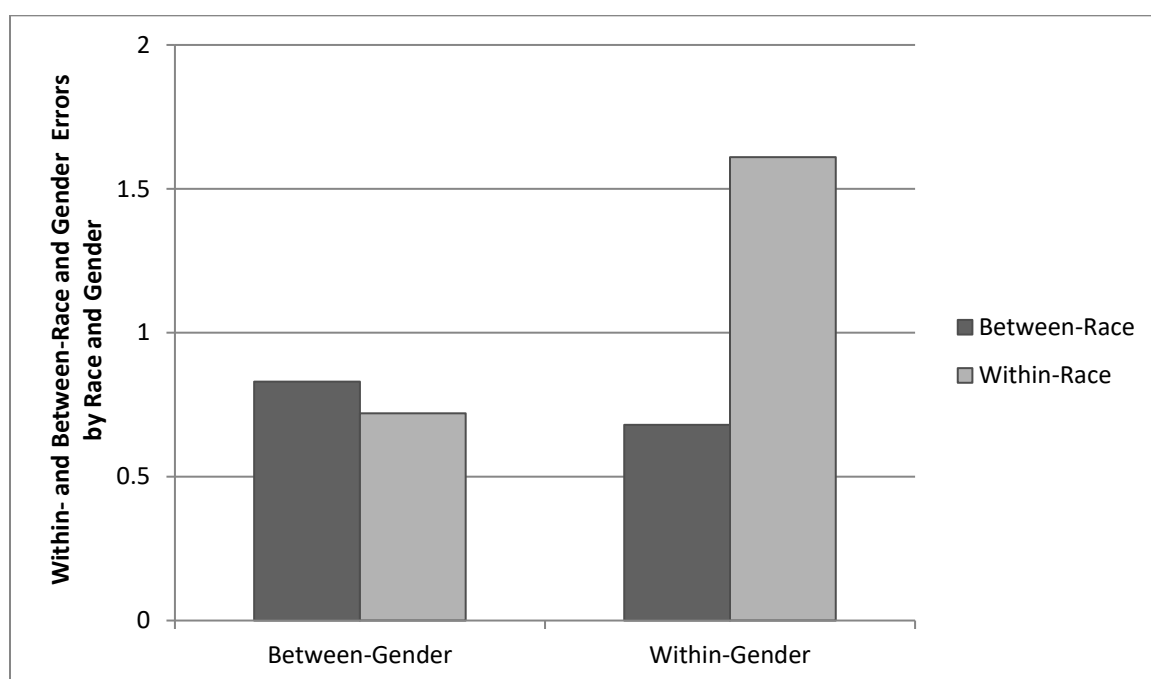


Figure 1. Within- and Between-race and gender error types.

Errors involving Black women versus others. The key test of the hypothesis was whether Black women had the least errors associated with them, due to becoming hypervisible. Thus, the effect of target race and target gender within each possible error type was examined using a series of 2 (Target Race: Black vs. White) x 2 (Target

Gender: Female vs. Male) x 2 (Statement Type: Mistake vs. Neutral) repeated measures ANOVAs.

Within-race/within-gender errors. For within-race/within-gender errors, results indicated that there was not an effect of race, gender, or statement type. There were also no significant interactions ($F_s < 1$).

Within-race/between-gender errors. Next, within-race/between-gender errors were examined. Results indicated that there was not a significant main effect of race $F(1, 96) = 2.45, p = .12, \eta^2_p = .01, ns$. There was also no effect for gender ($F < 1$), or statement type $F(1, 96) = 1.52, p = .22, \eta^2_p = .02, ns$. Additionally, there were no significant interactions between race, gender, and statement type.

Between-race/within-gender errors. The analysis of the between-race/within-gender errors did not yield a significant main effect of race, gender, or statement type ($F_s < 1$). However, there was a significant race x gender interaction $F(1, 96) = 6.84, p = .01, \eta^2_p = .07$ (See Figure 2).

Among women, the simple effect of race was significant such that people were more likely to make between-race/within-gender errors for Black women ($M = .25, SE = .03$) than for White women ($M = .16, SE = .03$), $F(1, 96) = 5.11, p = .03, \eta^2_p = .05$. In other words, people were more likely to attribute a statement associated with a Black woman to a White woman than they were to attribute a statement associated with a White woman to a Black woman. However, the difference between Black men ($M = .19, SE = .03$) and White men ($M = .24, SE = .03$) on between-race/within-gender errors was not significant, $F(1, 96) = 1.93, p = .17, \eta^2_p = .02 ns$.

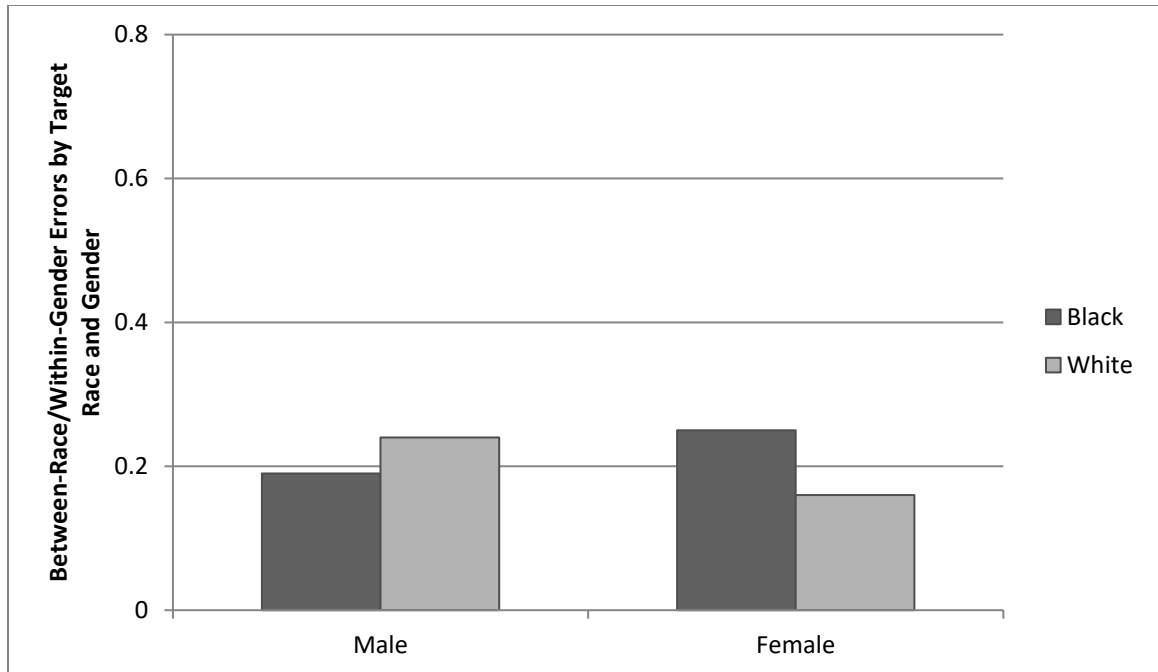


Figure 2. Between-Race/Within-Gender errors by target race and gender.

In addition, there was also a significant race x statement type interaction $F(1, 96) = 6.83, p = .01, \eta^2_p = .07$ (See Figure 3). The simple effect of race was significant such that people were more likely to make between-race/within-gender errors for mistake statements about Black targets ($M = .27, SE = .03$), as compared to mistake statements about White targets ($M = .17, SE = .03$), $F(1, 96) = 6.20, p = .02, \eta^2_p = .06$. The difference between Black ($M = .16, SE = .03$) and White ($M = .23, SE = .03$) targets for neutral statements was not significant, $F(1, 96) = 2.42, p = .12, \eta^2_p = .03, ns$.

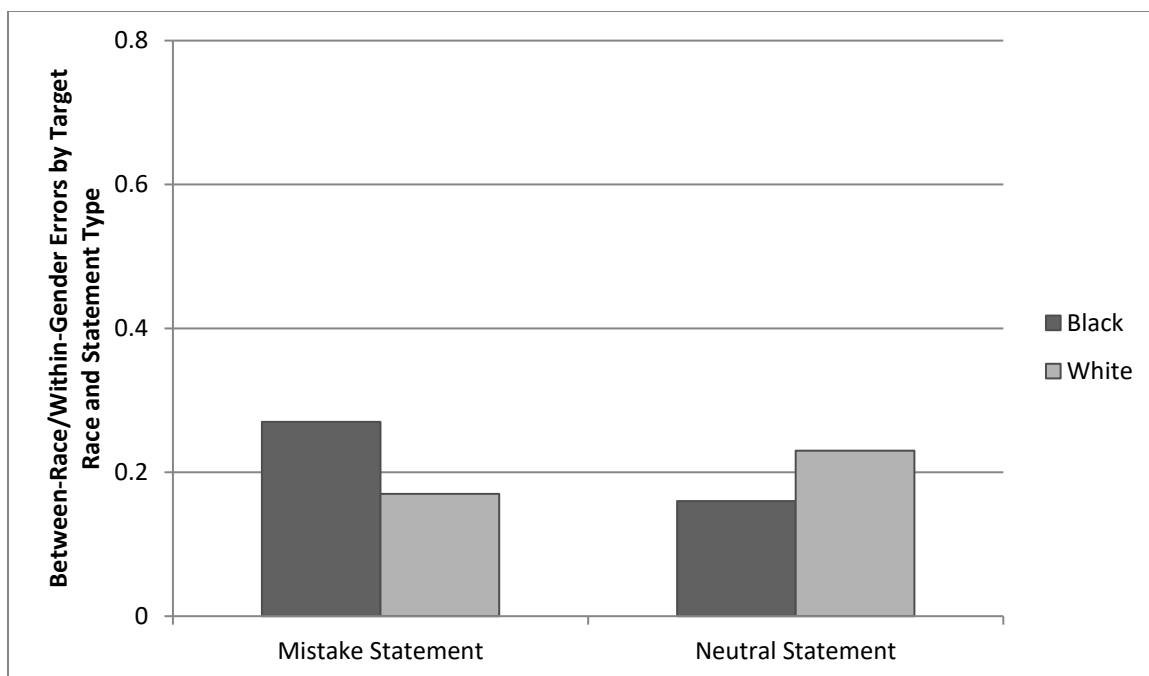


Figure 3. Between-Race/Within-Gender errors by target race and statement type.

Between-race/between-gender errors. The analysis of the between-race/between-gender errors yielded no significant main effects of race, gender, or statement type ($F_s < 1$). There was, however, there was a marginally significant race x gender interaction, $F(1, 96) = 3.33, p = .07, \eta^2_p = .03$. See Figure 4.

An analysis of simple effects revealed that participants were somewhat more likely to make between-race/between-gender errors for Black men ($M = .21, SE = .03$) as compared to White men ($M = .15, SE = .03$), $F(1, 96) = 2.7, p = .09, \eta^2_p = .03$. That is, people were somewhat more likely to attribute a statement associated with a Black man to a White woman than they were to attribute a statement associated with a White man to a Black woman. However, there were no differences in between-race/between-gender errors for women, ($F < 1$)

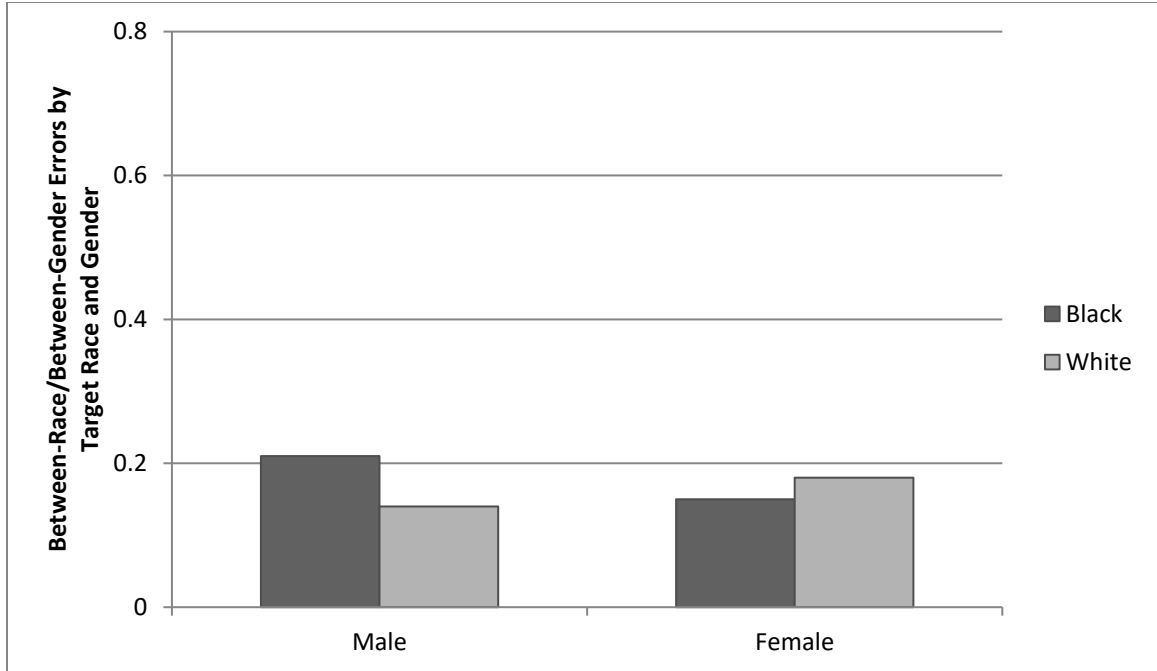


Figure 4. Between-Race/Target Race and Between-Gender/Gender errors by target race and gender.

In addition, there was also a significant race x statement interaction $F(1, 96) = 6.15, p = .02, \eta^2_p = .06$ (See Figure 5). An analysis of the simple effects test revealed that participants were more likely to make between-race/between-gender errors for mistake statements about Black targets ($M = .21, SE = .03$) as compared to mistake statements about White targets ($M = .12, SE = .02, F(1, 96) = 6.22, p = .01$). There were no significant differences between Black and White targets for neutral statements, ($F < 1$).

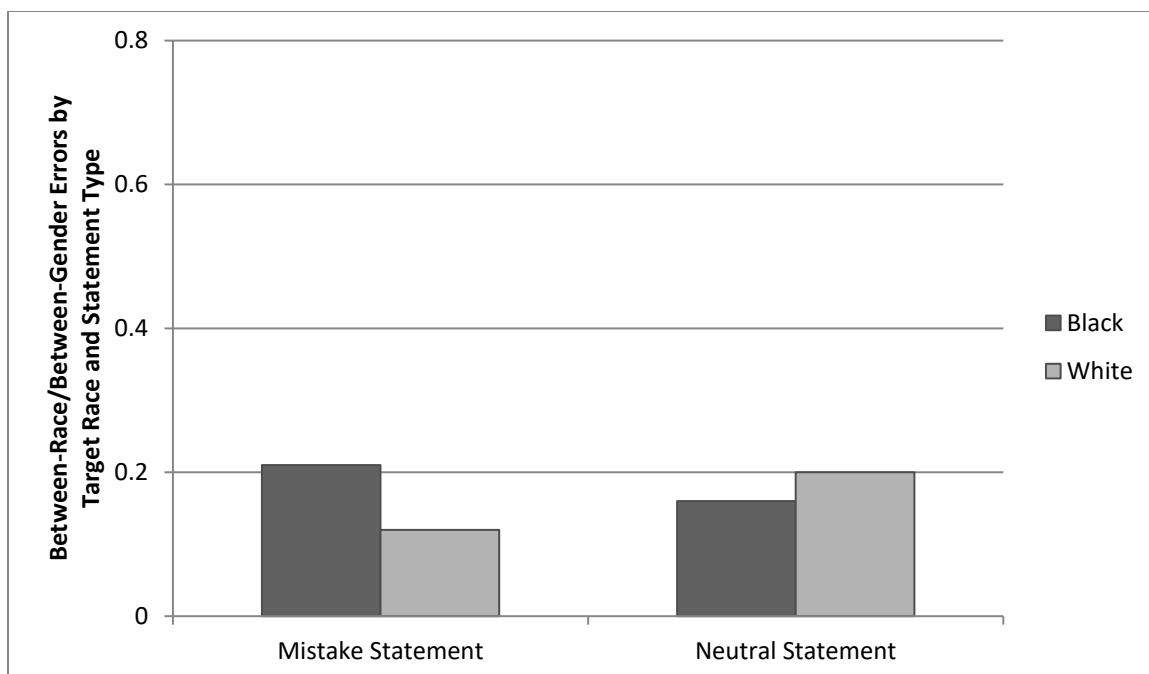


Figure 5. Between-Race/Between-Gender errors by target race and statement type.

Overall errors by target race and gender. Finally, the overall number of errors made for each type of target, regardless of error type, was analyzed. This was completed by first summing the four different error types. For example, for Black female targets, I summed within-race/within-gender, within-race/between-gender, between-race/within-gender, and between-race/between-gender errors. The results from this analysis yielded no significant main effects of race, gender, or statement type ($F_s < 1$). Yet, there was a significant race x statement type interaction $F(1, 96) = 7.22, p < .01, \eta^2_p = .07$.

The simple effects test revealed that participants were more likely to make errors for mistake statements about Black targets ($M = .89, SE = .06$) as compared to mistake statements about White targets ($M = .72, SE = .06$), $F(1, 96) = 6.44, p = .01, \eta^2_p = .06$. Neutral statements were not significantly different, $F(1, 96) = 2.34, p = .13, \eta^2_p = .02, ns$.

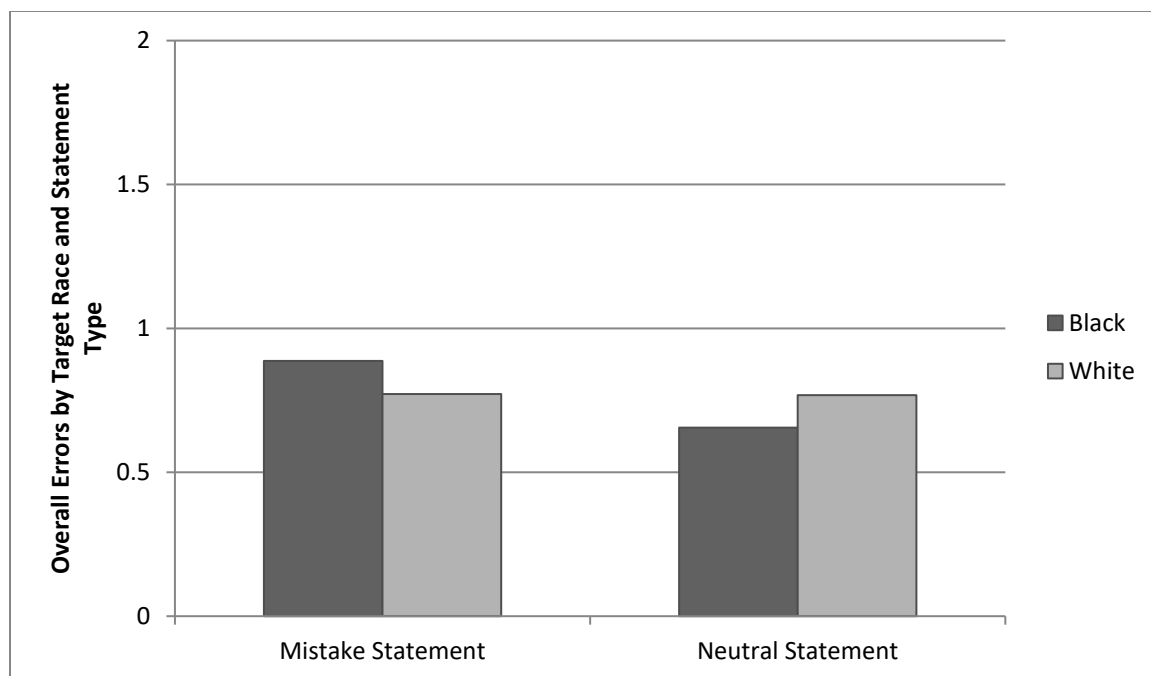


Figure 6. Overall errors by target race and statement type.

Evaluation of Target

Following the memory test, participants were asked to provide their subjective evaluation of each target on several different measures. Specifically, they were asked to rate on a scale of 1 to 7 how competent the target was, the extent to which the target belonged in STEM, and the extent to which the target has a future in STEM. Analyses of the evaluation measures were tested with a race (Black vs. White) x gender (Male vs. Female) repeated measures ANOVA.

Competence. Results indicated a significant main effect of race $F(1, 96) = 33.06$, $p < .001$, $\eta^2_p = .26$, and a significant main effect of gender $F(1, 96) = 5.59$, $p = .02$, $\eta^2_p = .06$. Both of these effects were qualified by a significant race x gender interaction, $F(1, 96) = 7.07$, $p = .009$, $\eta^2_p = .07$ (See Figure 7).

Among women, the simple effect of race was significant such that people were more likely to view Black women as competent ($M = 3.84$, $SD = .66$) than White women

($M = 3.65$, $SD = .64$), $F(1, 96) = 9.61$, $p = .003$, $\eta^2_p = .09$. The simple effect of race was also significant among men such that people were more likely to view Black men as competent ($M = 3.85$, $SD = .63$), than White men ($M = 3.46$, $SD = .76$), $F(1, 96) = 35.06$, $p < .001$, $\eta^2_p = .27$.

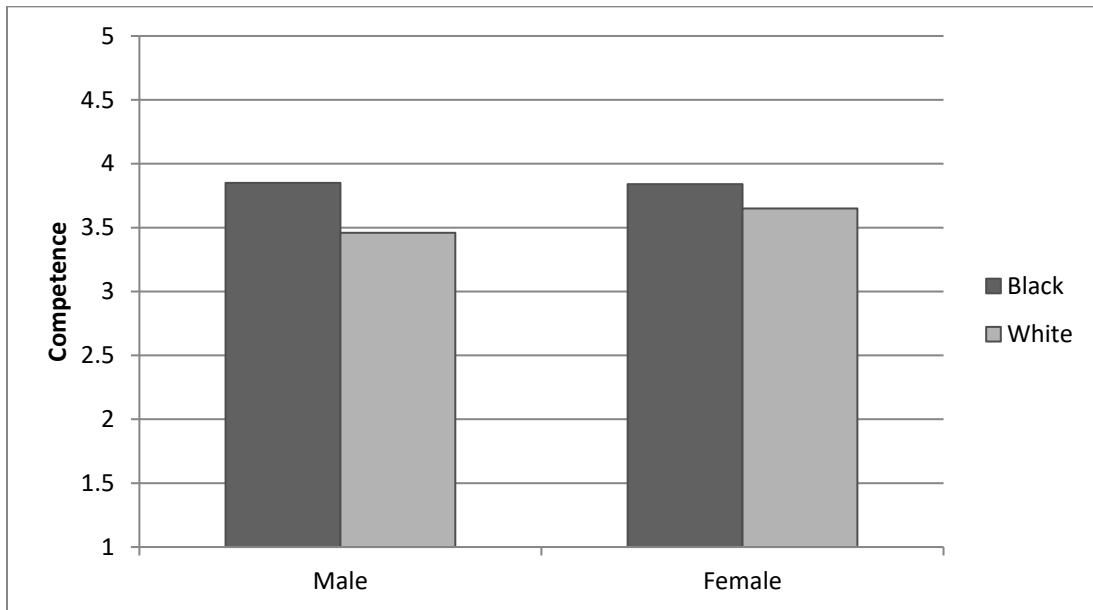


Figure 7. Level of perceived competence by target race and target gender.

Belonging. Results suggested that for belonging in STEM, there were no significant differences between targets ($F_s \leq 1$).

Future in STEM. Results for the future in STEM index indicated that there was a significant main effect of race $F(1, 96) = 6.79$, $p = .01$, $\eta^2_p = .07$, such that Black targets were perceived to have more of a future in STEM ($M = 3.43$, $SD = .78$) as compared to White targets ($M = 3.30$, $SD = .81$). There was neither a significant main effect of gender ($F < 1$), nor a significant race x gender interaction, $F(1, 96) = 2.30$, $p = .133$, $\eta^2_p = .02$, *ns*.

Discussion

The current research used a STEM context to understand people's memories for different targets. Specifically, it was hypothesized that people will be more likely to remember Black female targets when they make a mistake as compared to other targets. It was predicted that the contribution of a mistake statement would reduce invisibility and instead lead to hypervisibility for Black female targets. Further, this hypervisibility would lead to increased scrutiny and ultimately translate to a negative perception of the target in an evaluation context. However, the results provided no support of the hypotheses and participants were no more accurate at distinguishing statements associated with Black female targets as compared to other targets. This suggests that there was no support for hypervisibility.

For the accuracy measures, results indicated significant interactions for race and statement type for between-race/between-gender errors and between-race/within-gender errors. Patterns for between-race/between-gender and between-race/within-gender errors suggested that people misattribute statements associated with Blacks to Whites; however, this misattribution of statements from Blacks to Whites only occurs if the Black target makes a mistake. Therefore, the findings imply that participants either genuinely believed that a White target was associated with the statement or they were unwilling to attribute mistake statements to Black targets and risk making a false accusation. If the latter is true, the unwillingness to attribute mistakes to Black targets implies that participants had a strong concern for appearing prejudiced. This concern for appearing prejudiced is further evidenced in their decision to evaluate Black targets more positively than White

targets, suggesting that participants may be motivated to respond in a socially desirable way.

It is possible that social desirability and egalitarianism concerns drove participants to respond in a specific way toward Black targets. People are often pressured to act consistent with an egalitarian mindset (Plant & Devine, 1998). If participants were aware of the experiment's interest in perceptions of race and abilities, this may have lead participants to over-compensate their explicit ratings of the Black targets. This is supported in prior research, which has shown participants' over-compensation of favorable ratings for outgroup members at the explicit level, regardless of the participant's individual prejudice level (high vs. low prejudice; Dambrun & Guimond, 2004). In this case, the researchers examined French participants' ratings of Arabs (outgroup members) positive and negative traits. The results indicated that participants showed a more positive evaluation of Arabs as compared to their ingroup, the French. In reference to the current research, Dambrun and Guimond's (2004) findings support my finding that participants in the current research over-compensated by explicitly showing favorable ratings of outgroup members (Black targets). However, it is possible that if participants' evaluations were examined implicitly, these favorable ratings may disappear.

Previous research by Jussim, Coleman, and Lerch (1987) demonstrated an instance in which Blacks were evaluated more positively than Whites. Specifically, the researchers examined the effect of race (Black vs. White), dialects (Standard English vs. Nonstandard English), and class (Lower vs. Upper) for male job applicants on competence, job suitability, intelligence, etc. Results from this research indicated that on

average, Black applicants received more favorable evaluations than White applicants. However, class and dialect had a larger impact on the evaluation than did race.

To explain these findings the researchers cite several potential theories. One in particular is expectancy-violation theory, which suggests that evaluations become more extreme (either positively or negatively), when stereotype-based expectations are violated by an individual's characteristics (Kelley, 1971). An example they use to describe how Blacks may be perceived more favorably than similarly qualified Whites is that "the presence of obstacles augments the perceived role of positive personal qualities in the success of Blacks" (p. 537). Through the process of augmentation and discounting, participants saw fewer unfavorable qualities in Black applicants and more unfavorable qualities in White applicants. This principle could be applied to the current research to explain the decrease of mistake statements associated with Black targets, and the increased allocation of mistakes to White targets because a White target who made a mistake may have violated an expectation of competence, which may have lead participants to evaluate them more negatively. Analogously, Kelley's (1971) expectancy-violation theory could also be applied to explain the negative evaluations participants attributed to White targets on competence.

Implications

The results of this research have interesting implications for understanding the underrepresentation of ethnic minority students in STEM fields. Because there were no within-race/within-gender or within-race/between-gender differences, this suggests that Black women's experiences may not be completely different from Black men's experiences in STEM. This could inform researchers that perhaps race, as a category, is

more important to examine and consider when studying ethnic minorities in STEM, as compared to race and gender interactions.

Another implication of the research is the underlying motivation White people have toward maintaining an egalitarian position in society and how this may shape their evaluations and feedback toward outgroups. Specifically, the research suggests that when Black people made mistakes, people were unwilling to attribute the mistake to a member of the outgroup and instead would attribute the mistake to a White person. This misattribution of mistakes informs the researchers that participants are unwilling to risk false accusations of Blacks. Due to the fact that Black targets made an equal number of mistakes as the White targets, this suggests that participants were not willing to negatively evaluate Black targets. An unwillingness to provide honest evaluations to outgroups can have negative implications for the workplace and mentor-mentee relationships. This has been demonstrated in prior research, which has examined the tendency for Whites to hold a positivity bias when giving cross-race feedback (Ruscher, Wallace, Walker, & Bell, 2010). Specifically, White participants tend to be overly positive and essentially unhelpful in situations that require giving critical feedback to Blacks.

Future directions

Prior research by Hamilton and Gifford (1976) has shown that illusory correlations might arise from the tendency to over-associate infrequent co-occurrences. For instance, infrequent events, such as mistakes, are expected to be more distinctive than more common events. Furthermore, the infrequency of behaviors (regardless of valence or desirability) tended to be over-attributed to minority groups, thus creating an illusory

correlation between group membership and behavior). Due to the fact that mistakes occurred equally frequently in all conditions, future research should examine participants' accuracy for faces when only one target makes a mistake rather than all targets who make a mistake. This would be best tested with a between-subjects design methodology. Therefore, because mistakes would occur rarely and for a specific target, it may show a stronger effect for hypervisibility as opposed to when mistakes occur more frequently and equally amongst all targets.

Another avenue future research should examine is how the outcome of the study would differ if no mistakes occurred and whether a domain manipulation is enough to affect the memorability of Black women. The current study used a STEM domain to examine the extent to which Black women were remembered. Because there are stereotypes and expectations for women and ethnic minorities in STEM, it would be interesting to test if memory for target statements changed if domain was manipulated.

Conclusion

Previous research has demonstrated the negative effects of intersectional invisibility on Black women's outcomes. As women's representation continues to grow in STEM fields, it is important to understand how ethnic minority group members, including Black women, are remembered and perceived. Although the current research did not find support for either hypervisibility or invisibility further research is needed to understand the role of visibility in intersectional research. Future goals of researchers should seek to understand how to improve visibility for ethnic minorities in STEM so as to ensure more positive outcomes in STEM fields and broaden participation.

Appendix A:

Target Faces by Race and Gender

Black Female



Black Male





White Female



White Male



Appendix B:

Mistake Statements

Instructions: Please briefly describe a recent lab experience that occurred during your research assistantship.

1. This person didn't zero out the scale before weighing my compound.
2. This person incorrectly calculated the results and was off by several decimal points.
3. This person broke expensive glassware.
4. This person misidentified an unlabeled test tube.
5. This person didn't attend the mandatory safety training course.
6. This person didn't clean up the lab station before an inspection.
7. This person used the wrong automatic pipette and measured out the wrong amount of solution.
8. This person forgot to pre-weigh weigh boats/filter paper before measuring mass.

Appendix C:

Neutral Statements

1. This person ordered new glassware for the lab.
2. This person attended a training course.
3. This person labeled the test tubes so we would know which solution is which.
4. This person read a standard operating procedure (SOP) manual.
5. This person identified unknown solutions.
6. This person drew chemical structures.
7. This person performed metal extractions.
8. This person took technical notes of the procedure.

Appendix D:

Foil Statements

1. This person failed to maintain sterile conditions and accidentally caused contamination.
2. This person didn't wear the proper personal protective equipment when necessary (i.e., They didn't wear gloves when handling acid/base).
3. This person entered data into an excel spreadsheet for analysis.
4. This person made stock solutions.
5. This person filled boxes of pipette tips.
6. This person drew solutions into syringes.
7. This person mislabeled a test tube of solution so we didn't know what it was.
8. This person wrote a lab report.
9. This person gradually titrated a solution.
10. This person calculated concentrations of different compounds.
11. This person measured pH of unknown solutions.
12. This person skipped an important step in the procedure.
13. This person performed neutralization and reduced the acidity of a solution.
14. This person reviewed Material Safety Data Sheets.
15. This person worked in the fume hood.
16. This person calibrated machines for measurement.

Appendix E:

Evaluation Measures

Instructions: Please use the following scale to rate the following research assistants.

There are no right or wrong answers; we are simply interested in your individual opinion.

So please provide your honest evaluation.

1 – Not at all	2	3	4	5	6	7 – Very Much
-------------------	---	---	---	---	---	------------------

1. How likely is the student to pursue an undergraduate degree in the sciences?
2. How likely is the student to pursue a graduate degree in the sciences?
3. How likely is the student to be successful in science?
4. This student belongs in science.
5. This student is an outsider in science.
6. Other people in science are a lot like this student.
7. How competent is this student?
8. How qualified do you think the student is?
9. How skillful is this student?
10. How lazy is the student?
11. How hard working is the student?
12. This student is extremely careful in science labs.

References

- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, *94*(4), 991-1013.
- Chambers, D. W. (1983). Stereotypic images of the scientist: The Draw-a-Scientist Test. *Science education*, *67*(2), 255-265.
- Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 151-155.
- Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist*, *64*(3), 170.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631.
- Dambrun, M., & Guimond, S. (2004). Implicit and explicit measures of prejudice and stereotyping: Do they assess the same underlying knowledge structure?. *European Journal of Social Psychology*, *34*(6), 663-676.
- Davis, A. Y. (1981). *Women, race, and class*. New York: Random House.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5.
- Devos, T., & Banaji, M. R. (2005). American= white? *Journal of personality and social psychology*, *88*(3), 447.
- Ghavami, N., & Peplau, L. A. (2012). An intersectional analysis of gender and ethnic stereotypes testing three hypotheses. *Psychology of Women Quarterly*, *37*(1), 113-127.

- Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). "Ain't I a woman?": Towards an intersectional approach to person perception and group-based harms. *Sex Roles, 59*(5-6), 392-403.
- Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12*(4), 392-407.
- Hamilton, D. L., & Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology, 39*(5), 832.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*(3), 258.
- Howard, J. W., & Rothbart, M. (1980). Social categorization and memory for in-group and out-group behavior. *Journal of Personality and Social Psychology, 38*(2), 301.
- Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 52*(3), 536.
- Kelley, H. H. *Attribution in social interaction*. New York: General Learning Press, 1971.
- Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75*(5), 1155.
- Levin, S., Sinclair, S., Veniegas, R. C., & Taylor, P. L. (2002). Perceived discrimination in the context of multiple group memberships. *Psychological Science, 13*(6), 557-560.

- Ma, Correll, & Wittenbrink (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. *Behavior Research Methods*.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474-16479.
- National Science Foundation, National Center for Science and Engineering Statistics. 2015. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2015. Special Report NSF 15-311. Arlington, VA. Retrieved from: <http://www.nsf.gov/statistics/wmpd/>.
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... & Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593-10597.
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811.
- Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*(5-6), 377-391.
- Risen, J. L., Gilovich, T., & Dunning, D. (2007). One-shot illusory correlations and stereotype formation. *Personality and Social Psychology Bulletin*, *33*(11), 1492-1502.
- Rosette, A. S., & Livingston, R. W. (2012). Failure is not an option for Black women: Effects of organizational performance on leaders with single versus dual-

- subordinate identities. *Journal of Experimental Social Psychology*, 48(5), 1162-1167.
- Ruscher, J. B., Wallace, D. L., Walker, K. M., & Bell, L. H. (2010). Constructive feedback in cross-race interactions. *Group Processes & Intergroup Relations*.
- Schug, J., Alt, N. P., & Klauer, K. C. (2015). Gendered race prototypes: Evidence for the non-prototypicality of Asian men and Black women. *Journal of Experimental Social Psychology*, 56, 121-125.
- Sesko, A. K., & Biernat, M. (2010). Prototypes of race and gender: The invisibility of Black women. *Journal of Experimental Social Psychology*, 46(2), 356-360.
- Sidanius, J., & Pratto, F. (1999). Social dominance: An intergroup theory of social hierarchy and oppression. New York: Cambridge
- Steele, J. (2003). Children's Gender Stereotypes About Math: The Role of Stereotype Stratification. *Journal of Applied Social Psychology*, 33(12), 2587-2606.
- 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(7), 778.
- Thomas, E. L., Dovidio, J. F., & West, T. V. (2014). Lost in the categorical shuffle: Evidence for the social non-prototypicality of Black women. *Cultural Diversity and Ethnic Minority Psychology*, 20(3), 370.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232.
- Williams, J. C., Phillips, K. W., & Hall, E. V. (2014). Double Jeopardy? Gender Bias Against Women of Color in Science. *Tools For Change*.

Zarate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition*, 8(2), 161-185.