Fall 2016

Gender stereotypes and academic performance : the influence of salient role models on stereotype validation

Kelsey Thiem
University of Iowa

GENDER STEREOTYPES AND ACADEMIC PERFORMANCE:

THE INFLUENCE OF SALIENT ROLE MODELS ON STEREOTYPE VALIDATION

by

Kelsey Thiem

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Psychology in the
Graduate College of
The University of Iowa

December 2016

Thesis Supervisor: Assistant Professor Rebecca Neel

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Kelsey Thiem

has been approved by the Examining Committee for
the thesis requirement for the Doctor of Philosophy degree
in Psychology at the December 2016 graduation.

Thesis Committee: _____
Rebecca Neel, Thesis Supervisor

_____
C. Daryl Cameron

_____
Jason K. Clark

_____
Steven Hitlin

_____
Paul D. Windschitl

ACKNOWLEDGEMENTS

I would like to thank those who have supported me along the way to this stage in my career. Thank you to my parents, Pat and Debi, who taught me I could accomplish anything I set my mind to, and to my sister, Hailey, who has always been there to playfully remind me during hard times that "my name is Kelsey Thiem, and there is no one in the world better than me." Thank you to Ross Johnson, for loving and supporting me, for enabling my stress-fueled candy addiction, and for staying up and keeping me company while I worked into the early morning hours. Thank you to Lee Budesheim, who was the first to believe in my ability to complete a Ph.D., and to Jillian O'Rourke Stuart for being a constant companion, sounding board, and source of emotional support during my time at Iowa. I'd like to thank Becca Neel for acting as my secondary advisor, even before she became my official paper advisor, and for being an excellent example of a female role model. Finally, I'd like to thank my advisor, Jason Clark, for taking a chance on a quiet, naive undergraduate to entrust with the beginnings of his exciting new line of research. I could not have accomplished this degree without each and every one of them.

ABSTRACT

People commonly seek out role models when they want to achieve their goals because role models help people believe that success is possible and demonstrate how to achieve it. Because seeking out role models is a common occurrence, a great deal of research has been devoted to understanding the effects that they can have on those who look up to them. One effect that has not been previously examined is the extent to which role models can affect people's certainty in their previous performance perceptions. Evaluative certainty is often increased for people when their performance perceptions are confirmed by the presence of a congruent stereotype: a phenomenon known as stereotype validation. Stereotype validation has been shown to effect women within stereotypically male domains. Importantly, higher evaluative certainty predicts negative downstream outcomes for these women, including lower beliefs in their math ability and reduced expectations for a future math performance. The goal of the current research was to investigate whether the salience of female role models reduces or enhances women's evaluative certainty following stereotype validation. Four studies reveal partial support for the certainty reduction hypothesis. Stereotype-validated women are less certain of their poor performance perceptions when they are exposed to female role models.

PUBLIC ABSTRACT

Because people frequently seek out role models, a great deal of research has examined the effects that they can have on those who look up to them. One effect that has not been previously studied is the extent to which role models can affect people's certainty in their previous performance perceptions.

During a process called stereotype validation, people who perform poorly on a task (e.g., an academic test) become more certain that their performance was poor when they are afterwards reminded of a stereotype that supports this perception. For example, when a woman performs poorly on a math test, she will be more certain that she performed poorly if, after the test, she is reminded of the stereotype that women are bad at math. Because the stereotype fits with her perception that she performed poorly on the math test, she should be more certain that her performance was indeed poor. Importantly, when a woman is certain of her poor performance, she is less likely to believe in her math abilities and less likely to believe that she could perform well on a future math test.

The goal of the current research was to investigate whether women become less or more certain that they performed poorly after they read about female role models who have performed well in a stereotyped domain. The current research found that women become less certain in their poor performance after looking up to a female role model.

TABLE OF CONTENTS

LIST OF TABLES xi

LIST OF FIGURES xii

CHAPTER 1 ROLE MODELS, ACADEMIC PERFORMANCE, AND STEREOTYPE VALIDATION 1

The Beneficial Effects of Role Models 2

Similarity of The Role Model Matters 4

The Importance of Role Model Gender For Women in Stereotyped Contexts 5

The Potential Benefits of Role Models in Regards to Attitude and Evaluation Certainty 7

The Certainty Reduction Hypothesis 9

The Certainty Enhancement Hypothesis 11

The Current Work 13

CHAPTER 2 PRELIMINARY RESEARCH 17

Method 17

Participants and Design. 17

Procedure. 18

Independent Variable. 19

Role Model Salience. 19

Dependent Measures. 20

Actual Quiz Performance. 20

Perceived Quiz Performance. 20

Post-Manipulation Evaluative Certainty. 20

Ability Beliefs. 21

Attitudes. 21
Results 21
Actual Quiz Performance. 21
Perceived Quiz Performance. 22
Post-Manipulation Evaluative Certainty. 22
Direct and Indirect Effects on Downstream Outcomes. 23
Ability Beliefs. 23
Attitudes. 24
Discussion 24
CHAPTER 3 INITIAL EVIDENCE FOR THE EFFECTS OF ROLE MODELS ON EVALUATIVE CERTAINTY 26
Method 26
Participants and Design. 26
Procedure. 27
Independent Variable. 27
Role Model Salience. 27
Dependent Measures. 27
Results 27
Actual Quiz Performance. 27
Perceived Quiz Performance. 28
Evaluative Certainty. 28
Direct and Indirect Effects on Downstream Outcomes. 29
Ability Beliefs. 29
Attitudes. 29

Discussion 30
CHAPTER 4 POTENTIAL MODERATION BY ROLE MODEL GENDER 31
Method 34
Participants and Design 34
Procedure. 35
Independent Variable. 36
Role Models. 36
Dependent Measures. 36
Career Interest. 36
Post-Manipulation Performance Rating. 37
Post-Manipulation Quiz Performance 37
Implicit Association Test. 37
Attention Checks 37
Results 38
Actual Performance on the Pre-Manipulation Quiz 38
Perceived Performance on the Pre-Manipulation Quiz. 39
Evaluative Certainty 39
Direct and Indirect Effects on Downstream Outcomes. 39
Ability Beliefs. 39
Attitudes. 40
Career Interest. 40
Post-Manipulation Performance Rating. 41
Post-Manipulation Quiz Performance 41

Implicit Association Test. 41
Discussion 42
CHAPTER 5 AN EXAMINATION OF POTENTIAL BOUNDARY CONDITIONS .. 45
Method 46
Participants and Design. 46
Procedure. 47
Independent Variables. 48
Stereotype Information. 48
Role Model Salience. 48
Dependent Measures. 49
Actual Performance on the Pre-Manipulation Quiz. 49
Perceived Performance on the Pre-Manipulation Quiz. 49
Evaluative Certainty. 49
Ability Beliefs. 49
Future Performance Expectations. 49
Post-Manipulation Performance Rating. 50
Post-Manipulation Quiz Performance. 50
Attention Checks. 50
Results 51
Actual Performance on the Pre-Manipulation Quiz. 51
Perceived Performance on the Pre-Manipulation Quiz. 52
Evaluative Certainty. 52
Direct and Indirect Effects on Downstream Outcomes. 53

Ability Beliefs. 53

Future Performance Expectations. 55

Post-Manipulation Performance Rating. 57

Post-Manipulation Quiz Performance 58

Discussion 59

CHAPTER 6 POTENTIAL MODERATION BY LEVEL OF ROLE MODEL SUCCESS 62

Method 64

Participants and Design 64

Procedure. 64

Independent Variable. 65

Role Model Success Level. 65

Dependent Measures. 66

Actual Quiz Performance. 66

Perceived Quiz Performance 66

Evaluative Certainty 66

Ability Beliefs. 66

Future Performance Expectations. 66

Attitude Towards Business Economics. 66

Post-Manipulation Performance Rating. 67

Perceived Success Attainability. 67

Attention Checks 67

Results 68

Actual Quiz Performance..... 68

Perceived Quiz Performance..... 69

Perceived Success Attainability..... 69

Evaluative Certainty..... 70

Direct and Indirect Effects on Downstream Outcomes..... 70

Ability Beliefs..... 70

Future Performance Expectations..... 72

Attitude Towards Business Economics..... 73

Post-Manipulation Performance Rating..... 74

Discussion..... 75

CHAPTER 7 GENERAL DISCUSSION..... 77

Limitations..... 79

Future Directions..... 87

Conclusions..... 91

APPENDICES..... 92

Appendix A: Example essays used to manipulate role model presence in the pilot study, Study 1, and Study 3..... 92

Appendix B: Example photographs and background information for female and male role models used in Study 2..... 94

Appendix C: Example photographs and background information for mid-level and high-level female role models used in Study 4..... 95

TABLES..... 96

FIGURES..... 104

REFERENCES..... 113

LIST OF TABLES

Table 1. Results of bootstrapping mediation analyses in the pilot study 96

Table 2. Results of bootstrapping mediation analyses in Study 1 97

Table 3. Results of bootstrapping mediation analyses in Study 2 98

Table 4. Results of bootstrapping moderated mediation analyses for participants in the stereotype information absent condition in Study 3 100

Table 5. Results of the bootstrapping moderated mediation analyses for participants in the stereotype information present condition in Study 3 101

Table 6. Results of bootstrapping serial moderated mediation analyses for male participants in Study 4 102

Table 7. Results of bootstrapping serial moderated mediation analyses for female participants in Study 4 103

LIST OF FIGURES

Figure 1. Adjusted mean evaluative certainty as a function of role model condition in the pilot study. 104

Figure 2. PROCESS model 4 used in the pilot study, Study 1, and Study 2. 105

Figure 3. Mean evaluative certainty as a function of role model condition in Study 1. 106

Figure 4. Mean evaluative certainty as a function of role model condition in Study 2. 107

Figure 5. Adjusted mean evaluative certainty as a function of role model condition, stereotype information, and participant gender in Study 3. 108

Figure 6. PROCESS model 73 used in Study 3. 109

Figure 7. Adjusted mean perceived success attainability as a function of role model condition and participant gender in Study 4. 110

Figure 8. Adjusted mean evaluative certainty as a function of role model condition and participant gender in Study 4. 111

Figure 9. Modified PROCESS model 6 used in Study 4. 112

CHAPTER 1

ROLE MODELS, ACADEMIC PERFORMANCE, AND STEREOTYPE VALIDATION

People often seek out role models to look up to when they want to achieve success. For example, children often look up to professional athletes when they first embark on their own athletic endeavors, perhaps hoping one day to become a professional athlete themselves. Adults beginning a new career at an entry-level position typically want to rise through the company's ranks, and they often look up to their managers and supervisors for ways to do so. Because not everyone shares the same goal, not everyone shares the same role model. However, no matter who a person chooses as a role model, all role models generally share one same attribute: they are talented and successful in their respective fields—typically more so than those who look up to them (Marx & Ko, 2012).

Because role models are so frequently sought out, a great deal of research has been dedicated to understanding their influence. Past research has demonstrated that role models can impact people's attitudes towards the role model, attitudes towards the role model's respective field, and attitudes towards the self. However, no known research to date has examined how role models can influence people's certainty in their attitudes and evaluations. The current research aims to address this gap in the literature by examining the effects that role model salience can have within stereotype validation situations, with a focus on female participants and female role models.

The Beneficial Effects of Role Models

Research on role models is often conceived as social comparison research. Because role models are talented and successful—and customarily more talented and successful than the people who look up to them—role models serve as upward social comparisons. People partake in upward social comparisons fairly frequently (Nosanchuk & Erickson, 1985; Suls & Tesch, 1978; Wheeler, 1966; Wheeler & Koestner, 1984; Wood, 1989), even after people have experienced a failure (e.g., Friend & Gilbert, 1973; Samuel, 1973; Suls & Tesch, 1978). For example, in one study where participants were given the choice to compare their performance with those who performed better, those who performed worse, or to make no comparison at all, the majority of participants chose to make an upward social comparison (Collins, 1996). This research suggests that, when an upward social comparison can be made, it will be made.

Festinger (1954) proposed that people choose to compare themselves to superior others because they are motivated by the unidirectional upward drive. People want to improve themselves, so they look to a role model in order to determine how to assimilate toward (i.e., become similar to) that role model. This attempt to assimilate can be achieved because role models demonstrate to people that success is possible (Major, Testa, & Blysma, 1991). People often use role models as a proxy for themselves (see Wheeler, Martin, & Suls, 1997), believing they can accomplish what a role model can accomplish. Therefore, when people see a role model succeed, they come to believe that they can achieve the same success as well. Additionally, role models not only demonstrate that success can be achieved, but they also demonstrate how to achieve that

success (Meichenbaum, 1971). Role models can function as a sort of road map for others, demonstrating the actions necessary to accomplish success.

By establishing what level of success can be reached and by demonstrating how to reach it, role models can inspire people to improve their abilities and performances (Collins, 1996; Lockwood, 2006; Lockwood & Kunda, 1997). After reading about a role model who had achieved success in the students' desired career, university students explicitly reported feelings of inspiration (Lockwood & Kunda, 1997). Students wrote that the role model was "inspiring," "motivating," and caused them to want to "work harder to achieve my goals so that I can have what [the role model] has" (Lockwood & Kunda, 1997, p. 96). Perhaps it is this inspiration and motivation that prompts several of the benefits people experience from role model salience. For instance, exposure to role models increases feelings of self-efficacy (Bandura, Adams, & Beyer, 1977), beliefs in one's abilities (Marx & Roman, 2002), and people's expectations for future performance outcomes (e.g., better course grades; Stout et al., 2011). Upward social comparisons also prompt people to put in more effort towards a task and persist longer on a task (Nadler & Fisher, 1986). Furthermore, people who are exposed to role models often perform better on a task than people who are not exposed to role models (e.g., Blanton, Buunk, Gibbons, & Kuyper, 1999; Seta, 1982). In one study, for example, undergraduate students were asked in the middle of a course to compare their current course grades with another student who either had better or worse course grades than they did. Students who chose to compare their grades with a superior student achieved higher final grades at the end of the course than students who chose to compare their grades with an inferior student (Blanton et al., 1999). All of this past research suggests that people can assimilate

towards a role model; role models prompt psychological benefits that enable people to improve their abilities and their performances.

Similarity of The Role Model Matters

One important caveat of role model exposure is that not every role model is equally beneficial to every person. Certain qualities of a role model can influence who will benefit most from their salience. One aspect that has been found to be influential in past research is the similarity between the role model and the participant; comparisons with dissimilar others have less impact on people than comparisons with similar others (Asgari, Dasgupta, & Stout, 2012; Major et al., 1991; Marx & Ko, 2012; Wood, 1989). For instance, undergraduate participants were marginally more inspired by role models who shared their school affiliation and interests than by role models who did not share these qualities (Marx & Ko, 2012). Furthermore, participants in this study performed significantly better on a math test after interacting with the similar role model than with the dissimilar role model. This research indicates that having a role model who is similar to oneself is important for a person to be able to benefit from the role model's salience. One way similarity can be established is through gender.

People often use shared social identities, such as gender, as a way of perceiving similarity between each other (Brewer & Weber, 1994; Turner & Tajfel, 1986). Indeed, when allowed to choose a role model, people typically choose a role model who shares their gender (Lockwood, 2006; Wood, 1989) even when gender is not relevant to the comparison dimension (Major & Forcey, 1985). However, having a same-gender role model has shown to be more important for women than men. Men appear to benefit equally from male and female role models whereas women benefit more from the

salience of female role models than male role models (Hoyt & Simon, 2011; Latu, Mast, Lammers, & Bombari, 2013; Lockwood, 2006; Marx & Roman, 2002). For instance, men rated themselves as equally positive regardless of the gender of the salient role model; in contrast, women rated themselves more positively after reading about a female role model than a male role model (Lockwood, 2006). Lockwood (2006) speculates that this result is a consequence of women facing more barriers related to their gender than men. Because men typically have fewer obstacles related to their gender to overcome, they do not necessarily need a male role model to show them that success is possible for them to achieve; any role model will do. However, because women more often face impediments to success that can be attributed to their gender, they may require a female role model because female role models demonstrate that gender barriers can be overcome. This may be particularly true in contexts where women are stereotyped to perform poorly because of their gender, such as science, technology, engineering, and math (STEM) fields.

The Importance of Role Model Gender For Women in Stereotyped Contexts

A great deal of research has examined the impact of role model gender on women's performances within stereotypically male domains and found that, in these contexts, women benefit more from female role models than from male role models (Hoyt & Simon, 2011; Latu et al., 2013; Marx & Roman, 2002). As just one example, when performance on a math test was preceded by exposure to role models who were competent in math, women performed significantly better on the test when the role model was female rather than when the role model was male (Marx & Roman, 2002). In contrast, males performed equally well across role model gender.

Research also suggests that, in order for women to benefit from the salience of female role models in stereotypically male domains, the role models need not have succeeded in the exact same domain in which they are performing. As long as the female role model has succeeded in one stereotypically male field, that role model serves as evidence against negative stereotypes about women's performances. Thus, a woman may be able to benefit from a role model to achieve success in one domain even though that role model has achieved success in another domain.

In one study, for instance, participants read essays either about women who were successful in stereotypically masculine fields other than math (e.g., surgery, architecture, law) or about corporations in those fields (McIntyre et al., 2005; McIntyre, Paulson, & Lord, 2003). Participants were then threatened with the stereotype that women are bad at math before taking a math test. Although female participants performed worse than male participants on the math test when they had previously read about corporations, male and female participants performed equally well when they had previously read about successful women. Furthermore, female participants who read about successful women scored better on a subsequent math test than female participants who read about successful corporations. In this instance, women were able to benefit from the female role models and achieve success on a math test even though the role models had not been successful in math themselves. Because the role models had discounted other specific stereotypes about women in male-dominated fields (e.g., stereotypes about women in science or law), women were able to benefit from this model of success in another domain.

The Potential Benefits of Role Models in Regards to Attitude and Evaluation Certainty

The past research, reported above, demonstrates that role models can help people in many different ways (e.g., increasing task persistence, task performance, feelings of self-efficacy, etc.), particularly female role models for women in stereotypically male domains. However, past research has not examined whether role models may benefit people by influencing the certainty with which they hold their attitudes and evaluations. Attitude certainty is an important area of research because a person's certainty in their attitude determines how durable and influential that attitude is. Attitudes, evaluations, thoughts, and judgments that are held with high certainty are less susceptible to change (e.g., Bassili, 1996; Tormala & Petty, 2002) and endure over longer periods of time (e.g., Bizer, Tormala, Rucker, & Petty, 2006) than those held with low certainty. Furthermore, attitudes held with high certainty are more predictive of future behaviors that are based on the attitude (e.g., Fazio & Zanna, 1978).

There are many situations in which attitudes and evaluations held with high certainty can be detrimental to the people who hold them. One such situation is when people become certain of their poor performances in a process known as stereotype validation. Stereotype validation occurs when a negative stereotype is activated after a person's poor performance on a stereotype-relevant task, and the negative stereotype confirms—or validates—the person's evaluation of his or her stereotype-relevant performance as poor. Thus, the person becomes more certain of his or her poor performance perception (e.g., increased evaluative certainty; Clark, Thiem, Barden, Stuart, & Evans, 2015).

As an example of this phenomenon, consider a study (from Clark et al., 2015) in which men and women took a challenging math test before reporting whether they perceived their performance on the test to be poor or strong. Next, participants either did or did not report their gender. For those who did report their gender, the stereotype that women are bad at math was likely made salient. For those who did not report their gender, this negative stereotype was not made salient; there was nothing to prompt its recall. Participants then reported how certain they were in their original performance perception (i.e., how certain they were that their performance was poor or strong). Among women who perceived their test performance was poor, those who subsequently answered the gender demographic item reported being more certain of this perception than women who did not answer the gender demographic item. The gender demographic manipulation did not influence the evaluative certainty of women who reported their performance as strong; the stereotype (i.e., women are bad at math) did not confirm their original performance perception (i.e., "I did well on the math test"). Similarly, the gender demographic manipulation did not affect the evaluative certainty of men who perceived their performance to be poor, ostensibly because of the mismatch between the stereotype (i.e., men are good at math) and their performance perception (i.e., "I did poorly on this math test"). Although men who perceived their performance was strong tended to report higher certainty in their performance perceptions when they reported their gender information compared to when they did not, this effect was not significant.

Importantly, the women who were more certain of their poor performance experienced pernicious downstream consequences of this evaluation. Women who were highly certain of their poor performance reported lower beliefs in their math ability, less

positive attitudes towards math, less interest in pursuing a math-related career, and lower expectations for performance on a future math test than women who were less certain of their poor performance (Clark et al., 2015). Furthermore, in a study involving poor performance on a science test, women's higher certainty in their poor test performance predicted worse performance on a subsequent science test (Clark & Thiem, under review).

Stereotype validation research has demonstrated how detrimental high evaluative certainty can be when the attitude held with certainty is negative and self-relevant. This research has shown, in particular, how detrimental high evaluative certainty can be for women in stereotypically male domains. The goal of the current research was to examine the effects of role models on evaluative certainty within a stereotype validation context. Although stereotype validation can occur for both men and women (see Clark et al., 2015), the present work focused on women within stereotypically male domains.

The Certainty Reduction Hypothesis

A great deal of past research has demonstrated that women benefit from the salience of female role models. One previously unstudied way in which women may benefit from role models is by reducing their certainty in their poor performance perceptions. Specifically, exposure to female role models may reduce evaluative certainty for women who have had their poor performances validated by a negative performance stereotype.

One way that role models may reduce evaluative certainty is through the inspiration and motivation they incite. Although women become more certain of their poor performance after exposure to a performance-confirming stereotype, the inspiration

activated by role models may undercut this certainty. People often become inspired by role models and become motivated to improve their abilities in an attempt to assimilate towards the role model and achieve similar levels of success (e.g., Lockwood & Kunda, 1997). This inspiration and motivation may color women's performance perceptions when they reflect back on them. Whereas women who are not inspired may be more negative when reflecting on their past performances, women who are inspired and motivated may be more forgiving and optimistic in their reflections. In this way, women exposed to female role models may feel less certain that they performed poorly on a previous test than women who are not exposed to role models.

Perhaps a stronger reason to expect role model salience to reduce women's evaluative certainty is because of female role models' stereotype-disconfirming nature. Past research on the Stereotype Inoculation Model asserts that women benefit from female role models who have succeeded in male-dominated fields because they serve as evidence against the stereotype that "STEM is for men" (Stout et al., 2011). After exposure to stereotype-disconfirming female role models, women more strongly identified with STEM, had more positive attitudes towards STEM, and had higher motivation to pursue a career in STEM (Stout et al., 2011). Although the Stereotype Inoculation Model refers to situations in which stereotype-disconfirming female role models are made salient prior to stereotype-relevant tasks, these role models could serve a similar function after a stereotype-relevant performance, such as stereotype validation.

Within stereotype validation, women's certainty in their poor performance is increased by reminding them of the negative performance stereotype about their gender. However, female role models who have achieved success in a male-dominated field

directly oppose this stereotype. Because of this contradiction, women who are exposed to female role models may not consider the stereotype to be valid. If women do not consider the stereotype to be valid, they will be less likely to use it when assessing their certainty in their performance evaluation. By contradicting and undermining the negative stereotype, role models may prevent the increase in evaluative certainty typically produced by the stereotype's activation.

For these reasons, I propose the certainty reduction hypothesis: after performing poorly and after subsequently being exposed to a negative, performance-relevant gender stereotype (i.e., stereotype validation), women who are exposed to stereotype-disconfirming female role models will report less certainty in their poor performance than women who are not exposed to such role models. Furthermore, women exposed to role models should experience less of the negative downstream outcomes typically associated with stereotype validation compared to women who are not exposed to role models. This is because attitudes and evaluations held with lower certainty should be less predictive of downstream outcomes. For example, women who are exposed to role models should report ability beliefs and performance expectations that are not as low as those reported by women who are not exposed to role models.

The Certainty Enhancement Hypothesis

Although there is support for the certainty reduction hypothesis, there is also some research to support a competing certainty enhancement hypothesis: after a poor performance followed by exposure to a validating negative stereotype, women exposed to stereotype-disconfirming female role models will be more certain of their poor performance than women who are not exposed to such role models. Research has shown

that making upward social comparisons can sometimes be threatening (Major, Sciacchitano, & Crocker, 1993; Parks-Stamm, Heilman, & Hearns, 2008) and emotionally painful (Brickman & Bulman, 1977). When people compare themselves to superior others, they are made aware of their personal shortcomings (Dijksterhuis & Van Knippenberg, 1998; Rudman & Phelan, 2010). This awareness may lead people to contrast away from the superior other. Rather than assimilating to the role model by believing that they can improve and achieve the same amount of success, people may focus on the evidence that they are less competent than the role model. This focus on the differences between the self and the role model can lead to detrimental effects on self-perceptions, including low self-esteem, low ability beliefs, and depression (Hoyt & Simon, 2011; Major et al., 1993; Parks-Stamm et al., 2008; Rudman & Phelan, 2010)—a sharp contrast away from the positive evaluations of the role model. Furthermore, these contrast effects are all the more likely when the role model is an ingroup member because ingroup members cannot be dismissed as irrelevant comparisons in the same way that outgroup members can (Major et al., 1993). Thus, women may be especially likely to contrast away from female role models.

Because role model salience would immediately follow a poor performance and exposure to a negative, performance-relevant stereotype, the disparity between the role models' success and the participants' relative failure may be particularly salient. This obvious disparity may result in emotional pain and feelings of threat, causing women to evaluate themselves negatively. This negative evaluation may manifest in greater certainty of their poor performance. In this way, women who are exposed to role models may contrast away from role models rather than assimilating towards them.

The Current Work

Although a great deal of past research has examined the impact of role models, researchers have not yet examined role models' impact on people's certainty of their performance perceptions. Because high certainty of one's poor performance can lead to pernicious consequences, including low ability beliefs and poor performance on a subsequent test, it is important to determine what effect role models can have on a person's evaluative certainty. The studies within this dissertation use stereotype validation methods to examine whether exposure to female role models influences women's evaluative certainty, and in what direction.

The certainty reduction hypothesis predicts that, after having their poor performance validated by a negative stereotype, women who are exposed to female role models will be less certain of their poor performance than women who are not exposed to female role models. This reduced certainty should protect against the negative consequences that typically result from stereotype validation; for instance, women for whom female role models are salient should report ability beliefs and future performance expectations that are not as low as those reported by women for whom female role models are not salient.

Alternatively, the certainty enhancement hypothesis predicts that, after having their poor performance validated by a negative stereotype, women who are exposed to female role models will be more certain of their poor performance than women who are not exposed to female role models. This enhanced certainty should lead to greater severity of the negative outcomes typically associated with stereotype validation; women for whom female role models are made salient should report lower ability beliefs and

future performance expectations than women for whom female role models are not made salient.

Study 1 was designed to investigate these competing hypotheses within a science context. After performing poorly on a science test and reading about the gender-science stereotype (i.e., men are better at science than women), female participants either read about women who succeeded in stereotypically male fields or about successful corporations in those same fields. Based on the results of a pilot study described below, I predicted results in line with the certainty reduction hypothesis: women who read about role models should report less certainty in their poor performance than participants who read about corporations.

Study 2 was conducted to determine whether role model gender differentially influences women's evaluative certainty. All women performed poorly on a science test and then read about the negative gender-science stereotype before reading about either female or male role models who had achieved success within stereotypically male fields. Because women are stereotyped to perform negatively on science tests, I predicted women would need a female role model to demonstrate that this negative stereotype could be overcome. Whereas a female role model who has excelled in a stereotypically male domain directly contradicts the negative gender performance stereotypes, a male role model who has excelled in a stereotypically male domain may help to confirm these stereotypes. Thus, I predicted that evaluative certainty would be lower for women exposed to female participants than for women exposed to male participants.

The goal of Study 3 was to examine several boundary conditions of role models' effects on evaluative certainty. For instance, role model salience may not affect

participants for whom negative stereotypes are not made salient after a poor performance. If their poor performance is not validated by a negative stereotype, people may not experience an initial increase in evaluative certainty that role models work to reduce. Therefore, participants who do not read a negative stereotype after their poor performance should report similar levels of evaluative certainty regardless of whether they are exposed to role models. In order to examine this boundary condition, Study 3 included conditions in which negative stereotypes were not activated after poor quiz performance. I anticipated that female participants for whom negative stereotypes were not made salient would report similar levels of evaluative certainty regardless of role model salience.

Additionally, role model salience may not influence participants for whom negative stereotypes, activated after a poor performance, are not self-relevant. If a negative stereotype is presented after a poor performance, but doesn't validate that poor performance, then people's evaluative certainty should not increase. Without an increase in evaluative certainty, role models should have no evaluative certainty to reduce. Thus, participants for whom negative stereotypes are not self-relevant should report similar levels of evaluative certainty. In order to test this boundary condition, Study 3 collected data from male participants. I anticipated that, because the negative gender-science stereotype should not confirm men's poor science quiz performance, role models should have no certainty to reduce, causing men to report similar levels of evaluative certainty across role model salience conditions.

Finally, Study 4 was conducted to examine whether the level of success achieved by a role model moderates the role model's effect on evaluative certainty. Past research

has demonstrated that people benefit less from incredibly high-achieving role models than role models who have achieved a more moderate amount of success (Hoyt & Simon, 2011; Lockwood & Kunda, 1997). This is because people perceive high role models' success to be less attainable than medium role models' success (Lockwood & Kunda, 1997). In this study, participants performed poorly on a business economics quiz before reading the gender-business economics stereotype (i.e., men are better at business economics than women). Next, participants either read about mid-level female role models or high-level female role models. I predicted that women would perceive mid-level role models' success to be more attainable than high-level role models' success. Furthermore, I predicted that exposure to mid-level role models and higher perceived success attainability would predict lower evaluative certainty.

Across all studies, I predicted that lower evaluative certainty—produced by certain role model conditions—would predict lower levels of the detrimental downstream outcomes that typically result from stereotype validation. High evaluative certainty produced by stereotype validation usually predicts low ability beliefs in the stereotyped domain, low expectations for a future stereotype-relevant performance, and more negative attitudes towards the stereotyped domain. However, I expected that the reduced evaluative certainty produced by female role model salience would predict relatively higher ability beliefs, relatively higher future performance expectations, and relatively more positive attitudes towards the stereotyped domains.

Before planning the above studies, I conducted a small pilot study to determine if exposure to female role models can influence women's evaluative certainty after stereotype validation, and if so, in what direction.

CHAPTER 2

PRELIMINARY RESEARCH

Because no known work has previously examined the effect of role models on certainty of performance perceptions, I conducted a pilot study to determine if this effect could be found. An additional goal of this pilot study was to determine, if role model salience does influence evaluative certainty, in which direction it does so. According to the certainty reduction hypothesis, after having their poor performance validated by a negative performance-relevant stereotype, women for whom female role models were made salient should report feeling less certain that they performed poorly on an earlier science test than women for whom female role models are not made salient. On the contrary, the certainty enhancement hypothesis predicted that, after having their poor performance validated by a negative performance-relevant stereotype, women for whom female role models were made salient should report greater certainty in their poor test performance than women for whom female role models were not made salient.

Method

Participants and Design.

I set a goal to collect data from 100 participants, and data was collected until that number was met or exceeded. One hundred five University of Iowa female undergraduates (M_{age} = 18.91, SD = 1.15) participated in the study in exchange for partial fulfillment of a research requirement for their introductory psychology or research methods courses. Approximately 66% of the participants (69 participants) identified as White. All participants were randomly assigned to one of two conditions: role models-absent or role models-present.

Procedure.

Each participant was seated at an individual computer station in groups no larger than three by a White female experimenter. Participants began the study by completing a short science quiz. The quiz consisted of 10 multiple-choice questions on different science topics (e.g., physics, earth sciences, etc.) and was designed to be challenging in order to produce primarily negative performance perceptions. Participants were required to answer all quiz questions and were allowed to take as much time as necessary to do so. Immediately after the quiz, participants rated their performance on a scaled measure.

At this point, participants were meant to read stereotype information in order to elicit stereotype validation:

> "Research suggests men tend to perform better than women on tests of scientific knowledge. The research you are participating in is aimed at a better understanding of this."

Similar statements have been used in past research as a means to induce stereotype validation (Clark et al., 2015, in press; Clark & Thiem, under review). Participants were then meant to complete measures of evaluative certainty. The purpose of including evaluative certainty items at this point in the study was to be able to compute differences in certainty from before the role model manipulation to after the role model manipulation. However, the order of the stereotype information and the evaluative certainty items were inadvertently switched so that the stereotype information followed the evaluative certainty items. Because of this error, the pre-manipulation evaluative certainty measures were not analyzed and difference scores between the pre-manipulation and post-

manipulation evaluative certainty measures were not created. Thus, the pre-manipulation evaluative certainty measure will not be discussed further.

On the next screen, participants read a statement asserting that the next part of the study was unrelated to the task they just completed. Participants were told they would read several short essays and were instructed to pay attention because they would be asked questions about each essay. Participants were then randomly assigned to either the role models-present or role models-absent condition. Participants assigned to the role models-present condition read four essays about women who had achieved success in stereotypically male fields; participants assigned to the role models-absent condition read four essays about successful corporations in those same fields. Immediately after each essay, participants answered four questions meant to highlight the success of the woman or corporation. Participants were then asked to think back to the science questions they had answered earlier before responding to measures of certainty in their performance evaluation, beliefs in their own science ability, and attitudes towards science. Finally, participants were thanked and debriefed.

Independent Variable.

Role Model Salience. Participants read one of two sets of essays taken directly from past research (McIntyre et al., 2003, 2005; see Appendix A). Participants assigned to the role models-present condition read four essays detailing the success of fictional women in stereotypically male domains (i.e., architecture, law, surgery, and engineering and business). Participants in the role models-absent condition read four similar essays, but these essays described successful corporations rather than successful women. After reading each essay, participants answered four open-ended questions about the content of

each essay. These questions were also taken directly from past research and were intended to highlight the success of the women and corporations. The questions were as follows, depending on condition: 1) Who [What profession] does the selection describe? 2) What is the person's occupation? [What philosophy is highlighted?] 3) How successful has the person [the philosophy] been? 4) What obstacles did the person have to overcome? [What obstacles stand in the way of this philosophy?]

Dependent Measures.

Actual Quiz Performance. The total number of correct responses to the 10 quiz questions was used as a composite score of actual quiz performance.

Perceived Quiz Performance. After the science quiz but prior to the stereotype information and the role model salience manipulation, participants rated their performance on a continuous measure: "Overall, how well do you think you performed on the science quiz?" (1 = *performed extremely poorly* to 11 = *performed extremely well*).

Post-Manipulation Evaluative Certainty. After the role model manipulation, participants responded to four items assessing how certain they were about their science quiz performance. All items were responded to on an 11-point scale (1 = *strongly disagree* to 11 = *strongly agree*) and were as follows: "Please express how much you agree with the following statement:" (1) "I am certain that I performed poorly on the test." (2) "I am positive that I performed poorly on the test." (3) "I am certain that I performed well on the test." (reverse-scored) and (4) "I am positive that I performed well on the test." (reverse-scored). An index of evaluative certainty was formed by averaging responses to the four items ($\alpha = .87$).

Ability Beliefs. Following the evaluative certainty measure, participants rated their level of scientific ability on three items: "Please rate your own scientific skills on the following scale." (1 = *very weak* to 11 = *very strong*); "Please rate your own scientific knowledge on the following scale." (1 = *very low* to 11 = *very high*); "I believe that I am very knowledgeable about science." (1 = *strongly disagree* to 11 = *strongly agree*). Responses on these items were averaged to form an index of ability beliefs ($\alpha = .92$).

Attitudes. Attitudes towards science were assessed with three items following the ability beliefs measures: "To what extent do you like science?" (1 = *not at all* to 11 = *very much*); "I enjoy learning about and engaging in science." (1 = *strongly disagree* to 11 = *strongly agree*); "I find science to be an interesting subject" (1 = *strongly disagree* to 11 = *strongly agree*). Responses were averaged to form an index of attitudes toward science ($\alpha = .97$).

Results

Actual Quiz Performance.

A one-way analysis of variance (ANOVA) was conducted on participants' science quiz scores. There were no differences in performance across role model salience condition, $F(1, 103) = 1.28$, $p = .26$. Participants in the role models-present condition scored approximately the same as participants in the role models-absent condition ($M_{present} = 3.00$ [$SD = 1.55$], $M_{absent} = 3.35$ [$SD = 1.58$]). According to a one-sample t-test, participants' average quiz score ($M = 3.18$, $SD = 1.57$) was significantly below 50% (5 out of 10 questions correct), $t(104) = -11.89$, $p < .001$. This suggests that, on average, participants performed poorly on the science quiz.

Perceived Quiz Performance.

A one-way ANOVA was conducted on participants' perceived science quiz performance. Perceived performance did not differ as a function of role model condition, $F(1, 103) = 1.58$, $p = .21$. Participants in both the role models-present and role models-absent condition perceived similar levels of quiz performance ($M_{present} = 3.76$ [$SD = 1.60$], $M_{absent} = 4.16$ [$SD = 1.69$]). Additionally, on average, participants rated their performance as poor (i.e., significantly below the midpoint (6) of the scale), $M = 3.97$ ($SD = 1.65$), $t(104) = -12.60$, $p < .001$.

Because neither actual nor perceived quiz performance differed across role model salience condition, the remainder of analyses did not control for them.

Post-Manipulation Evaluative Certainty.

A one-way ANOVA was conducted on the index of evaluative certainty. Although the effect of role model salience was not significant ($F < 1$, $p = .47$), the direction of the means were in line with the certainty reduction hypothesis. Participants in the role models-present condition reported non-significantly lower evaluative certainty than participants in the role models-absent condition ($M_{present} = 7.07$ [$SD = 2.14$], $M_{absent} = 7.35$ [$SD = 1.82$])[1]. See Figure 1.

[1] A one-way analysis of covariance (ANCOVA) controlling for actual and perceived quiz performance was also performed on the index of evaluative certainty. A marginal effect of role model condition emerged, $F(1, 101) = 2.98$, $p = .09$, $\eta_p^2 = .029$. Participants in the role models-present condition tended to be less certain that they performed poorly on the science quiz than participants in the role models-absent condition (adjusted $M_{present} = 6.92$ [$SE = .22$], adjusted $M_{absent} = 7.49$ [$SE = .24$]). There was also a significant effect of the perceived performance covariate, $F(1, 101) = 41.96$, $p < .001$.

Direct and Indirect Effects on Downstream Outcomes.

Ability Beliefs. An ANOVA conducted on the index of ability beliefs revealed no direct effect of the role model manipulation, $F < 1$, $p = .50$. Participants who were exposed to role models reported similar levels of beliefs in their science ability as participants who were not exposed to role models ($M_{present} = 5.61$ [$SD = 2.02$], $M_{absent} = 5.87$ [$SD = 1.90$]).

Because greater evaluative certainty has been shown to lead to reduced ability beliefs and more negative attitudes towards the stereotyped domain (Clark et al., 2015), I anticipated a similar effect in this study; whereas higher evaluative certainty—predicted by role model absence—should predict lower ability beliefs and more negative attitudes towards science, lower evaluative certainty—predicted by role model presence—should predict less detrimental outcomes, including ability beliefs that are not as low and attitudes that are not as negative.

This mediation model was tested using the PROCESS macro for SPSS (Hayes, 2014). Although there was no direct effect of the role model manipulation (the independent variable) on ability beliefs (the dependent variable), meditational analyses are still appropriate (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Rucker, Preacher, Tormala, & Petty, 2011; Shrout & Bolger, 2002). PROCESS Model 4 was used to determine whether evaluative certainty mediated the relationship between the role model manipulation and ability beliefs (see Figure 2). The obtained data were treated as the population, and 10,000 bootstrap samples were randomly drawn with replacement to produce 95% bias-corrected confidence intervals (BC CIs). The role model manipulation did not significantly predict evaluative certainty, and evaluative certainty did not

significantly predict ability beliefs. Thus, the indirect effect of role model salience on science ability beliefs was not found (95 BC CI: -.1990 - .4315). See Table 1.

Attitudes. An ANOVA conducted on participants' attitudes towards science found no direct effect of the role model manipulation, $F < 1$, $p = .68$. Participants in the role models-present and role models-absent conditions reported similar attitudes towards science ($M_{present} = 6.90$ [$SD = 2.82$], $M_{absent} = 6.68$ [$SD = 2.59$]). A PROCESS analysis congruent to the one used on ability beliefs was conducted to examine the indirect effect of the role model manipulation on attitudes towards science. The indirect effect was not significant (95 BC CI: -.1054 - .4284). The role model manipulation did not predict evaluative certainty, and evaluative certainty did not predict attitudes towards science. See Table 1.

Discussion

The goal of this pilot study was to determine whether role model salience could influence people's evaluative certainty after stereotype validation, and if so, in what direction. Although there was no effect of role model condition on the evaluative certainty index, the direction of the means was in line with the certainty reduction hypothesis. After performing poorly on the science quiz and reading the stereotype that men are better at science than women, women who read about successful female role models reported non-significantly less certainty in their negative performance perceptions than women who read about successful corporations.

One reason that this effect may not be significant is likely low power. Because this experiment was conducted as a pilot study, I aimed to collect data from only 100 participants, which was probably an insufficient number to detect the effect of the

manipulation. Additionally, the effect of the role model manipulation may have been dampened by the misplaced pre-manipulation certainty measures. For this reason, I expected the true effect to be much larger, and more significant, than what was reported here. Thus, a replication was necessary.

CHAPTER 3

INITIAL EVIDENCE FOR THE EFFECTS OF ROLE MODELS ON EVALUATIVE CERTAINTY

The purpose of Study 1 was to replicate the pilot study with a sufficient sample size to detect the hypothesized effect. Because the pilot study had limited power (53%) to detect a small-to-medium effect ($\eta_p^2 = .039$) such as this effect was likely to be, it was crucial to replicate it with a sample size large enough to ensure adequate power to detect a significant effect of the role model manipulation.

Although not significant, the pattern of results of the pilot study suggested that exposure to stereotype-disconfirming role models reduced, rather than enhanced, evaluative certainty of stereotype-validated women. Therefore, I predicted a similar, but larger and more significant, result for Study 1. Specifically, I hypothesized that women who read about female role models would report less certainty in their poor performance than participants who read about successful corporations. Furthermore, I anticipated that lower evaluative certainty would mitigate the negative consequences associated with stereotype validation.

Method

Participants and Design.

Power analyses were conducted using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the appropriate sample size for this study. Assuming a small-to-medium effect size of $\eta_p^2 = .039$, power is .80 or greater for $p < .05$ in a sample of 199 participants (approximately 100 participants per cell). Data was collected until the target number of 200 participants was reached or surpassed, which resulted in data from a

total of 219 participants (M_{age} = 19.18, SD = 1.78). Approximately 75% of participants (164 participants) identified as White. The study was a two-cell between-participants design: role models-absent vs. role models-present.

Procedure.

The procedure of Study 1 was identical to the procedure of the pilot study except that the pre-manipulation measures of certainty were not collected.

Independent Variable.

Role Model Salience. The same role model manipulation used in the pilot study was used here in Study 1. Participants in the role models-present condition read four essays about women who achieved success in stereotypically male fields; participants in the role models-absent condition read four essays about successful corporations in those same fields. After each essay, participants responded to the same four questions used in the pilot study that were meant to highlight the success of the role model or corporation.

Dependent Measures.

The measures of actual quiz performance, perceived quiz performance, evaluative certainty, beliefs in science ability, and attitudes towards science were identical to the measures used in the pilot study.

Results

Actual Quiz Performance.

A one-way ANOVA was conducted on participants' science quiz scores. As anticipated, because the quiz preceded the role model manipulation, performance on the science quiz did not differ across condition, $F < 1$, $p = .61$ ($M_{present}$ = 3.41 [SD = 1.51],

M_{absent} = 3.31 [*SD* = 1.49]). Additionally, results of a one-sample t-test revealed that participants performed poorly on the quiz, with an average quiz score (*M* = 3.36, *SD* = 1.50) significantly below 50% (5 out of 10 questions correct), $t(218) = -16.21$, $p < .001$.

Perceived Quiz Performance.

A one-way ANOVA was conducted on participants' rating of their perceived performance on the quiz. Perceived quiz performance did not differ as a function of role model condition, $F(1, 217) = 2.05$, $p = .15$, $\eta_p^2 = .009$ ($M_{present}$ = 4.26 [*SD* = 1.73], M_{absent} = 3.93 [*SD* = 1.67]). In addition, results of a one-sample t-test revealed that participants' average perceived performance (*M* = 4.09, *SD* = 1.71) was significantly lower than the midpoint of the scale (6), $t(218) = -16.54$, $p < .001$. This result demonstrates that the quiz was perceived as difficult, on average.

Because neither actual nor perceived quiz performance differed across role model condition, the remaining analyses did not control for their effects.

Evaluative Certainty.

A one-way ANOVA was conducted on the index of evaluative certainty (α = .83). Consistent with hypotheses, a significant effect of role model condition emerged, $F(1, 217) = 14.72$, $p < .001$, $\eta_p^2 = .064$[2]. Participants in the role models-present condition reported significantly lower evaluative certainty than participants in the role models-absent condition ($M_{present}$ = 6.49 [*SD* = 2.14], M_{absent} = 7.51 [*SD* = 1.79]). See Figure 3.

[2] A one-way ANCOVA that controlled for actual and perceived quiz performance revealed similar results, $F(1, 215) = 12.58$, $p < .001$, $\eta_p^2 = .055$. Participants in the role models-present condition (adjusted $M_{present}$ = 6.58 [*SE* = .17]) reported lower evaluative certainty than participants in the role models-absent condition (adjusted M_{absent} = 7.42 [*SE* = .17]).

Direct and Indirect Effects on Downstream Outcomes.

Ability Beliefs. A one-way ANOVA was conducted to test the direct effect of role model salience on the index of beliefs in science ability (α = .92). No direct effect of role model condition emerged on participants' ability beliefs, $F < 1$, $p = .35$. Participants in the role models-present condition ($M_{present} = 5.81$ [$SD = 2.12$]) and participants in the role models-absent condition ($M_{absent} = 5.56$ [$SD = 1.80$]) reported similar beliefs in their science abilities.

As in the pilot study, PROCESS model 4 was used to examine the indirect effect of role model salience on ability beliefs through evaluative certainty. Ten thousand bootstrap samples were randomly drawn with replacement to produce 95% bias-corrected confidence intervals (BC CIs). As displayed in Table 2, the analysis revealed a significant indirect effect of role model salience on ability beliefs. Women who were not exposed to role models reported higher evaluative certainty than women exposed to role models, and higher evaluative certainty, in turn, predicted lower beliefs in science ability. This result is consistent with hypotheses because the presence of the role model attenuated the negative effects typically associated with stereotype validation.

Attitudes. I conducted analyses to examine both the direct and indirect effect of the role model manipulation on the index of participants' attitudes towards science (α = .98). There was no direct effect of role model condition on attitudes, $F < 1$, $p = .58$. Participants reported similar attitudes towards science whether role models were present ($M_{present} = 7.09$ [$SD = 2.74$]) or absent ($M_{absent} = 6.89$ [$SD = 2.71$]). Although the direct effect of the role model manipulation was not significant, the indirect effect was significant (see Table 2). Role model absence predicted higher evaluative certainty, and

higher evaluative certainty predicted more negative attitudes towards science. This result suggests that role model salience helped to reduce the negative attitudes that typically stem from stereotype validation.

Discussion

The primary goal of Study 1 was to demonstrate that exposure to role models can significantly reduce how certain a person is that they performed poorly on a stereotyped task. The data supported this hypothesis. After performing poorly on a science quiz, and after having their performance perceptions validated by the presence of a negative stereotype, women who read about successful corporations reported more certainty in their poor performance than women who read about role models. Furthermore, consistent with past research (e.g., Clark & Thiem, under review), higher evaluative certainty predicted lower science ability beliefs and more negative attitudes towards science. Framed differently, the presence of female role models predicted lower evaluative certainty and helped to attenuate the negative outcomes typically found as a result of stereotype validation.

In order to ensure that the effect of role model salience on women's evaluative certainty was not idiosyncratic, Study 2 utilized a new set of female role models. More notably, Study 2 was designed to investigate whether role models' gender differentially influences women's evaluative certainty.

CHAPTER 4

POTENTIAL MODERATION BY ROLE MODEL GENDER

The primary goal of Study 2 was to determine whether women experience different amounts of evaluative certainty based on the gender of the role models they are exposed to. Upward social comparisons, like those that occur when a person is exposed to a role model, are more impactful when the comparison other is similar to the participant (Major et al., 1991; Schunk & Hanson, 1985; Wood, 1989), and gender is one dimension of similarity that people use when making social comparisons. People often choose a role model of the same gender rather than a role model of the opposite gender (Lockwood, 2006; Wood, 1989), and women have been shown to benefit more from female role models than from male role models, especially in domains in which women are stereotyped to perform poorly. For instance, women more strongly identified with math after interacting with a female math major than a male math major (Stout et al., 2011), and some research has shown that women perform better on a math test when the test is administered by a female math major rather than a male math major (Marx & Roman, 2002; but see Stout et al., 2011). Additionally, women who were exposed to a female role model reported significantly higher self-ratings (e.g., perceiving the self as smart or capable) than women who were exposed to a male role model (Lockwood, 2006).

In addition to the ways that women have benefitted more from female role models than from male role models in past research, women may also benefit more from female role models in the extent to which the role models reduce women's evaluative certainty; women may be less certain of their poor performance after they are exposed to female role models than after they are exposed to male role models. Evaluative certainty is

increased for women when a negative stereotype confirms their perceptions of their performance as poor. If women are exposed to a female role model who directly contradicts gender stereotypes about women's performances (i.e., by excelling in a male-dominated field), women may be less likely to perceive the stereotype as valid. If they do not perceive the stereotype as valid, women may not use the stereotype when evaluating how certain they are in their performance perceptions. Thus, these women may be protected from the increase in evaluative certainty that usually results from stereotype validation.

On the other hand, male role models who have succeeded in a stereotypical male domain may signal to women that the stereotype is true: only men can excel in stereotypically male domains. Because male role models fit with this negative stereotype, women may be more likely to believe the stereotype and use it when forming their evaluative certainty. In this way, male role models may be unlikely to reduce women's evaluative certainty and may actually work to increase it.

For these reasons, I hypothesized that participants exposed to female role models would report less evaluative certainty than participants exposed to male role models. Furthermore, I predicted that lower evaluative certainty would mitigate the downstream outcomes that typically result from stereotype validation. Because of their lower evaluative certainty, women for whom female role models are salient should report science ability beliefs, attitudes towards science, and science career-interests that are not as low as those reported by women for whom male role models are salient. Furthermore, lower evaluative certainty, predicted by the salience of female role models, should help protect against poor performance on a subsequent stereotype-relevant task.

An additional goal of Study 2 was to examine dependent variables not previously assessed as a result of evaluative certainty but that may nevertheless be influenced by evaluative certainty and role model gender. One of these variables is participants' rating of their performance on the first science quiz. If exposure to role models can reduce women's certainty in their performance perceptions, it may also influence women's actual perceptions of their performance. Attitudes are more susceptible to change when they are held with less certainty (e.g., Bassili, 1996; Tormala & Petty, 2002), so by reducing how certain women are that they performed poorly, role model salience may also provide the opportunity for women's performance perceptions to change. Women may report having performed better on the pre-manipulation quiz after reading about a female role model than after reading about a male role model.

I also included a measure to examine whether role model gender influences implicit associations with science. Past research has examined the influence of role models on implicit measures such as the extent to which participants implicitly identify with a stereotyped domain (Stout et al., 2011). In two of three studies involving female participants, Stout et al. (2011) found that exposure to a female role model produced greater identification with math than exposure to a male role model. Perhaps a similar effect could result in this study: exposure to female role models may result in greater implicit identification with science than exposure to male role models. Furthermore, this effect may be mediated by evaluative certainty.

However, there are reasons to expect that role model exposure would not influence implicit measures. In the two studies from Stout et al. (2011) in which the authors found effects of role models on implicit identification with science, the role

models were physically present in the room. In the third study, in which the authors did not find the effect, participants had only read about role models. Stout et al. (2011) speculated that reading about role models is not as "psychologically powerful" as personally interacting with them (p. 263); therefore, simply reading about role models is not impactful enough to influence self-conceptions. Because participants only read about role models in this study, I had reason to anticipate that role model gender would not influence participants' identification with science.

Additionally, past research has determined that implicit and explicit attitudes change through different means. Whereas explicit attitudes can be changed through brief exposure to new information, implicit attitudes change more slowly after repeated exposure to new associations (e.g., the association between the concepts "self" and "science"; Rydell & McConnell, 2006). Because Study 2's procedures involved brief exposure to new information (in the form of role model biographies) and did not involve long-term exposure to repeated associations, I had further reason to expect role model gender would not influence participants' implicit identification with science.

The final goal of Study 2 was to determine whether the effects of Study 1 could be replicated with a different set of female role models. This was important in order to confirm that the role models from Study 1 were not idiosyncratic and did not produce idiosyncratic results.

Method

Participants and Design.

As in Study 1, I set a goal of 100 participants per cell. However, data was collected beyond 200 participants in order to account for participants who would be

dropped from analysis for incorrectly responding to attention checks. Thus, data was collected from 301 female undergraduates from the University of Iowa psychology research subject pool. Four participants were excluded for not completing the study, and 46 participants were excluded for incorrectly answering at least one of the two attention check questions. The final sample consisted of data from 251 participants (M_{age} = 18.68, *SD* = 1.32). Approximately 80% of participants (200 participants) identified as White. Participants were randomly assigned to one of two conditions: female role models or male role models.

Procedure.

The procedure of Study 2 was similar to that of the pilot study and Study 1. All participants were greeted by a White female experimenter and completed the same challenging science quiz used in the previous two studies before rating their quiz performance on the same 11-point scale. After this rating, participants read the same stereotype information used in the previous studies. Next, instead of reading essays about successful women or corporations, participants viewed a series of 16 brief biographies paired with photographs of fictitious women or men, the majority of whom had achieved success in stereotypically male fields. Participants then completed the same evaluative certainty, ability belief, and attitude ratings as in the previous studies before completing additional measures. Participants responded to an item that assessed their interest in pursuing a science-related career and an item that asked them to rate their pre-manipulation quiz performance for a second time. After participants responded to these items, they took a second science quiz and then completed a single-category implicit

association test (IAT) to assess their level of personal identification with science. Finally, participants answered two attention check items before being thanked and debriefed.

Independent Variable.

Role Models. Sixteen short biographies, accompanied by photographs of the biographies' supposed subjects, were presented sequentially in a randomized order. Each biography described the person's background and accomplishments. Half of the participants read descriptions of women (adapted from Hoyt & Simon, 2011) and half of the participants read descriptions of men. Fifteen of the descriptions involved success in a stereotypically male field (e.g., physics, astronautics), and only the names of the subjects and the gendered pronouns used throughout these biographies differed across role model gender. The sixteenth female role model was a successful activist for feminism, and the sixteenth male role model was a successful activist for racial equality. Additionally, photographs were matched across gender on approximate race, age, facial expression, posture, and manner of dress (see Appendix B).

Dependent Measures.

Participants' actual and perceived performance on the pre-manipulation quiz was assessed in the same way as the previous studies. Additionally, the same measures from the previous studies were used to assess evaluative certainty, beliefs in scientific ability, and attitudes towards science.

Career Interest. After the attitude items, participants responded to one item assessing their interest in a science-related career: **"**I am interested in pursuing a career that requires strong scientific ability" (1 = *strongly disagree* to 11= *strongly agree*).

Post-Manipulation Performance Rating. To assess whether exposure to role models influenced participants' perceptions of their quiz performance, participants responded to the following item on an 11-point scale: "Rate your performance on the science quiz you took earlier" (1 = *I performed very badly* to 11 = *I performed very well*).

Post-Manipulation Quiz Performance. Participants completed a second science quiz that consisted of seven questions and was meant to be moderately easy so as to provide an opportunity for performance variability. The number of correctly-answered questions served as the index of performance on the post-manipulation quiz.

Implicit Association Test. In order to assess participants' implicit level of identification with the science domain, participants completed a single-category IAT (Karpinski & Steinman, 2006; adapted from Stout et al., 2011). In one critical block of trials, the categories "science" and "me" were mapped to one response key and the category "liberal arts" was assigned to another response key. In a second critical block of trials, the category "me" was mapped onto the same response key as "liberal arts" instead of "science." Block order was counterbalanced across participants. In each block of trials, participants saw words representing each category (e.g., "chemistry," "literature," "mine") appear sequentially in the center of the screen. Participants' task was to quickly categorize each word using the corresponding response keys. After an incorrect response, a red X appeared on the screen and remained there until participants pressed the correct response key. A blank, black screen appeared for 250 ms between each trial.

Attention Checks. One item was used to assess whether participants accurately recalled the stereotype information they read after rating their performance on the first science quiz. This question asked, "after taking the first science quiz, you were told

which of the following?" Participants were given the following response options: A) "performance on tests of math ability tends to vary as a function of some personality variables," B) "performance on tests of math ability often varies based on the gender of the respondent," C) "men tend to perform better than women on tests of math ability," D) "performance on the math test will be compared to that of male respondents," and E) "performance on the math test will be compared to that of female respondents." Participants who did not select item C were excluded from the final data set.

Additionally, one item was used to assess whether participants accurately recalled what type of biographies they read: "Think back to the biographies you read earlier. What type of people did you read about?" Participants were given the following response options: A) "men only," B) "women only," and C) "both men and women." Participants in the male role model condition who did not choose item A and participants in the female role model condition who did not choose item B were excluded from analyses.

Results

Actual Performance on the Pre-Manipulation Quiz.

A one-way ANOVA was conducted on participants' performance on the pre-manipulation science quiz. Mean performance did not differ across role model gender, $F < 1$, $p = .68$ ($M_{female} = 3.20$ [$SD = 1.58$], $M_{male} = 3.12$ [$SD = 1.36$]). A one-sample t-test was also conducted to determine whether, on average, participants' quiz performance was poor. As anticipated, participants' average pre-manipulation quiz performance ($M = 3.16$, $SD = 1.47$) was significantly lower than 50% (5 out of 10 questions correct), $t(250) = -19.90$, $p < .001$.

Perceived Performance on the Pre-Manipulation Quiz.

A one-way ANOVA was also conducted on participants' pre-manipulation rating of their performance. Participants in the female role models condition reported similar perceived performances as participants in the male role models condition, $F < 1$, $p = .41$ ($M_{female} = 3.99$ [$SD = 1.59$], $M_{male} = 3.82$ [$SD = 1.66$]). Additionally, results of a one-sample t-test revealed that participants' mean perceived performance ($M = 3.90$, $SD = 1.63$) was significantly lower than the scale's midpoint (6), $t(250) = -20.41$, $p < .001$, signaling that participants perceived their quiz performance to be poor, overall.

Because neither actual performance nor perceived performance on the pre-manipulation quiz differed across role model gender, the following analyses did not control for their effects.

Evaluative Certainty.

Consistent with hypotheses, a one-way ANOVA on the index of evaluative certainty index was significant, $F(1, 249) = 4.10$, $p = .044$, $\eta_p^2 = .016$[3]. Participants in the female role model condition reported less certainty in their poor performance than participants in the male role model condition ($M_{female} = 6.99$ [$SD = 1.81$], $M_{male} = 7.46$ [$SD = 1.81$]). See Figure 4.

Direct and Indirect Effects on Downstream Outcomes.

Ability Beliefs. A one-way ANOVA was conducted to test whether role model gender directly affected participants' beliefs in their science ability. No effect of role

[3] Results of a one-way ANCOVA controlling for perceived and actual performance produced similar results, $F(1, 249) = 3.60$, $p = .059$, $\eta_p^2 = .014$. Participants in the female role models condition tended to report lower evaluative certainty than participants in the male role models condition (adjusted $M_{female} = 7.05$ [$SE = .13$], adjusted $M_{male} = 7.40$ [$SE = .13$]).

model gender was found, $F < 1$, $p = .41$ ($M_{female} = 5.59$ [$SD = 1.85$], $M_{male} = 5.40$ [$SD = 1.71$]).

The indirect effect of role model gender on ability beliefs was examined using the same analysis as in the pilot study and Study 1 (i.e., PROCESS model 4). As displayed in Table 3, evaluative certainty significantly mediated the relationship between role model gender and ability beliefs. Participants exposed to male role models reported higher certainty in their poor performance than participants exposed to female role models, and higher certainty predicted lower beliefs in science ability. This result was consistent with predictions.

Attitudes. There was no direct effect of role model gender on participants' attitudes towards science, $F(1, 251) = 1.33$, $p = .25$ ($M_{female} = 6.76$ [$SD = 2.49$], $M_{male} = 7.11$ [$SD = 2.40$]). However, role model gender indirectly influenced attitudes towards science in a manner that was consistent with predictions (see Table 3). Role model gender significantly predicted evaluative certainty such that participants exposed to male role models reported higher evaluative certainty than participants exposed to female role models. Higher evaluative certainty, in turn, predicted more negative attitudes towards science.

Career Interest. Role model gender did not directly affect participants' interest in pursuing a science-related career, $F < 1$, $p = .77$ ($M_{female} = 6.28$ [$SD = 3.34$], $M_{male} = 6.41$ [$SD = 3.32$]). However, the indirect effect of role model gender on participants' career interest was significant (see Table 3). Consistent with hypotheses, exposure to male role models predicted higher evaluative certainty than exposure to female role models, and

higher evaluative certainty was associated with less interest in pursuing a science-related career.

Post-Manipulation Performance Rating. No direct effect of role model gender emerged on participants' post-manipulation rating of their performance on the pre-manipulation quiz, $F(1, 251) = 2.29$, $p = .13$ ($M_{female} = 4.20$ [$SD = 1.93$], $M_{male} = 3.86$ [$SD = 1.66$]). However, this rating was indirectly affected by role model gender (see Table 3). Exposure to male role models predicted higher evaluative certainty, which in turn predicted lower post-manipulation performance ratings, as predicted.

Post-Manipulation Quiz Performance. The results of a one-way ANOVA revealed that role model gender did not directly influence performance on the post-manipulation quiz, $F < 1$, $p = .62$. Participants in the female role models condition ($M_{female} = 3.31$ [$SD = 1.17$]) performed similarly to participants in the male role models condition ($M_{male} = 3.23$ [$SD = 1.19$]) on the post-manipulation science quiz. Additionally, role model gender did not indirectly influence post-manipulation quiz performance (see Table 3). Although role model gender did predict levels of evaluative certainty, evaluative certainty did not significantly predict performance on the post-manipulation quiz. This result was inconsistent with hypotheses.

Implicit Association Test. D scores for the IAT were calculated such that positive numbers indicate a stronger association of the self with science than with liberal arts. Negative numbers indicate a stronger association of the self with liberal arts than with science. The results of a one-way ANOVA determined that role model gender did not directly influence participants' self-associations with science, $F < 1$, $p = .86$ ($M_{female} = -.12$ [$SD = .31$], $M_{male} = -.12$ [$SD = .32$]). Furthermore, IAT scores were not indirectly

influenced by role model gender (see Table 3). Role model gender significantly predicted participants' levels of evaluative certainty, but evaluative certainty did not predict participants' implicit self-associations with science.

Discussion

The purpose of Study 2 was to determine whether role models' gender differentially influences the evaluative certainty of women after stereotype validation. Women who were exposed to male role models reported greater certainty in their poor performance than women who were exposed to female role models. Furthermore, consistent with past research, higher evaluative certainty was associated with the negative downstream outcomes that typically result from stereotype validation, including reduced ability beliefs, more negative attitudes towards science, and lower interest in persuading a science-related career. Higher evaluative certainty was also associated with lower post-manipulation ratings of performance on the first science quiz, which is a downstream outcome of stereotype validation that has not been previously examined. These results are consistent with the certainty reduction hypothesis since exposure to female role models decreased the amount of certainty women felt about their poor performance and helped to alleviate the negative consequences of stereotype validation.

Inconsistent with hypotheses, evaluative certainty did not predict performance on the second science quiz. From a theoretical perspective, there was ample reason to expect this effect. Evaluations of past academic performance significantly predict future academic performance, including within a science domain (Singh, Granville, & Dika, 2002). Furthermore, because attitudes held with high certainty have a greater influence on subsequent behavior (Tormala & Rucker, 2007), evaluations of past academic

performance held with certainty should be especially predictive of future academic performance. Indeed, past research has found this effect. Women who were more certain of their poor performance on a previous science test performed worse on a subsequent science test than women who were less certain of their poor performance (Clark & Thiem, under review).

However, there is a practical perspective to consider. The effects found in Clark & Thiem (under review) were conducted with a sample from Amazon's Mechanical Turk and used a different post-manipulation science test than the test used here. The questions were changed for this test in order to account for the likelihood that college students have more recent experiences with science than Mechanical Turk participants and thus may be more likely to perform well on science tests than Mechanical Turk participants. Ironically, in an attempt to avoid a ceiling effect, the post-manipulation science quiz used in Study 2 may have been too challenging for the college student sample. Participants scored significantly below 50% on the quiz, $t(251) = -3.07$, $p = .002$. Perhaps if participants were given a post-manipulation science test with a different, slightly easier set of questions, performance on that quiz would be significantly predicted by evaluative certainty.

Additionally, neither role model gender nor evaluative certainty influenced participants' implicit identification with science. This result is not surprising since Stout et al. (2011) were unable to find this effect when participants merely read about role models rather than interacted with them. Furthermore, implicit measures are typically influenced by long-term exposure to repeated associations (Rydell & McConnell, 2006), which did not occur in this study. Finally, this implicit measure was assessed after the

battery of explicit measures. The delay between the manipulation and the IAT could have been too long for any effects of an implicit nature to remain and be detected.

Up to this point, each study has examined the effect that role models have on female participants who have had their poor performance validated by a negative stereotype. The purpose of Study 3 was to extend these past findings by examining the effects of role models in situations where negative stereotypes are either not activated after a poor performance or are not self-relevant. To examine these boundary conditions, Study 3 utilized male and female participants and included conditions in which negative stereotypes were not activated after a poor performance.

CHAPTER 5

AN EXAMINATION OF POTENTIAL BOUNDARY CONDITIONS

The main purpose of Study 3 was to examine several boundary conditions of the effect that role models have on evaluative certainty. One boundary condition may be whether negative stereotypes are activated after a poor performance. If a person's poor quiz performance is not followed by the negative validating stereotype, they will be unlikely to experience an initial increase in evaluative certainty. Without an initial increase in certainty, there will be no certainty for role model salience to reduce. To address this boundary condition, Study 3 included conditions in which negative stereotypes were not activated after a poor quiz performance.

An additional boundary condition may be the extent to which negative stereotypes, activated after a poor performance, are relevant to the performer. For instance, men's poor performance on a stereotypically male task (e.g., science), should not be validated by the subsequent activation of the negative gender stereotype (e.g., men are good at science). Because the stereotype does not fit with their performance perceptions, men should not experience an initial increase in certainty. Therefore, there should be no evaluative certainty for role model salience to reduce. Male participants were recruited for participation in Study 3 in order to assess this boundary condition.

Study 3 therefore involved conditions for male and female participants in which negative stereotype information was either present or absent and in which role models were either present or absent. I anticipated a three-way interaction where both participant gender and the presence versus absence of stereotype information would moderate the effect of role model salience on evaluative certainty. Among female participants, I

expected a significant two-way interaction. For women who read the stereotype information after their poor performance, I expected that those exposed to female role models would report lower evaluative certainty than those exposed to successful corporations. However, for women who did not read the stereotype information, I expected women to report similar levels of evaluative certainty regardless of role model salience. I did not anticipate any effects for male participants. I expected men to report similar levels of evaluative certainty whether the stereotype information was absent or present and whether they were exposed to role models or not.

Additionally, as in the previous studies, I anticipated that reduced evaluative certainty would attenuate the downstream consequences of stereotype validation, such as protection against low ability beliefs and low expectations for a future quiz performance.

Method

Participants and Design.

In order to determine the appropriate sample size for this study, an a priori power analysis was conducted using G*Power 3. By using the observed effect size from Study 1 ($\eta_p^2 = .064$), it was determined that approximately 59 participants per cell was necessary to achieve 80% power (approximately 472 participants). However, I collected data from an additional amount of participants in order to account for data that would need to be excluded for incomplete survey responses and incorrect responses to the attention checks. A total of 612 U.S. citizens participated in the study through Amazon's Mechanical Turk. Data exclusions were made based on the following criteria: 1) failing to complete the study (42 participants), 2) incorrectly answering either of the attention check items (87 participants), and 3) holding a post-graduate degree in science, technology, engineering,

or mathematics (19 participants). The final data set included data from 464 participants (218 males, 246 females; M_{age} = 35.46, SD = 10.88). Approximately 75% of participants (346 participants) identified as White. The study represented a 2 (role models: absent, present) x 2 (stereotype information: absent, present) x 2 (participant gender) between-participants design.

Procedure.

Participants were informed that the purpose of the study was to examine existing knowledge of science topics, and they began the study by completing the first of two multiple-choice science quizzes. The pre-manipulation quiz consisted of 10 questions utilized in past research (Clark & Thiem, under review). Six of the ten questions were from the quiz used in the pilot study and Studies 1 and 2; the remaining four questions from the previously used quiz were replaced by slightly easier questions. As discussed earlier, I did not expect Mechanical Turk participants to be as well-versed in science as the college students who comprised the samples of the previous studies, and I did not want participants to perceive that they had performed so poorly that their evaluative certainty could not be influenced by role model salience. However, the quiz was designed to be difficult enough so that most participants would perceive their quiz performance to be poor. All participants then rated their performance on the same scaled measure used in the previous studies. Next, participants were randomly assigned to a stereotype information condition. Participants in the stereotype information-absent condition read a statement asserting that performance on science tests differs as a function of personality variables. Participants in the stereotype information-present condition read the same statement used to validate performances in the previous studies. Participants were then

randomly assigned to the same role model condition from Study 1. Next, all participants responded to measures of evaluative certainty, perceived success attainability, beliefs in science ability, and expectations for a future science quiz performance. Participants rated their quiz performance once again before completing a second science quiz composed of five multiple-choice items. Two of these items were used in the post-manipulation quiz from Study 2. As with the pre-manipulation quiz, several questions from the post-manipulation quiz in Study 2 were replaced by easier questions. Participants also completed two attention check items. Finally, participants were debriefed and paid for their participation.

Independent Variables.

Stereotype Information. After rating their performance perceptions, half of participants read a statement asserting that "research suggests performance on tests of scientific knowledge tends to vary as a function of some personality variables. The research you are participating in is aimed at a better understanding of this." This statement served as the stereotype information-absent condition. The remainder of participants read the same stereotype information statement used in the previous studies: "Research suggests men tend to perform better than women on tests of scientific knowledge. The research you are participating in is aimed at a better understanding of this." This statement served as the stereotype information-present condition.

Role Model Salience. The role model manipulation in Study 3 was identical to the role model manipulation used in the pilot study and Study 1. Half of participants—those in the role models-present condition—read four essays about women who had succeeded in stereotypically male careers. The other half of participants—those in the role models-

absent condition—read about successful corporations. After each essay, participants responded to the same four questions used in the pilot study and Study 1 to highlight the success of the role model or corporation.

Dependent Measures.

Actual Performance on the Pre-Manipulation Quiz. The total number of correctly-answered questions on the first science quiz served as a measure of pre-validation quiz performance.

Perceived Performance on the Pre-Manipulation Quiz. Participants rated their pre-validation quiz performance on an item identical to the one used in the previous studies.

Evaluative Certainty. Two of the certainty items used in the previous studies were also used in Study 3: "I am certain that I performed poorly on the quiz" and "I am certain that I performed well on the quiz" (reverse-coded). Two additional items were also assessed: "I am sure that I performed poorly on the quiz" and "I am sure that I performed well on the quiz" (reverse-coded). All items were answered on a scale from 1 = *strongly disagree* to 11 = *strongly agree*. Responses on these items were averaged to form an index of evaluative certainty ($\alpha = .96$).

Ability Beliefs. Participants answered the same three ability belief items used in previous studies ($\alpha = .96$).

Future Performance Expectations. Two items were used to measure participants' expectations for their performance on a future science quiz. These items were: "Imagine taking a similar science quiz in the future." 1) "I predict my performance would be:" (1 = *very poor* to 11 = *very strong*) and 2) "How do you think you would perform on it?" (1 =

would perform very poorly to 11 = *would perform very well*). An index of future performance expectations was created by averaging participants' responses on these two items ($\alpha = .98$).

Post-Manipulation Performance Rating. In order to assess whether the role model manipulation influenced participants' perceptions of their pre-manipulation science quiz performance, all participants rated their performance on the following item using a scale from 1 = *I performed very badly* to 11 = *I performed very well*: "Once again, please rate your performance on the science quiz you took earlier."

Post-Manipulation Quiz Performance. The number of items answered correctly on the second quiz served as a measure of post-manipulation quiz performance.

Attention Checks. The same item from Study 2 was used in Study 3 to assess whether participants accurately recalled the stereotype information. Participants in the stereotype information-absent condition who did not choose option A and participants in the stereotype information-present condition who did not choose option C were excluded from the sample. Participants were asked one item to assess whether they had paid attention to the role model manipulation essays they read: "Think back to the essays you read earlier. Which of the following professions was NOT represented among the essays?" Participants were given the following response options: A) "surgery", B) "architecture", C) "thermal regulation unit design", and D) "journalism." Participants who did not choose option D were removed from the sample.

Results

Actual Performance on the Pre-Manipulation Quiz.

The results of a one-sample t-test revealed that participants, on average, did not score significantly better than 50% (5 out of 10 questions correct) on the pre-manipulation science quiz, $t(463) = -.42$, $p = .68$ ($M = 4.96$, $SD = 1.99$). This signifies that overall performance on the science quiz was poor.

A three-way ANOVA was also performed on participants' pre-manipulation quiz performance. A main effect of participant gender emerged ($F[1, 456] = 22.61$, $p < .001$, $\eta_p^2 = .047$) such that women performed more poorly than men ($M_{\text{women}} = 4.55$ [SD = 1.88], $M_{\text{men}} = 5.42$ [$SD = 2.02$]). However, this effect was qualified by a Stereotype Information x Participant Gender interaction, $F(1, 456) = 8.55$, $p = .004$, $\eta_p^2 = .018$. Women in the stereotype information-present condition ($M_{\text{stereotype present}} = 4.93$ [SD = 1.95]) performed better than women in the stereotype information absent condition ($M_{\text{stereotype absent}} = 4.19$ [SD = 1.74]), $F(1, 456) = 9.03$, $p = .003$. However, there were no differences in men's performance across stereotype information condition, $F(1, 456) = 1.35$, $p = .25$ ($M_{\text{stereotype present}} = 5.28$ [SD = 2.06], $M_{\text{stereotype absent}} = 5.58$ [SD = 1.97]). No other significant effects emerged (Fs < 1.28, ps > .26).

The main effect of gender is not surprising since men typically outperform women on science tests (Gonzales et al., 2004), but the Stereotype Information x Participant Gender interaction was unexpected. Because the quiz was taken prior to the stereotype information manipulation, I did not anticipate that the manipulation would affect quiz performance for women or for men. This interaction may signify a failure of random assignment.

Perceived Performance on the Pre-Manipulation Quiz.

A one-sample t-test found that participants believed they had performed poorly on the science quiz; their mean perceived performance ($M = 4.62$, $SD = 2.59$) was significantly below the midpoint (6) of the scale, $t(463) = -11.51$, $p < .001$.

A three-way ANOVA was also conducted and revealed a main effect of gender, $F(1, 456) = 24.57$, $p < .001$, $\eta_p^2 = .051$. Women reported lower performance perceptions than men ($M_{women} = 4.07$ [SD = 2.50], $M_{men} = 5.24$ [$SD = 2.54$]). This effect is not surprising since women performed more poorly than men on the science test, and their performance perceptions should reflect their actual performances. No other effects were significant (Fs < 2.56, ps > .11).

Because actual performance on the pre-manipulation quiz differed as a function of the Stereotype Information x Participant Gender interaction and because perceived performance on the pre-manipulation quiz differed across participant gender, the remainder of the analyses controlled for actual and perceived pre-manipulation quiz performance.

Evaluative Certainty.

The results of a three-way ANCOVA on the index of evaluative certainty revealed that the anticipated three-way interaction between role model salience, stereotype information condition, and participant gender was not significant ($F < 1$, $p = .86$; see Figure 5). Women in the stereotype-present condition reported similar levels of evaluative certainty across role model condition (adjusted $M_{\text{role models present}} = 7.31$ [$SE = .20$], adjusted $M_{\text{role models absent}} = 7.20$ [$SE = .19$]), as did women in the stereotype-absent condition (adjusted $M_{\text{role models present}} = 7.12$ [$SE = .19$], adjusted $M_{\text{role models absent}} = 7.43$ [SE

= .20]). Similarly, men in the stereotype-present condition reported similar levels of evaluative certainty regardless of role model salience (adjusted $M_{\text{role models present}}$ = 6.78 [*SE* = .19], adjusted $M_{\text{role models absent}}$ = 6.58 [*SE* = .21]), and so did men in the stereotype-absent condition (adjusted $M_{\text{role models present}}$ = 6.76 [*SE* = .20], adjusted $M_{\text{role models absent}}$ = 6.88 [*SE* = .22]). However, a main effect of participant gender emerged, $F(1, 454) = 12.51$, $p < .001$, $\eta_p^2 = .027$. Women were more certain of their poor performance than men (adjusted M_{women} = 7.27 [*SE* = .10], adjusted M_{men} = 6.75 [*SE* = .11]). The only other significant effect to emerge from this analysis was that of the perceived performance covariate ($F[1, 454] = 838.10$, $p < .001$, $\eta_p^2 = .649$; all other Fs < 1.71, ps > .19)[4].

Direct and Indirect Effects on Downstream Outcomes.

Ability Beliefs. A three-way ANCOVA was conducted to determine if role model salience, stereotype information, participant gender, or any of their interactions directly affected participants' beliefs in their science ability. As with evaluative certainty, the three-way interaction was not significant, $F < 1$, $p = .37$. However, many lower order effects emerged as significant. A main effect of gender was found ($F[1, 454] = 15.25$, $p < .001$, $\eta_p^2 = .033$) such that women had lower ability beliefs than men (adjusted M_{women} = 4.70 [*SE* = .11], adjusted M_{men} = 5.33 [*SE* = .12]). This main effect was qualified by the Role Model Salience x Participant Gender interaction, $F(1, 454) = 4.30$, $p = .04$, $\eta_p^2 = .009$. Among women, role model salience did not influence ability beliefs, $F < 1$, $p = .46$

[4] Results of a three-way ANOVA that did not control for actual and perceived performance on the pre-manipulation quiz revealed similar results. The three-way interaction between role model salience, the stereotype information condition, and participant gender was not significant ($F < 1$, $p = .81$). However, a main effect of gender emerged ($F[1, 456] = 38.59$, $p < .001$, $\eta_p^2 = .078$) such that women reported higher certainty in their poor performance (M_{women} = 7.76 [*SE* = .17]) than men (M_{men} = 6.18 [*SE* = .19]). No other significant effects emerged (Fs < 1.16, ps > .28).

(adjusted $M_{\text{role models present}} = 4.78$ [$SE = .15$], adjusted $M_{\text{role models absent}} = 4.61$ [$SE = .15$]). However, men in the role models-present condition reported lower ability beliefs (adjusted $M_{\text{role models present}} = 5.09$ [$SE = .15$]) than men in the role models-absent condition (adjusted $M_{\text{role models absent}} = 5.56$ [$SE = .17$]), $F(1, 454) = 4.34$, $p = .04$.

Additionally, there was a marginally significant interaction between role model salience and stereotype information condition, $F(1, 454) = 2.68$, $p = .10$, $\eta_p^2 = .006$. Role model salience did not influence ability beliefs when the stereotype information was present, $F < 1$, $p = .50$ (adjusted $M_{\text{role models present}} = 5.17$ [$SE = .15$], adjusted $M_{\text{role models absent}} = 5.07$ [$SE = .16$]), but when the stereotype information was absent, participants in the role models-absent condition tended to report higher ability beliefs than participants in the role models-present condition (adjusted $M_{\text{role models absent}} = 5.11$ [$SE = .16$], adjusted $M_{\text{role models present}} = 4.70$ [$SE = .15$]), $F(1, 454) = 2.91$, $p = .09$. Although the effect of the actual performance covariate was marginally significant ($F[1, 454] = 3.40$, $p = .07$, $\eta_p^2 = .007$) and the effect of the perceived performance covariate was significant ($F[1, 454] = 318.94$, $p < .001$, $\eta_p^2 = .413$), all other effects were nonsignificant (Fs < 1.95, ps > .16).

An analysis was conducted to determine the indirect effect of role model salience on ability beliefs. As discussed above, I anticipated that both the stereotype information condition and participant gender would moderate the effect of role model salience on evaluative certainty. Thus, PROCESS Model 73 was used to test this predicted moderated-mediation (Hayes, 2014). This model allowed both the stereotype information condition and participant gender to moderate each path of the proposed mediation (see Figure 6). As with the previous PROCESS analyses, the data were treated as the

population, and 10,000 bootstrap samples were drawn with replacement to produce 95% bias-corrected confidence intervals.

As seen in Tables 4 and 5, none of the indirect effects were significant. Evaluative certainty did not mediate the relationship between role model salience and ability beliefs for men or women regardless of the presence of the stereotype information. Although higher evaluative certainty did predict lower ability beliefs for women in the stereotype information present and stereotype information absent conditions and for men in the stereotype information absent condition, evaluative certainty was not predicted by role model salience for any of the four groups.

Future Performance Expectations. The Role Model Salience x Stereotype Information x Participant Gender interaction was not significant, $F < 1$, $p = .49$. Several lower-order effects were significant, however, including a main effect of participant gender, $F[1, 454] = 13.50$, $p < .001$, $\eta_p^2 = .029$. Women reported lower expectations for performance on a future science test than men (adjusted $M_{\text{women}} = 5.03$ [$SE = .10$], adjusted $M_{\text{men}} = 5.58$ [$SE = .11$]). A marginally significant main effect of stereotype information condition also emerged, $F(1, 454) = 3.65$, $p = .06$, $\eta_p^2 = .008$. Participants had marginally higher future performance expectations in the stereotype information present (adjusted $M_{\text{stereotype present}} = 5.44$ [$SE = .10$]) than in the stereotype information absent condition (adjusted $M_{\text{stereotype absent}} = 5.17$ [$SE = .10$]). The main effect of gender was qualified by the Role Model Salience x Participant Gender interaction, $F(1, 454) = 5.28$, $p = .02$, $\eta_p^2 = .011$. Women in the role models-present condition reported marginally higher expectations for future performance than women in the role models-absent condition, (adjusted $M_{\text{role models present}} = 5.20$ [$SE = .14$], adjusted $M_{\text{role models absent}} =$

4.86 [*SE* = .14]). $F(1, 454) = 2.79$, $p = .10$. However, men's future performance expectations did not differ as a function of role model salience, $F(1, 454) = 2.20$, $p = .14$ (adjusted $M_{\text{role models present}} = 5.42$ [*SE* = .14], adjusted $M_{\text{role models absent}} = 5.74$ [*SE* = .16]). Additionally, a marginally significant interaction between role model salience and stereotype information emerged, $F(1, 454) = 3.13$, $p = .08$, $\eta_p^2 = .007$. Although it appeared that, when the stereotype information was absent, participants reported lower future performance expectations in the role models-present condition (adjusted $M_{\text{role models present}} = 5.04$ [*SE* = .14]) than the role models-absent condition (adjusted $M_{\text{role models absent}} = 5.29$ [*SE* = .15]), this simple effect was not significant, $F(1, 454) = 1.00$, $p = .32$. Additionally, when the stereotype information was present, participants appeared to report higher future performance expectations in the role models-present condition than the role models-absent condition (adjusted $M_{\text{role models resent}} = 5.58$ [*SE* = .14], adjusted $M_{\text{role models absent}} = 5.31$ [*SE* = .15]), but this simple effect was also not significant, $F(1, 454) = 2.41$, $p = .12$. Finally, the effect of the perceived performance covariate was also significant ($F[1, 454] = 514.94$, $p < .001$, $\eta_p^2 = .531$). All other remaining effects were nonsignificant (Fs < .47, ps > .49).

A moderated-mediation analysis congruent to the analysis performed on the ability beliefs index was conducted on the future performance expectation index, and similar results emerged (see Tables 4 and 5). Even though higher evaluative certainty predicted lower future performance expectations for men and women in each stereotype information condition, evaluative certainty was not predicted by role model salience for women or men regardless of stereotype information condition. Thus, evaluative certainty

did not mediate the relationship between role model salience and future performance expectations for any participant group.

Post-Manipulation Performance Rating. Results of the three-way ANCOVA showed that the three-way interaction between role model salience, stereotype information condition, and participant gender was not significant, $F < 1$, $p = 94$. The three-way ANCOVA did reveal several lower-order effects, such as a main effect of gender, $F(1, 454) = 19.39$, $p < .001$, $\eta_p^2 = .041$. After being subjected to the stereotype information and role model salience manipulations, women perceived their performance on the pre-manipulation quiz to be lower than men did (adjusted $M_{\text{women}} = 4.47$ [$SE = .08$], adjusted $M_{\text{men}} = 5.00$ [$SE = .09$]). A marginally significant interaction between role model salience and participant gender also emerged, $F(1, 454) = 3.02$, $p < .083$, $\eta_p^2 = .007$. It appeared that women rated their performance as better in the role models-present condition (adjusted $M_{\text{role models present}} = 4.55$ [$SE = .11$]) than the role models-absent condition (adjusted $M_{\text{role models absent}} = 4.39$ [$SE = .11$]), but this simple effect was not significant, $F < 1$, $p = .35$. It also appeared that men rated their performance as worse in the role models-present condition (adjusted $M_{\text{role models present}} = 4.88$ [$SE = .11$]) than the role models-absent condition (adjusted $M_{\text{role models absent}} = 5.12$ [$SE = .13$]), but this simple effect was also not significant, $F(1, 454) = 2.39$, $p < .12$. The effect of the perceived performance covariate was also significant, $F(1, 454) = 1150.73$, $p < .001$, $\eta_p^2 = .717$, but no other direct effects emerged (Fs < 1.67, ps > .20).

The same moderated-mediation analysis used on the other dependent variables in Study 3 was conducted on participants' post-manipulation rating on their pre-manipulation quiz performance. Evaluative certainty did not mediate the relationship

between role model salience and this performance rating. Role model salience did not predict evaluative certainty even though higher evaluative certainty did predict lower post-manipulation performance ratings for men and women when the stereotype was absent and present. See Tables 4 and 5.

Post-Manipulation Quiz Performance. The three-way interaction between role model salience, stereotype information condition, and participant gender was not significant, $F < 1$, $p = .36$. Women performed similarly across role model salience conditions when the stereotype was present (adjusted $M_{\text{role models present}} = 2.78$ [$SE = .14$], adjusted $M_{\text{role models absent}} = 2.75$ [$SE = .14$]) and when the stereotype was absent (adjusted $M_{\text{role models present}} = 2.42$ [$SE = .14$], adjusted $M_{\text{role models absent}} = 2.72$ [$SE = .14$]). Men also performed approximately the same on the second science quiz across role model salience regardless of whether the stereotype information was present (adjusted $M_{\text{role models present}} = 2.79$ [$SE = .14$], adjusted $M_{\text{role models absent}} = 2.90$ [$SE = .15$]) or absent (adjusted $M_{\text{role models present}} = 2.72$ [$SE = .15$], adjusted $M_{\text{role models absent}} = 2.78$ [$SE = .16$]).

The only direct effects to emerge from the three-way ANCOVA on post-manipulation quiz performance were the effects of the two covariates: pre-manipulation quiz actual performance ($F[1, 454] = 58.95$, $p < .001$, $\eta_p^2 = .115$) and pre-manipulation perceived performance ($F[1, 454] = 5.06$, $p = .03$, $\eta_p^2 = .011$). All other effects were nonsignificant (Fs < 2.03, ps > .16).

The results of the same moderated-mediation analysis on post-manipulation quiz performance revealed that evaluative certainty did not mediate the relationship between role model salience and performance on the second science quiz. Not only did role model salience not predict evaluative certainty for men and women regardless of stereotype

information presence, but evaluative certainty did not predict subsequent quiz performance for any of the four groups of participants. See Tables 4 and 5.

Discussion

Study 3 was conducted to examine two boundary conditions of the effects that role models have on people's evaluative certainty. One boundary condition included situations in which a negative stereotype was not activated after a person's poor performance. The second boundary condition included situations in which the negative stereotype, activated after a poor performance, was not self-relevant. In each of these circumstances, participants should not experience an increase in certainty for which the role models would work to reduce. Although, as predicted, role model salience did not influence women's evaluative certainty when the negative stereotype was not activated (testing potential boundary condition one), and, as predicted, role model salience did not influence men's evaluative certainty (testing potential boundary condition two), role model salience also did not influence women's evaluative certainty when the negative stereotype was activated, which was contrary to hypotheses. In other words, the anticipated three-way interaction on evaluative certainty did not emerge.

Of course, one practical reason for not finding this effect could be lack of statistical power. The three-way interaction, if a true effect, is likely a small one, and with 464 participants, I achieved only 58% power to detect a small effect ($\eta_p^2 = .01$). However, there is also a theoretical reason to consider: participants' age. This Mechanical Turk sample had a much higher average age than the college-student samples from the pilot study and Studies 1 and 2 (approximately 35 years old on average compared to approximately 19 years old on average). Whereas many 19-year olds have likely not

chosen a career path, much less started work in that chosen career, 35 year-olds are much more likely to have been working at their respective careers for many years. These older adults may be less likely to feel inspired or motivated by role models since they may anticipate their future success levels to be similar to their current success levels, which are likely moderate at best.

Another aspect of participant age to consider is their age relative to the perceived age of the role models. Past research has demonstrated that people benefit less from role models when they are closer in age to the role model than when they are younger than the role model (Lockwood & Kunda, 1997). Being approximately the same age as a role model can make the role model's success appear less attainable because people may believe that they have little time left to accomplish what the role model has accomplished. If the success appears unattainable, people may not be motivated or inspired to go after that success. Additionally, if a person believes a role model has achieved more success than them in the same amount of time, they may feel threatened by the role model. As discussed in the introduction, feelings of threat may lead to contrast effects rather than assimilation effects (e.g., Rudman & Phelan, 2010), leading participants to have negative self-perceptions which may result in feeling more certain rather than less certain that they performed poorly on the quiz. It is possible that participants in this study perceived the role models to be the same age or younger than themselves. However, role model age was not specified in this study, and no data were collected to be able to ascertain what age participants perceived the role model to be. Thus, I can only speculate that relative age of the participants to the role models was a factor that influenced this study's results.

Although participant age may have influenced the extent to which participants felt motivated or inspired by the role models, participant age should not affect the perception that the role models disconfirmed the gender stereotype. Recall that I theorized that role models not only influence evaluative certainty by increasing motivation, inspiration, and perceived success attainability, but also by demonstrating that the negative gender stereotype is false, rendering the stereotype less influential when women are exposed to female role models. Thus, even if women were not motivated or inspired by the role models, the stereotype-disconfirming nature of the role models should still have worked to reduce evaluative certainty. This study's null results show that this wasn't the case. Therefore, other hidden moderators may be at play with a sample of Mechanical Turk participants.

In spite of these potential issues of a Mechanical Turk sample, Study 4 also used a sample of participants from Mechanical Turk to investigate whether the level of success achieved by a role model influences the extent to which role model salience reduces evaluative certainty.

CHAPTER 6

POTENTIAL MODERATION BY LEVEL OF ROLE MODEL SUCCESS

Study 4 was conducted to examine whether the level of success achieved by female role models moderates the influence that they have on women's evaluative certainty. Past research has demonstrated that role models who have achieved a very high level of success—those who can be considered "superstars," "elite," or "outstanding"—have a less positive impact on people than role models who have achieved a more moderate level of success (Hoyt & Simon, 2011; Lockwood & Kunda, 1997). Highly-successful role models may be less beneficial than moderately-successful role models because their success is perceived as less attainable (Hoyt & Simon, 2011; Lockwood & Kunda, 1997). When people perceive a role model's success to be unattainable, they are precluded from identifying with and relating to the role model (Hoyt & Simon, 2011). Importantly, identification with the role model is a key component of what allows people to benefit from role model exposure (Lockwood, 2006). This explains why positive outcomes of role models are most pronounced for participants who believe it is possible to achieve a similar levels of success as the role models (e.g., Dasgupta & Asgari, 2004; Lockwood & Kunda, 1997).

Based on this past research, I predicted that exposure to mid-level role models would result in lower evaluative certainty for stereotype-validated women than exposure to high-level female role models, and perceived success attainability should mediate this effect. Research discussed above has previously demonstrated that the success of mid-level role models is perceived to be more attainable than the success of high-level role models. If a woman believes that it is possible to achieve a certain level of success, she

may feel less certain that she previously performed poorly in that domain. For this reason, higher perceived success attainability should predict lower evaluative certainty for women. Moreover, I expected that both high perceived success attainability and low evaluative certainty would help mitigate the negative downstream outcomes for women.

In regards to male participants, I anticipated a much different pattern of results. I predicted that men would perceive differences in success attainability as a function of role model success level (i.e., men will report higher success attainability for mid-level role models than for high-level role models), but I did not expect perceived success attainability to significantly predict levels of evaluative certainty for men. Although men tend to benefit from female role models to the same extent as male role models (e.g., Lockwood, 2006), the men in Study 4 should not have their poor performance perceptions validated by the presence of the negative stereotype, so they should not experience an initial increase in evaluative certainty. With men's evaluative certainty already at a low point, there should be little room for role models to reduce this evaluative certainty even further.

The second purpose of Study 4 was to examine whether the effects of role models on evaluative certainty extend outside the domain of science. The past three studies have examined effects within the realm of science, but it is important to determine whether the influence of role models remains localized to evaluative certainty within a science domain or if the effects can generalize to other stereotypes. As with science, women are often stereotyped to be less competent than men in business and business-related skills like leadership and negotiation (e.g., Kray, Thompson, & Galinsky, 2001; von Hippel, Sekaquaptewa, & McFarlane, 2015), and past research has demonstrated that stereotype

validation can occur for women in business domains (Clark et al., in press). Thus, Study 4 examined the effects of role models within a business economics context.

Method

Participants and Design.

As in Study 3, G*Power 3 was used to conduct an a priori power analysis to determine the target sample size for Study 4. Using the observed effect size from Study 1 ($\eta_p^2 = .064$), it was determined that approximately 59 participants per cell was necessary to achieve 80% power. However, the target sample size was increased in order to account for data that would need to be excluded for incomplete survey responses and incorrect responses one the attention check items. A total of 400 U.S. citizens were recruited through Amazon's Mechanical Turk. Participants' data was removed from the sample if they did not complete the study (10 participants) or if they answered the first attention check incorrectly (55 participants). Data from 335 participants comprised the final data set (149 males, 186 females; $M_{age} = 34.37$, $SD = 10.79$). Approximately 76% of participants (256 participants) identified as White. The study represents a 2 (role model: mid-level, high-level) x 2 (participant gender) between-participants design.

Procedure.

Participants read that the study was being conducted in order to examine existing knowledge of business economics. Participants' first task was to complete a multiple-choice quiz, consisting of eight questions about business economics used in past research (Clark et al., 2017). The quiz was designed to be difficult so that almost all of the participants would believe they performed poorly. All participants then rated their

performance on a scaled measure before reading a statement about gender stereotypes about business. This statement was similar to the stereotype information statements used in past studies and read: "research suggests that men tend to perform better than women on tests of business economics. The research you are participating in is aimed at a better understanding of this." Immediately after reading this statement, participants were randomly assigned to read about role models who had achieved either a moderate level of success or a high level of success. Participants then responded to measures of evaluative certainty, beliefs in business economics ability, expectations for a future business economics quiz performance, and attitudes towards business economics. Participants also rated their performance for a second time, responded to an item assessing the extent to which participants believed that the role models' success was achievable, and answered two attention check questions. Finally, participants were debriefed and paid for their participation.

Independent Variable.

Role Model Success Level. Half of the participants read four of the female role model biographies used in Study 2 (adapted from Hoyt & Simon, 2011). These role models functioned as the mid-level role models because they were portrayed as average women who achieved a relatively moderate level of success. The other half of participants read biographies of four famous women who achieved a high level of success (adapted from Dasgupta & Asgari, 2004 and Hoyt & Simon, 2011). The role models were matched across condition on occupation, with three of the role models having achieved success in a stereotypically male field and one role model having achieved success as an

activist for feminism. The role models were also matched across condition on apparent race, facial expression, posture, and manner of dress (see Appendix C).

Dependent Measures.

Actual Quiz Performance. The total number of questions answered correctly on the business economics quiz served as the measure of actual quiz performance.

Perceived Quiz Performance. Participants rated their quiz performance on an item similar to the one used in past studies except the word "science" was replaced with "business economics."

Evaluative Certainty. The evaluative certainty items were identical to those used in Study 3 ($\alpha = .95$).

Ability Beliefs. Participants answered three ability belief items that were similar to those used in past studies ($\alpha = .97$). These items included: "Please rate your own skills in business economics on the following scale:" (1 = *very weak* to 11 = *very strong*), "Please rate your own knowledge of business economics on the following scale:" (1 = *very low* to 11 = *very high*), and "I believe that I am very knowledgeable about business economics." (1 = *strongly disagree* to 11 = *strongly agree*).

Future Performance Expectations. The same two items used to measure participants' expectations for their performance in past studies were also used here, except the word "science" was replaced with "business economics" ($\alpha = .98$).

Attitude Towards Business Economics. Participants responded to two items similar to those used in the previous studies, on an 11-point scale ($\alpha = .94$). One item read "to what extent do you like business economics?" (1 = *not at all* to 11 = *very much*), and

the other item read "I find business economics to be an interesting subject" (1 = *strongly disagree* to 11 = *strongly agree*).

Post-Manipulation Performance Rating. The same item used to assess participants' quiz performances after the role model manipulation in Study 3 was used here as well, except the word "science" was replaced with "business economics."

Perceived Success Attainability. Participants responded to one item used to assess perceived attainability of the role models' success: "Think about the women you viewed earlier. To what extent do you think that someday in the future you might reach a similar level of success in your own field?" (1 = *not at all likely* to 11 = *very likely*).

Attention Checks. The item from Studies 2 and 3 that was used to assess whether participants accurately recalled the stereotype information they read was also used in here in Study 4. Participants who did not choose option C were excluded from the sample. Additionally, participants were asked one item to assess whether participants accurately recognized the apparent age of the role models which reflected role model condition: "Think back to the descriptions of several women you read earlier. Which type of women did you read about?" Participants were asked to choose between "young women" and "middle-aged women." Upon analysis, I determined that this item was poorly written. Whereas approximately 1% of participants in the high-level role model condition responded incorrectly to this item, approximately 25% of participants in the mid-level role model condition responded incorrectly. This difference in the proportion of incorrect responses across condition implies that the age of role models was likely unclear to participants in the mid-level role model condition. This does not necessarily pose a problem for the interpretation of the study's results since age was merely a proxy for the

level of success a role model had achieved. Thus, this attention check was not used to exclude any participants from the final data set.

Results

Actual Quiz Performance.

A one-sample t-test revealed that participants performed poorly on the quiz; on average they scored significantly below 50% (4 out of 8 questions correct), $t(334) = -5.67$, $p < .001$ ($M = 3.56$, $SD = 1.42$).

A two-way ANOVA was conducted on participant's business economics quiz performance. A marginal interaction emerged between role model level and participant gender, $F(1, 331) = 3.15$, $p = .08$, $\eta_p^2 = .009$. Although it appeared as if women performed worse in the mid-level role model condition (adjusted $M_{mid} = 3.38$ [$SE = .15$]) than in the high-level role model condition (adjusted $M_{high} = 3.70$ [$SE = .15$]), this simple effect was not significant, $F(1, 331) = 2.27$, $p = .13$. It also appeared as if men performed better in the mid-level role model condition (adjusted $M_{mid} = 3.71$ [$SE = .16$]) than in the high-level role model condition (adjusted $M_{high} = 3.47$ [$SE = .17$]), but this simple effect was also not significant, $F(1, 331) = 1.07$, $p = .30$. There were no significant main effects (Fs $< .09$, ps $> .77$).

The Role Model Success Level x Participant Gender interaction was unexpected since the quiz was administered prior to the role model manipulation. Furthermore, this effect mirrors the unexpected Stereotype Information x Participant Gender interaction found in Study 3. Although it is highly unlikely that failure of random assignment would occur for two studies in a row, it appears to be the case here.

Perceived Quiz Performance.

The results of a one-sample t-test revealed that participants' average perceived quiz performance ($M = 4.80$, $SD = 2.40$) was significantly below the midpoint (6) of the scale, $t(334) = -9.17$, $p < .001$, which signifies that participants believed their performance to be poor, on average.

Additionally, a two-way ANOVA revealed a significant main effect of gender, $F(1, 331) = 14.40$, $p < .001$, $\eta_p^2 = .042$. Women reported lower perceived performance than men ($M_{women} = 4.36$ [$SD = 2.24$], $M_{men} = 5.34$ [$SD = 2.49$]). No other effects were significant (Fs < 1.13, ps > .29).

Because actual quiz performance differed as a function of the Role Model Success Level x Participant Gender interaction, and because perceived quiz performance differed across participant gender, these two results were controlled for in the remainder of the analyses.

Perceived Success Attainability.

I conducted a two-way ANCOVA on participants' ratings of attainability of the role models' success. Although the anticipated Role Model Level x Participant Gender interaction was not significant, $F < 1$, $p = .78$, a significant main effect of role model success level emerged, $F(1, 329) = 7.43$, $p = .007$, $\eta_p^2 = .022$. See Figure 7. Participants in the mid-level role model condition (adjusted $M_{mid} = 5.64$ [$SE = .23$]) rated the role models' success as more attainable than participants in the high-level role model condition (adjusted $M_{high} = 4.73$ [$SE = .24$])[5]. Additionally, the effect of the perceived

[5] A two-way ANOVA that did not control for actual and perceived performance on the pre-manipulation quiz revealed similar results. A main effect of role model condition emerged, $F(1,329) = 8.27$, $p = .004$, $\eta_p^2 = .024$, such that participants in the mid-level role

performance covariate was significant (F[1, 329] = 19.24, p = .001, η_p^2 = .055) and the effect of the actual performance covariate was marginally significant (F[1,329] = 3.19, p = .075, η_p^2 = .01). No other effects emerged, Fs < 2.50, ps > .12.

Evaluative Certainty.

A two-way ANCOVA concluded that neither role model success level, participant gender, nor their interaction affected participants' certainty of their poor quiz performance, Fs < 1, ps > .67[6]. Women's evaluative certainty was similar across role model condition (adjusted M_{mid} = 6.89 [SE = .14], adjusted M_{high} = 6.86 [SE = .17]), as was men's (adjusted M_{mid} = 7.01 [SE = .16], adjusted M_{high} = 6.86 [SE = .17]). See Figure 8. The actual performance covariate was not significant (F < 1, p = .24), but the perceived performance covariate was, F(1, 329) = 767.56, p < .001, η_p^2 = .70.

Direct and Indirect Effects on Downstream Outcomes.

Ability Beliefs. A two-way ANCOVA was conducted to examine the direct effect of role model level and participant gender on participants' beliefs in their business economics ability. Although the Role Model Level x Participant Gender interaction did not emerge (F < 1, p = .45), the analysis revealed a main effect of participant gender on

---

model condition reported higher perceived success attainability than participants in the high-level role model condition (M_{mid} = 5.65 [SD = 3.10], M_{high} = 4.66 [SD = 3.19]). However, a main effect of gender also emerged F(1,329) = 6.05, p = .014, η_p^2 = .018. Men reported higher perceived success attainability than women (M_{men} = 5.65 [SD = 3.20], M_{women} = 4.78 [SD = 3.11]). The interaction between level of role model success and participant gender was not significant, F < 1, p = .72.

[6] The results of a two-way ANOVA that did not control for actual and perceived performance revealed a main effect of participant gender, F(1, 329) = 9.00, p = .003, η_p^2 = .026, such that women were more certain of their poor performance than men (M_{women} = 7.29 [SD = 2.59], M_{men} = 6.42 [SD = 2.64]). Neither the effect of role model level nor the interaction between participant gender and role model level were significant, Fs < 1, ps > .47.

ability beliefs, $F(1, 329) = 6.45$, $p = .01$, $\eta_p^2 = .019$. Women reported lower beliefs in their business economics ability than men (adjusted $M_{women} = 3.97$ [$SE = .12$], adjusted $M_{men} = 4.42$ [$SE = .13$]). The only other effect to emerge was that of the perceived performance covariate, $F(1, 329) = 358.31$, $p < .001$, $\eta_p^2 = .521$ (all remaining Fs < 1, ps > .45).

I anticipated that role model success level would indirectly influence ability beliefs and the other downstream outcomes through both perceived success attainability and evaluative certainty, but only for female participants. Women should perceive mid-level role models' success to be more attainable than high-level role models' success, and higher perceived success attainability should lead to lower certainty in their poor performance perceptions. Additionally, I expected both high perceived success attainability and low evaluative certainty to protect against the negative downstream outcomes typically associated with stereotype validation.

Although male participants may perceive different levels of success attainability based on the role models' level of achieved success, female role models should not be particularly relevant or beneficial to them, and their evaluative certainty should already be low. Thus, I did not expect perceived success attainability to predict evaluative certainty or ability beliefs for male participants. Without these meditational paths, no indirect effects of role model success level should emerge on the downstream outcomes.

To examine these indirect effects, I utilized a modified version of PROCESS model 6. This model examined whether perceived success attainability and evaluative certainty mediated the path between role model success level and ability beliefs. Participant gender was allowed to moderate the effects of role model success level on

perceived success attainability, evaluative certainty, and the downstream outcomes (see Figure 9). The data were treated as the population from which 10,000 bootstrap samples were drawn with replacement.

The results of this serial mediation analyses revealed that two meditational paths were significant for both male and female participants (see Tables 6 and 7, respectively). Role model success level did not significantly affect evaluative certainty for either men or women, but role model success level did significantly predict perceived success attainability: for both men and women, role models' success was perceived as less attainable for high-level role models than for mid-level role models. In turn, lower perceived success attainability predicted higher evaluative certainty. Both lower perceived success attainability and higher evaluative certainty were associated with low ability beliefs. Framed differently, mid-level role models predicted higher perceived success attainability, which predicted lower evaluative certainty. Both higher perceived success attainability and lower evaluative certainty were associated with an attenuation of low ability beliefs, which is consistent with predictions.

Future Performance Expectations. Results of a two-way ANCOVA determined that neither role model level, participant gender, nor their interaction had an effect on participants' expectations for performance on a future business economics quiz, Fs < 1.32, ps > .25. Women reported similar expectations for a future business economics performance regardless of role model success level (adjusted M_{mid} = 4.99 [SE = .17], adjusted M_{high} = 4.89 [SE = .17]), and the same is true for men (adjusted M_{mid} = 5.03 [SE = .19], adjusted M_{high} = 5.28 [SE = .20]). Although the effect of the actual performance

covariate was not significant ($F = 1.25$, $p = .27$) the effect of the perceived performance covariate was, $F(1, 329) = 443.94$, $p < .001$, $\eta_p^2 = .574$.

The same modified PROCESS model 6 that was used on the ability beliefs index was used to examine the indirect effects of role model success level on future performance expectations. For both male and female participants, high-level role model exposure predicted lower perceived success attainability but did not significantly predict evaluative certainty. Lower perceived success attainability predicted higher evaluative certainty, and both lower perceived success attainability and higher evaluative certainty predicted lower future performance expectations. In other words, mid-level role models produced higher perceived success attainability, which in turn predicted lower evaluative certainty. Together, higher perceived success attainability and lower evaluative certainty protected against low future performance expectations. See Tables 6 and 7.

Attitude Towards Business Economics. Similar to the results on the ability beliefs and future performance expectations indexes, no direct effects emerged on participants' attitudes towards business economics, Fs < 2.27, ps > .13. Women reported similar attitudes towards business economics when the role models had achieved a moderate amount of success as when the role models had achieved a high level of success level (adjusted $M_{mid} = 5.01$ [$SE = .25$], adjusted $M_{high} = 4.86$ [$SE = .25$]). Men also reported similar attitudes across role models success level (adjusted $M_{mid} = 5.36$ [$SE = .27$], adjusted $M_{high} = 5.32$ [$SE = .28$]). The perceived performance covariate was significant, $F(1, 329) = 102.64$, $p < .001$, $\eta_p^2 = .238$, however the actual performance covariate was not, $F < 1$, p = .67.

The indirect effects of role model success level on attitudes towards business economics were examined using the modified PROCESS model 6 and revealed the same pattern of results as the indirect effect analyses for ability beliefs and future performance indexes: two meditational paths were significant for both male and female participants (see Tables 6 and 7). High-level role models predicted lower perceived success attainability than mid-level role models, but role model success level did not influence evaluative certainty. Lower perceived success attainability predicted higher evaluative certainty. In turn, lower perceived success attainability and higher evaluative certainty predicted lower attitudes towards business economics. Described in a different way, mid-level role models predicted higher perceived success attainability, which predicted lower evaluative certainty. Together, higher perceived success attainability and lower evaluative certainty helped to protect against negative attitudes.

Post-Manipulation Performance Rating. Although the effect of the perceived performance covariate was significant, $F(1, 329) = 1179.89$, $p < .001$, $\eta_p^2 = .782$, the effect of the actual performance covariate was not ($F = 2.37$, $p = .13$) and no other effects emerged, Fs < 1.32, ps > .25. After the role model manipulation, women reported similar ratings of their quiz performance whether the role models had achieved a moderate amount of success (adjusted $M_{mid} = 4.72$ [$SE = .12$]) or a high amount of success (adjusted $M_{high} = 4.65$ [$SE = .12$]). Similarly, men reported equal quiz performance ratings after the role model manipulation whether the role model had achieved a moderate (adjusted $M_{mid} = 4.82$ [$SE = .13$]) or a high level of success (adjusted $M_{high} = 4.84$ [$SE = .14$]).

Indirect effects on post-manipulation perceived performance were examined in the same way as the ability beliefs, future performance expectations, and attitudes indexes using a modified version of PROCESS model 6. However, a slightly different patter of results emerged: only one meditational path was significant for both male and female participants. For each gender, exposure to high-level role models predicted lower perceived success attainability than exposure to mid-level role models, but role model success level did not significantly predict evaluative certainty. Lower perceived success attainability predicted higher evaluative certainty. As in the earlier analyses, higher evaluative certainty predicted lower post-manipulation performance ratings, but unlike in the earlier analyses, lower perceived success attainability did not predict lower post-manipulation performance ratings. Put more simply, the higher perceived success attainability and lower evaluative certainty predicted by mid-level role models buffered against low post-manipulation performance ratings. See Tables 6 and 7.

Discussion

Study 4 was conducted to examine whether the level of success achieved by a role model influences the extent to which women and men are certain of their poor performance perceptions. Consistent with hypotheses, women perceived the success of mid-level role models to be more attainable than the success of high-level role models, and higher perceived success attainability predicted lower evaluative certainty. Furthermore, through two meditational paths, mid-level role models helped to alleviate the negative downstream outcomes that typically result from stereotype validation. Mid-level role models, through higher perceived success attainability and lower evaluative certainty, protected against the low ability beliefs, low future performance expectations,

and negative attitudes that that typically result from stereotype validation. Only one meditational path predicted post-manipulation performance ratings: lower evaluative certainty, but not higher perceived success attainability, protected against lower post-manipulation performance ratings.

This last result makes theoretical sense because attitudes and beliefs, such as a person's rating of their quiz performance, are more susceptible to change when they are held with less certainty (e.g., Tormala & Petty, 2002). Although perceived success attainability should influence downstream outcomes related to the present and future, there is no theoretical reason for it to change participants' perceptions of the past.

The results for male participants were inconsistent with hypotheses. Although I did anticipate that role model success level would influence perceived success attainability for men, I did not expect men's higher perceived success attainability to predict lower evaluative certainty as it did. Even though men may not have experienced an initial increase in evaluative certainty after reading the stereotype information, exposure to mid-level role models may have still worked to reduce evaluative certainty even further.

CHAPTER 7

GENERAL DISCUSSION

The effects of role model exposure have been extensively examined in the past. However, the current studies are the first to investigate whether role models influence people's certainty in their performance perceptions. By studying women in a stereotype validation context, I was able to examine how exposure to role models influences evaluative certainty. Past research supported two competing hypotheses: the certainty reduction hypothesis and the certainty enhancement hypothesis. According to the certainty reduction hypothesis, role models should inspire and motivate women after a poor performance. This inspiration and motivation should color their perceptions of their past performance, causing them to be seen in a more favorable light and leading them to be less certain that they had performed poorly. Additionally, female role models should serve as evidence against the negative gender stereotype. This should reduce the validating nature of the negative stereotype, causing women to experience lower evaluative certainty. Alternatively, the certainty enhancement hypothesis predicted that the temporal closeness of women's poor performance evaluations and the salient success of the role models would cause women to be threatened and emotionally hurt by the role model. Because of this threat and emotional pain, women should perceive themselves more negatively, which could manifest in being more certain that they had previously performed poorly.

The current set of studies provides partial support for the certainty reduction hypothesis. Across two studies that used a college student sample, women performed poorly on a science quiz and read about the negative gender-science stereotype. Next,

participants either read about female role models (Studies 1 and 2), successful corporations (Study 1), or male role models (Study 2). In both studies, women for whom female role models were made salient were less certain that they performed poorly on the science quiz than women for whom female role models were not made salient. Furthermore, lower evaluative certainty helped to alleviate the negative downstream outcomes that typically result from stereotype validation (e.g., low ability beliefs, low future performance expectations, and low interest in science-related careers).

However, among one study that used a sample of participants collected through Amazon's Mechanical Turk, women and men reported similar levels of evaluative certainty regardless of whether they read the negative gender-stereotype after their poor performance and regardless of whether they were exposed to female role models or not. The results of this study were inconsistent with hypotheses.

A final study using a Mechanical Turk participant sample found that the level of success achieved by the role model did not directly influence the evaluative certainty reported by stereotype-validated women, but it did indirectly influence evaluative certainty through perceived success attainability. Women exposed to mid-level role models perceived the role models' success to be more attainable than women exposed to high-level role models. Higher perceived success attainability in turn predicted lower evaluative certainty. Moreover, both higher perceived success attainability and lower evaluative certainty helped protect against the negative downstream outcomes that have been found to result from stereotype validation. These results held for male participants, as well, which was counter to predictions.

Limitations

The studies reported here have several limitations that should be addressed. One of these limitations, previously discussed, is the age range of the sample in Study 3. Participant samples collected through Mechanical Turk typically have a higher mean age than samples collected through university subject pools, which was the case in this set of studies. The majority of past research on role models has used college-aged participants, and it is unclear to what extent role models can benefit women who are older and further along in their careers. Theoretically, role models may not be as inspiring and motivating to older women or women who are towards the middle of their careers as they are to young, college-aged women who have not yet begun or not yet chosen their careers. Because their futures are not as full of possibilities, older women may not perceive themselves to have as much potential for future success as younger women. If this is the case, then perhaps the mean age of the sample in Study 3 can help to explain why the salience of role models did not work to reduce women's evaluative certainty.

Alternatively, it may not be that older women do not benefit from role models at all; it could be that older women only benefit from role models who are significantly older than they are. Past research on a college student sample demonstrated that participants benefitted less from role models who were closer in age to them (Lockwood & Kunda, 1997). When role models are closer in age to participants, participants may be more likely to feel threatened by their success and less likely to perceive that the role models' success as attainable. Whereas the college-student participants in Studies 1 and 2 were likely to perceive the role models as older than themselves, it is unclear what age the Mechanical Turk participants perceived the role models in Study 3 to be (no

information was given as to the role models' ages). Perhaps participants in this study perceived the role models to be the same age or younger than themselves. If this was the case, then participants may not have been inspired or motivated by the role models, but may have been threatened by them. This could explain why the presence of role models did not reduce women's evaluative certainty. Unfortunately, with no data to illuminate participants' perceived age of the role models, I cannot discern if this was the case.

Of course, the relative age of role models to participants should be considered for Study 4 as well. In this study, role model age was more apparent because pictures were included with the biographical information. If relative age of participants to role models influences the extent to which role models are beneficial, one might expect the older role models to be more beneficial to female participants than the younger role models in this sample because of their higher mean age (approximately 34 years old). However, in this study, role model age was conflated with level of success; whereas the younger role models had all achieved a moderate level of success, the older role models had all achieved outstanding levels of success that most people could not expect to achieve. Only a study that orthogonally manipulates role model age with role model success level, for a sample varying widely in age, could determine how participant age interacts with role model age. However, relative age should be a consideration in any role model studies moving forward.

Another limitation for Study 3 could be that participants were not naïve to experimental materials. For Studies 1 and 2, most participants were taking the study as part of an Elementary Psychology course requirement and had not participated in psychology studies in past semesters. They were unlikely to have ever participated in a

stereotype validation study or to have read any of the role model materials before. Therefore, I can reasonably assume that the majority of participants in Studies 1 and 2 were truly naïve to the experimental materials. The same assumption cannot be made for the samples of Mechanical Turk participants, particularly in regards to Study 3. For Study 4, I was able to block out all participants who had participated in Study 3 and all participants who had previously taken a business economics-based stereotype validation study for my lab. Although I cannot ensure that participants in Study 4 had not previously viewed its specific role model manipulation materials through a different lab's study, I was able to limit participation to people who had never seen the materials from our lab. However, because of the large number of science-based stereotype validation studies my lab has conducted on Mechanical Turk, I was unable to block out all participants who had previously taken a science-based validation study from participating in Study 3. Thus, there could be a consequential number of participants in Study 3 who had previously been exposed to the experimental materials as well as the debriefing forms from previous science-based stereotype validation studies. This is important because research shows that when studies include non-naïve research participants, effect sizes are markedly lower than when studies involve only naïve participants (Chandler, Paolacci, Peer, Mueller, & Ratliff, 2015). Moving forward, it will be essential to block out potential repeat participants in order to ensure that participants are truly naïve to study materials and hypotheses. This should protect the internal validity of the study and ensure that resources are not being wasted by chasing deflated effects.

An additional concern with the participant samples collected through Amazon's Mechanical Turk is their attrition rates. In Studies 3 and 4, 188 and 43 participants

(respectively) began the study, but dropped out somewhere between the consent process and the attention check items. This corresponds to a 19% attrition rate in Study 3 and a 10% attrition rate in Study 4.[7] Having a high attrition rate can be detrimental to a study because it can result in failure of random assignment (Zhou & Fishbach, 2016). As dropout rates increase, so does the likelihood that participants drop out from experimental conditions at different rates. This condition-dependent attrition is a form of self-selection into conditions, which leads to confounds with the experimental manipulation. Although it is unclear, from the Studies 3 and 4 data sets, at which rates participants dropped out of each condition, the unanticipated effects of experimental condition on participants' actual quiz performance in both studies suggests failure of random assignment; condition-dependent attrition could be the cause. Future research involving Mechanical Turk participants should follow Zhou and Fishbach's (2016) recommendations for reducing attrition rates in an attempt to prevent failure of random assignment.

Another important limitation is that Studies 2 and 4 did not include a control group. Including a condition in which participants were not exposed to role models could have been beneficial in these two studies in order to clarify the extent to which different types of role models affect participants' levels of evaluative certainty. From the design of Study 2, for instance, I can only conclude that women's evaluative certainty is lower after exposure to female role models than after exposure to male role models. Although one may assume, based off of the results of Study 1, that being exposed to female role models reduced evaluative certainty relative to not being exposed to any role models, the data in Study 2 cannot speak to this point. Furthermore, it is unclear to what extent exposure to

[7] Studies 1 and 2, conducted in-person on the University of Iowa campus with a sample of student participants, each had an attrition rate of 0%.

male role models influenced women's evaluative certainty. It could be the case that exposure to male role models does not influence evaluative certainty relative to when participants are not exposed to any role models. Alternatively, it could be the case that exposure to male role models increases evaluative certainty relative to when participants are not exposed to role models. Including a control condition (in which role models were not made salient) could have illuminated which of these opposing possibilities was the case. As the design of Study 2 stands, however, this question cannot be answered.

Additionally, because no control condition was included in Study 4, many questions about the effects of high-level role models remain unanswered. Does exposure to both mid-level and high-level role models decrease evaluative certainty relative to no role model exposure, but the mid-level role models do so to a greater extent? Or does exposure to high-level role models produce levels of evaluative certainty similar to those experienced by women who are not exposed to role models? Or, even still, does exposure to high-level role models increase levels evaluative certainty relative to those experienced by women who are not exposed to role models? Including a control condition could have determined which of these possibilities was the case.

One concern across each of the current studies is how the role model manipulations indirectly influenced the downstream outcomes without any evidence of a direct effect. In none of the analyses reported above did the experimental manipulations directly influence the downstream outcomes (e.g., ability beliefs). Several factors could contribute to this pattern of effects (see Rucker et al., 2011), including sample size, precision of construct measurement, and the presence of undetected suppressor variables. Sample size was likely not an issue across these four studies since large samples were

collected in an effort to achieve high power. Additionally, measurement precision likely did not affect this mediational pattern, considering the alpha levels of evaluative certainty and the downstream outcomes across studies were reasonably high. However, it is possible that another variable, not measured in the current studies, acted to suppress the direct effect of the role model manipulation on the downstream outcomes. This suppressor variable may have worked in the opposite direction of evaluative certainty to influence the downstream outcomes, rendering the direct effect of the role model manipulations non-significant. Future research should make an effort to identify and measure potential suppressor variables in order to investigate their influence on the direct effect of role model manipulations on downstream outcomes of stereotype validation.

Each of the current studies also lacks evidence of several of the potential mechanisms mentioned throughout the introduction. I previously discussed how inspiration and motivation to improve, incited by role models, could perhaps influence participants' levels of evaluative certainty. However, I did not measure participants' levels of inspiration nor the extent to which participants are motivated to improve their future performances. Because I did not collect any empirical data to support the hypothesis that inspiration and motivation work to influence evaluative certainty, these claims remain unsupported.

An additional concern across studies is whether participants experienced demand characteristics. Although I made an effort to dissociate the stereotype validation portion of the experiment from the role model portion of the experiment by telling participants the two parts were unrelated, it is possible that participants may have surmised that the two portions were connected. Participants may have concluded that they were supposed

to react positively to the role models and answered items accordingly, which would mean the significant effects reported above are not true effects. However, several results across these studies suggest that demand characteristics were not likely at play. First, although the role model manipulations in Studies 1, 2, and 3 directly affected participants' evaluative certainty, they did not directly affect the downstream outcomes (e.g., ability beliefs). If demand characteristics of the role model manipulations were in effect, I would expect participants in the role models-present and female role models conditions to respond to all dependent measures more positively than participants in the role models-absent or male role models conditions, but this was not the case. Additionally, if demand characteristics were present, I would expect participants to have responded positively to role models regardless of the source of the participant sample. Although the role model manipulation was identical between Study 1 and Study 3, the role model manipulation did not influence any of the dependent variables in Study 3, including evaluative certainty. There is little reason to anticipate that demand characteristics would be invoked for university students but not Mechanical Turk workers, which suggests that demand characteristics should not be responsible for the significant effects of Study 1. Finally, if participants were responding in line with demand characteristics, I would anticipate that participants would respond equally positively to mid-level and high-level role models. However, as the results of Study 4 show, this was not the case. For these reasons, it is unlikely that demand characteristics were at play in the current research.

Furthermore, I previously asserted one reason why evaluative certainty should not be influenced by role models for male participants: because the presence of the stereotype information should not fit with men's performance perceptions, so men should not

experience an initial increase in evaluative certainty. Without this initial increase in certainty, there should be little room for the salience of role models to further reduce it. However, I did not measure evaluative certainty after exposure to the stereotype information manipulation but before the role model manipulation. Because this measurement was not assessed, I cannot determine if men's evaluative certainty was low after exposure to the stereotype information or whether it remained unchanged between pre- and post-exposure to role models. Without this measurement, my hypothesized reason for a lack of role model effect for male participants remains unsupported.

One final limitation of these studies is that the participants were not allowed to choose their own role models; rather, role models were selected for the participants as part of the study design. Outside of a laboratory, women would likely seek out their own role models rather than have specific role models forced on them. For this reason, these studies have limited external validity. Additionally, from a theoretical standpoint, I would anticipate stronger effects on evaluative certainty if women were allowed to choose whom to look up to after a poor performance. If given the choice of role model, I expect that women would select someone who they are highly similar to (e.g., same race, same career, etc.). Importantly, being highly similar to a role model has been shown to increase inspiration and perceptions of success attainability, and to protect against poor performances on stereotyped tasks (Marx & Ko, 2012).

Choice of role model may be especially important for participants recruited through Mechanical Turk. The perceived age of the role models relative to the participants in the Mechanical Turk sample in Study 3 may have dampened the effects of role model salience. But if given the choice of role model, perhaps relatively older

participants would choose relatively older role models than those we presented in the current studies. By choosing role models older than themselves, middle-aged participants may be more inspired and motivated, leading to a more beneficial effect of the role model and reduced levels of evaluative certainty.

Future Directions

In light of these limitations, future research on this topic should make an effort to investigate role model age effects, include control groups, investigate proposed mediators like inspiration and motivation, and examine whether choice of role model strengthens the effects reported here. However, the current research also prompts several other directions for further experiments.

One question that remains unanswered from these studies is whether exposure to role models can reduce the evaluative certainty of people from other negatively stereotyped social groups. With the current studies, I focused exclusively on women within science and business domains: two areas in which women are stereotyped to perform poorly. However, there are certain contexts in which men are stereotyped to perform poorly, including childcare. Past research has shown that men can become certain of their poor performance on a childcare quiz when they are subsequently reminded of the stereotype that women are inherently better at childcare than men (Clark et al., 2015). Perhaps exposure to male role models who have excelled in childcare would help to reduce men's evaluative certainty in this context. Additionally, Black Americans are stereotyped to perform poorly on academic tasks, including tests of math and verbal abilities, and they can be susceptible to stereotype validation in such contexts (Clark et al., 2015). Perhaps their evaluative certainty could be reduced through exposure to Black

role models who have succeeded in academics. Future research should examine the effectiveness of role models in reducing the evaluative certainty for members of these other negatively stereotyped groups.

One area of future research should examine whether certain qualities of a role model moderate the extent to which role models influence evaluative certainty. For instance, past research has shown that exposure to role models is not beneficial when the role models are perceived as undeserving of their success (e.g., McIntyre, Paulson, Taylor, Morin, & Lord, 2011; Taylor, Lord, McIntyre, & Paulson, 2011). Women who believed that a role model was undeserving of her success performed worse on a math test than women who believed the role model was deserving (Taylor et al., 2011). Exposure to undeserving role models may also lead to greater evaluative certainty than role models perceived as deserving of their success. When women believe that a role model has achieved her success through external causes rather than through hard work and determination, that success may be discounted. Women may not interpret the role model as evidence against the negative stereotype, and they may not perceive the role model's success to be attainable for themselves. For these reasons, women's evaluative certainty may not be reduced through exposure to role models if those role models are perceived as undeserving of their success.

How masculine or feminine a role model appears may also influence the extent to which they help to reduce women's evaluative certainty. Women who interacted with a female role model who was perceived as stereotypical of STEM domains (e.g., having masculine hobbies) reported lower beliefs in their ability to succeed in computer science than women who interacted with a non-stereotypical role model who had more gender-

neutral interests (Cheryan, Siy, Vichayapai, Drury, & Kim, 2011). Also, among middle school girls who did not already identify with STEM fields, those who read about an overtly feminine female role model reported lower expectations that they would take math classes in high school and lower perceived attainability of the role model's success than those who read about a more gender-neutral female role model (Betz & Sekaquaptewa, 2012). Because women may not identify with and perceive themselves to be similar to overly masculine or overly feminine role models, they may not be able to benefit from their salience. Thus, overly masculine and overly feminine role models may represent boundary conditions of role models' effect on evaluative certainty.

Another direction for future research should be to examine individual differences that may moderate the extent to which role models can reduce a person's evaluative certainty. One such individual difference may be how important the stereotyped domain is to a person. The more important a performance domain is to a person, the more they feel threatened when another person outperforms them (Pleban & Tesser, 1981; Tesser, 1988). Because academic majors and careers tend to be an important part of a person's self-concept, women who are majoring in or who have a career in a stereotypically male field may feel threatened by a role model who has succeeded in a stereotypically male field. This threat may manifest in contrast effects away from the role model, potentially causing women who strongly identify with science to report greater evaluative certainty after exposure to a role model than women for whom science is not a part of their identity.

Implicit theories of intelligence (Dweck, 1999) may be another individual difference that moderates the extent to which role models reduce evaluative certainty.

Past role model research has demonstrated that role models are more beneficial to incremental theorists (i.e., people who believe intelligence is malleable and can be changed) than to entity theorists (i.e., people who believe that intelligence is fixed and cannot be changed). Incremental theorists are more likely to perceive challenges as a means to improve and more likely to partake in upward social comparisons than entity theorists (Hoyt, Burnette, & Innella, 2012), which means they may be more likely to look to role models after a failure. Furthermore, incremental theorists are more likely to identify with and admire role models (Hoyt et al., 2012), and are more likely to perceive the role models' success as attainable (Lockwood & Kunda, 1997) than entity theorists. Thus, incremental theorists may be more susceptible to the psychological benefits of role models, including reduced evaluative certainty, than entity theorists.

Finally, individual differences in stigma consciousness may moderate the amount of evaluative certainty experienced after role model exposure. Stigma consciousness is the extent to which a person is concerned about stereotypes of their group (Pinel, 1999). People high in stigma consciousness are more focused on stereotypes of their group and believe they are affected by stereotypes of their group more than people low in stigma consciousness. Because women high in stigma consciousness are more attuned to stereotypes about their gender, they may be more likely to recognize the stereotype-disconfirming nature of female role models. For this reason, women high in stigma consciousness may be less likely to use the negative stereotype when evaluating their certainty in their performance perceptions, resulting in lower evaluative certainty than women low in stigma consciousness would likely experience.

Conclusions

Although the effects of exposure to role models have been extensively studied, particularly for women in stereotypically male domains, the current research adds to this literature by examining whether exposure to role models can influence people's certainty in their performance perceptions. The current research provides mixed evidence that exposure to female role models can reduce women's evaluative certainty following stereotype validation. My hope is that this research will serve as a starting point for future investigations into the boundary conditions of this effect: when, how, and for whom do role models influence evaluative certainty?

APPENDICES

Appendix A: Example essays used to manipulate role model presence in the pilot study, Study 1, and Study 3

Example essay from the role models-present condition
Janet Haley is a highly successful architect who lives and works in London, England. Born in Houston, Texas, she received an undergraduate degree in art history from the University of Houston and a Master's degree in architecture from the Minnesota School of Design. She was the only woman among nine males in her entering class at MSD and feels that she "had to prove myself" to the faculty, most of whom were male. In her graduate courses in architecture, however, Janet was more frequently the top student in her class. On graduating with her master's degree, she worked briefly for the well-known Houghton and associates architectural firm, which has its corporate headquarters in New York City and London, England. Although the firm made Janet responsible for designing a few small structures in the west counties of England, she began to notice that the most prestigious contracts were invariably awarded to her male counterparts in the "old boy network."

Taking a chance, Janet resigned from the firm to start her own architectural consulting business. For the first two years, she was unable to secure a contract. Almost out of money, she briefly considered quitting and retraining for a new career. In 1996, however, the Tate Museum asked Janet Haley to design their modern sculpture facility in Bradford-on-Avon, which recently opened to enormous critical acclaim. Since the opening of that art facility, Janet has been showered with offers, many of them from clients who used to take all their business to Houghton. According to a recent review in the London Times, "Janet Haley's architectural designs constitute a significant creative step forward from the old-school Houghton approach, which is so mired in past ideas that it has lost its utility and appeal."

Example essay from the role models-absent condition
ART-itectural Associates is a highly successful architectural firm based in London, England. In fact the success of the firm is so great that high pressures have been placed upon other firms to live up to the philosophy of "quality before everything else" that ART-itectural has. ART-itectural only hires people who are the best at what they do, and approach their designs and endeavors with the attitude that "Quality is Pride-and Pride is Everything." This attitude is sought after because it is the same attitude that ART-itectural had when the firm was first established. In fact, the founder of ART-itectural was originally employed by the renowned Houghton and associates architectural firm, which has its corporate headquarters in New York City and London, England. Although the Houghton made the ART-itectural's founder responsible for designing the newest structures in England, their attitude for quality was second to making money which was a disappointment.

Taking a chance, ART-itectural Associates-architectural consulting business was formed. For the first two years, ART-itectural Associates was unable to secure a contract. Almost out of money, in 1996, however, the Tate Museum asked ART-itectural to design their modern sculpture facility in Bradford-on-Avon, which recently opened to enormous critical acclaim. Since the opening of that art facility, the firm has been showered with offers, many of them from clients who used to take all their business to Houghton. According to a recent review in the London Times, "ART-itectural designs constitute a significant creative step forward from the old-school Houghton approach, which is so mired in past ideas that it has lost its utility and appeal."

Appendix B: Example photographs and background information for female and male role models used in Study 2

Amanda Baltrig excelled in political science at Colby College. She was so taken with the subject that she had several internships at Capitol Hill over the summers of her college career. While on the Hill, she took an interest in international affairs. Because she was born in Germany, she is fluent in both French and German, a skill which assisted her in her field of interest. Directly after receiving a B.A. from Colby College, Baltrig went on to pursue a Master's degree from Yale University. At present, she is obtaining a Doctorate degree from Yale and ultimately wants to pursue a high-ranking political position in the US government.

Andrew Baltrig excelled in political science at Colby College. He was so taken with the subject that he had several internships at Capitol Hill over the summers of his college career. While on the Hill, he took an interest in international affairs. Because he was born in Germany, he is fluent in both French and German, a skill which assisted him in his field of interest. Directly after receiving a B.A. from Colby College, Baltrig went on to pursue a Master's degree from Yale University. At present, he is obtaining a Doctorate degree from Yale and ultimately wants to pursue a high-ranking political position in the US government.

Appendix C: Example photographs and background information for mid-level and high-level female role models used in Study 4

Eva Rodriguez attended Georgetown University where she studied public health and biology, with the hope of obtaining admission to a top medical school and pursuing her dreams as a doctor. While in Washington, D.C., she volunteered at the local free health clinic and used her Spanish speaking skills to translate for the Spanish speaking patients. Before attending medical school, Rodriguez took two years to work for the Peace Corps, stationed in East Africa, where she provided health care to children with AIDS. She returned to medical school and obtained her MD, and re-kindled her volunteer work at the free clinic. She is presently a resident at Johns Hopkins University medical center, and is doing research on infantile transmissions of HIV with women and children.

Antonia Novello was the first woman and the first Hispanic to become the Surgeon General of the United States from 1990 to 1993. She brought to her work strong empathy for people and used her position and power to alleviate suffering, especially for women and children. Trained as a pediatric nephrologist as well as in public health, she worked in public practice and later in the U.S. Public Health Service. She became a clinical professor of pediatrics at Georgetown University Hospital in 1986. In 1987, she was named the Director for AIDS research at the National Institute of Child Health and Human Development and then became its Deputy Director. As Surgeon General, Dr. Novello was among the first to recognize the need to focus on women with AIDS and on the neonatal transmission of HIV.

TABLES

Table 1. Results of bootstrapping mediation analyses in the pilot study

	Ability Beliefs *B* (*SE*)	Attitudes *B* (*SE*)
Role Model Presence → Evaluative Certainty	-.28 (.39)	-.28 (.39)
Evaluative Certainty → Outcomes	-.40*** (.09)	-.29* (.13)
Estimated Indirect Effect (BC CI)	.11 (.16) (-.1990, .4315)	.08 (.12) (-.1054, .4284)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p < 0.05$, **$p <$ .01, ***$p < .001$.

Table 2. Results of bootstrapping mediation analyses in Study 1

	Ability Beliefs *B* (*SE*)	Attitudes *B* (*SE*)
Role Model Presence → Evaluative Certainty	-1.02*** (.27)	-1.02*** (.27)
Evaluative Certainty → Outcomes	-.32*** (.06)	-.23* (.09)
Estimated Indirect Effect (BC CI)	**.33** (.11) (.1491, .6006)	**.23** (.12) (.0327, .5209)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p <$ 0.05, **$p < .01$, ***$p < .001$.

Table 3. Results of bootstrapping mediation analyses in Study 2

	Ability Beliefs *B* (*SE*)	Attitudes *B* (*SE*)	Career Interest *B* (*SE*)	Post-Manipulation Performance Rating *B* (*SE*)	Post-Manipulation Quiz Performance *B* (*SE*)
Role Model Gender → Evaluative Certainty	-.46* (.23)	-.46* (.23)	-.46* (.23)	-.46* (.23)	-.46* (.23)
Evaluative Certainty → Outcomes	-.29*** (.06)	-.18* (.09)	-.28* (.12)	-.57*** (.05)	-.05 (.04)
Estimated Indirect Effect (BC CI)	**.13** (.08) (.0120, .3290)	**.08** (.06) (.0020, .2681)	**.13** (.09) (.0089, .3698)	**.26** (.14) (.0111, .5620)	.03 (.03) (-.0113, .1090)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p < 0.05$, **$p < .01$, ***$p < .001$.

Table 3 - continued. Results of bootstrapping mediation analyses in Study 2

	IAT D Score *B* (*SE*)
Role Model Presence → Evaluative Certainty	-.46* (.23)
Evaluative Certainty → Outcomes	.00 (.01)
Estimated Indirect Effect (BC CI)	.00 (.01) (-.0154, .0109)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *p < 0.05, **p < .01, ***p < .001.

Table 4. Results of bootstrapping moderated mediation analyses for participants in the stereotype information absent condition in Study 3

	Ability Beliefs *B* (*SE*)	Future Performance Expectations *B* (*SE*)	Post-Manipulation Performance Rating *B* (*SE*)	Post-Manipulation Quiz Performance *B* (*SE*)
		Men		
Role Model Level → Evaluative Certainty	-.12 (.31)	-.12 (.31)	-.12 (.31)	-.12 (.31)
Evaluative Certainty → Downstream Outcome	-.21* (.09)	-.36*** (.09)	-.34*** (.06)	.03 (.06)
Estimated Indirect Effect (BC CI)	.04 (.10) (-.1674, .2407)	.05 (.13) (-.2199, .3097)	.06 (.15) (-.2531, .3422)	.00 (.01) (-.30159, .0449)
		Women		
Role Model Level → Evaluative Certainty	-.26 (.28)	-.26 (.28)	-.26 (.28)	-.26 (.28)
Evaluative Certainty → Downstream Outcome	-.34*** (.09)	-.49*** (.08)	-.57*** (.06)	-.07 (.07)
Estimated Indirect Effect (BC CI)	.07 (.06) (-.0429, .2210)	.13 (.11) (-.1075, .3386)	.16 (.14) (-.1274, .4027)	.01 (.02) (-.0217, .0923)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p < 0.05$, **$p < .01$, ***$p < .001$.

Table 5. Results of the bootstrapping moderated mediation analyses for participants in the stereotype information present condition in Study 3

	Ability Beliefs *B* (*SE*)	Future Performance Expectations *B* (*SE*)	Post-Manipulation Performance Rating *B* (*SE*)	Post-Manipulation Quiz Performance *B* (*SE*)
		Men		
Role Model Level → Evaluative Certainty	.23 (.28)	.23 (.28)	.23 (.28)	.23 (.28)
Evaluative Certainty → Downstream Outcome	-.16 (.11)	-.27* (.10)	-.53*** (.06)	.06 (.07)
Estimated Indirect Effect (BC CI)	-.04 (.07) (-.2150, .0590)	-.07 (.10) (-.2946, .1179)	-.10 (.15) (-.3957, .1744)	.01 (.02) (-.0151, .0755)
		Women		
Role Model Level → Evaluative Certainty	.13 (.26)	.13 (.26)	.13 (.26)	.13 (.26)
Evaluative Certainty → Downstream Outcome	-.23* (.11)	-.47*** (.09)	-.53*** (.06)	-.06 (.07)
Estimated Indirect Effect (BC CI)	-.03 (.07) (-.1826, .1015)	-.05 (.11) (-.2617, .1660)	-.06 (.13) (-.3220, .1903)	-.01 (.02) (-.0593, .0135)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p < 0.05$, **$p < .01$, ***$p < .001$.

Table 6. Results of bootstrapping serial moderated mediation analyses for male participants in Study 4

	Ability Beliefs *B* (*SE*)	Future Performance Expectations *B* (*SE*)	Attitudes *B* (*SE*)	Post-Manipulation Performance Rating *B* (*SE*)
Role Model Level → Success Attainability	1.03* (.52)	1.03* (.52)	1.03* (.52)	1.03* (.52)
Role Model Level → Evaluative Certainty	.22 (.25)	.22 (.25)	.22 (.25)	.22 (.25)
Success Attainability → Evaluative Certainty	-.11** (.04)	-.11** (.04)	-.11** (.04)	-.11** (.04)
Success Attainability → Downstream Outcome	.11** (.04)	.09* (.04)	.17** (.06)	.00 (.02)
Evaluative Certainty → Downstream Outcome	-.45*** (.07)	-.62*** (.08)	-.26* (.13)	-.59*** (.05)
Estimated Indirect Effect Path 1 (BC CI)	**.11** (.06) (.0083, .2661)	**.10** (.06) (.0116, .2363)	**.14** (.08) (.0122, .3483)	.01 (.02) (-.0132, .0596)
Estimated Indirect Effect Path 2 (BC CI)	-0.10 (.12) (-.3606, .0950)	-.14 (.16) (-.4559, .1543)	-.07 (.09) (-.3079, .0665)	-.13 (.15) (-.4443, .1568)
Estimated Indirect Effect Path 3 (BC CI)	**.04** (.02) (.0044, .0995)	**.05** (.03) (.0058, .1402)	**.03** (.02) (.0032, .0905)	**.05** (.03) (.0059, .1365)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Path 1: Role Model Level → Success Attainability → Downstream Outcome. Path 2: Role Model Level → Evaluative Certainty → Downstream Outcome. Path 3: Role Model Level → Success Attainability → Evaluative Certainty → Downstream Outcome. Bold indicates a reliable indirect effect, where BC CI does not include zero. *$p < 0.05$, **$p < .01$, ***$p < .001$.

Table 7. Results of bootstrapping serial moderated mediation analyses for female participants in Study 4

	Ability Beliefs *B* (*SE*)	Future Performance Expectations *B* (*SE*)	Attitudes *B* (*SE*)	Post-Manipulation Performance Rating *B* (*SE*)
Role Model Level → Success Attainability	.81† (.44)	.81† (.44)	.81† (.44)	.81† (.44)
Role Model Level → Evaluative Certainty	.04 (.18)	.04 (.18)	.04 (.18)	.04 (.18)
Success Attainability → Evaluative Certainty	-.06† (.03)	-.06† (.03)	-.06† (.03)	-.06† (.03)
Success Attainability → Downstream Outcome	.11** (.04)	.10** (.03)	.12* (.06)	.02 (.02)
Evaluative Certainty → Downstream Outcome	-.42*** (.09)	-.57*** (.08)	-.32* (.14)	-.58*** (.05)
Estimated Indirect Effect Path 1 (BC CI)	**.09** (.05) (.0081, .2389)	**.08** (.05) (.0076, .2143)	**.12** (.07) (.0113, .3236)	.01 (.02) (-.0076, .0646)
Estimated Indirect Effect Path 2 (BC CI)	-0.04 (.08) (-.2121, .1066)	-.05 (.11) (-.2726, .1552)	-.03 (.06) (-.1701, .0768)	-.05 (.11) (-.2515, .1504)
Estimated Indirect Effect Path 3 (BC CI)	**.03** (.02) (.0032, .0859)	**.04** (.03) (.0032, .1171)	**.02** (.02) (.0015, .0719)	**.04** (.03) (.0037, .1109)

Note: Indirect effects reported with 95% bias-corrected confidence intervals (BC CI). Path 1: Role Model Level → Success Attainability → Downstream Outcome. Path 2: Role Model Level → Evaluative Certainty → Downstream Outcome. Path 3: Role Model Level → Success Attainability → Evaluative Certainty → Downstream Outcome. Bold indicates a reliable indirect effect, where BC CI does not include zero. $^{\dagger}p < 0.10$, $*p < 0.05$, $**p < .01$, $***p < .001$.

FIGURES

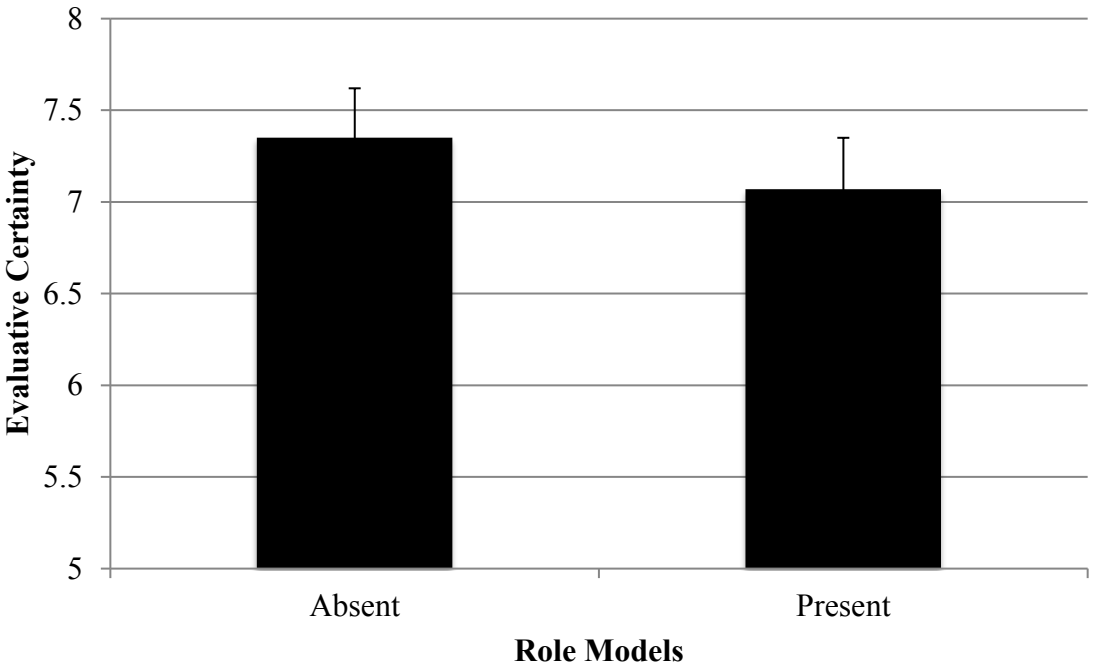

Figure 1. Adjusted mean evaluative certainty as a function of role model condition in the pilot study. Error bars represent standard errors.

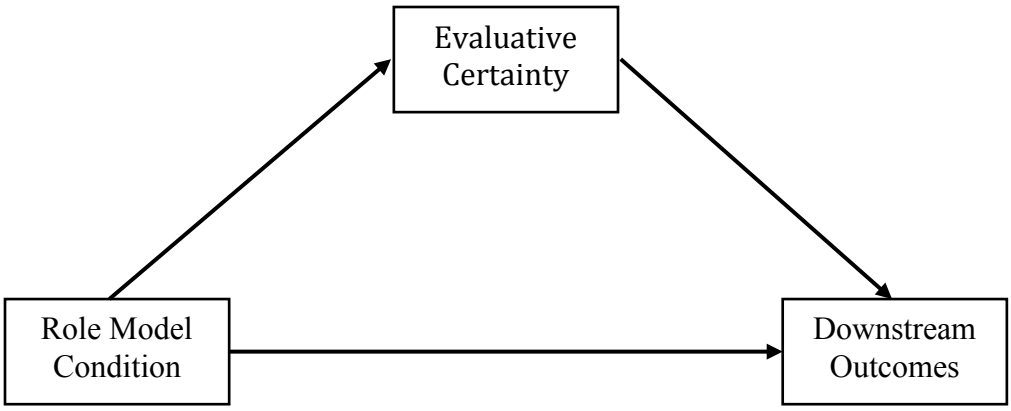

Figure 2. PROCESS model 4 used in the pilot study, Study 1, and Study 2.

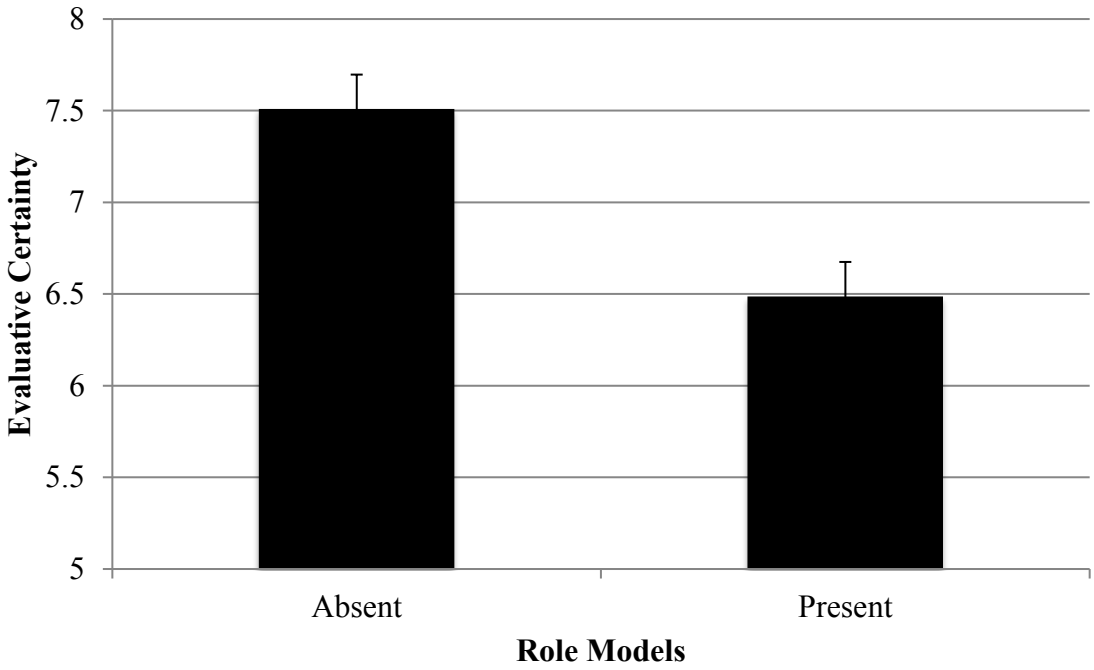

Figure 3. Mean evaluative certainty as a function of role model condition in Study 1. Error bars represent standard errors.

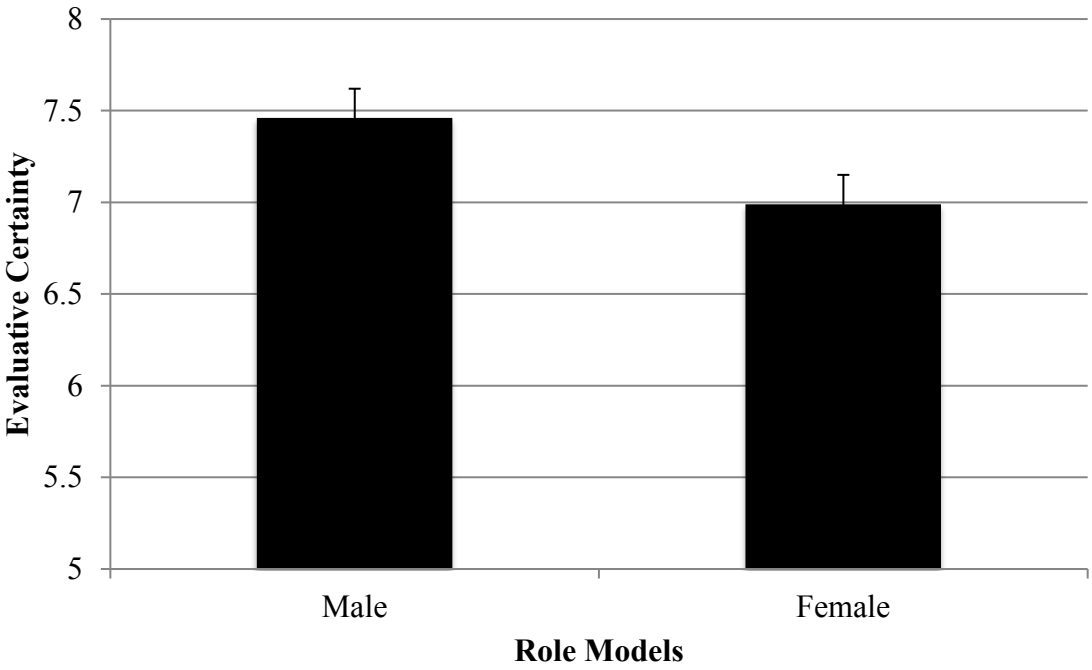

Figure 4. Mean evaluative certainty as a function of role model condition in Study 2. Error bars represent standard errors.

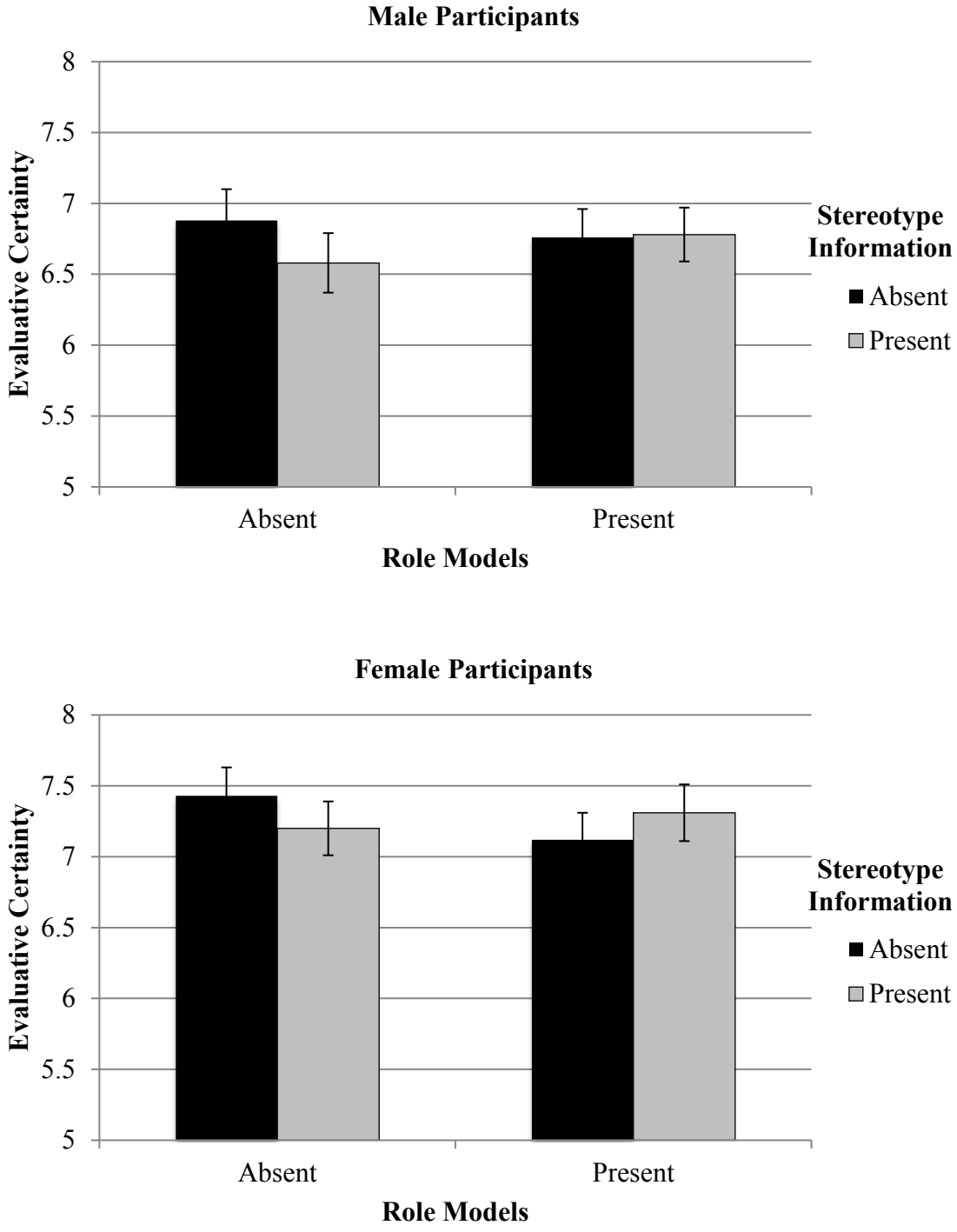

Figure 5. Adjusted mean evaluative certainty as a function of role model condition, stereotype information, and participant gender in Study 3. Error bars represent standard errors.

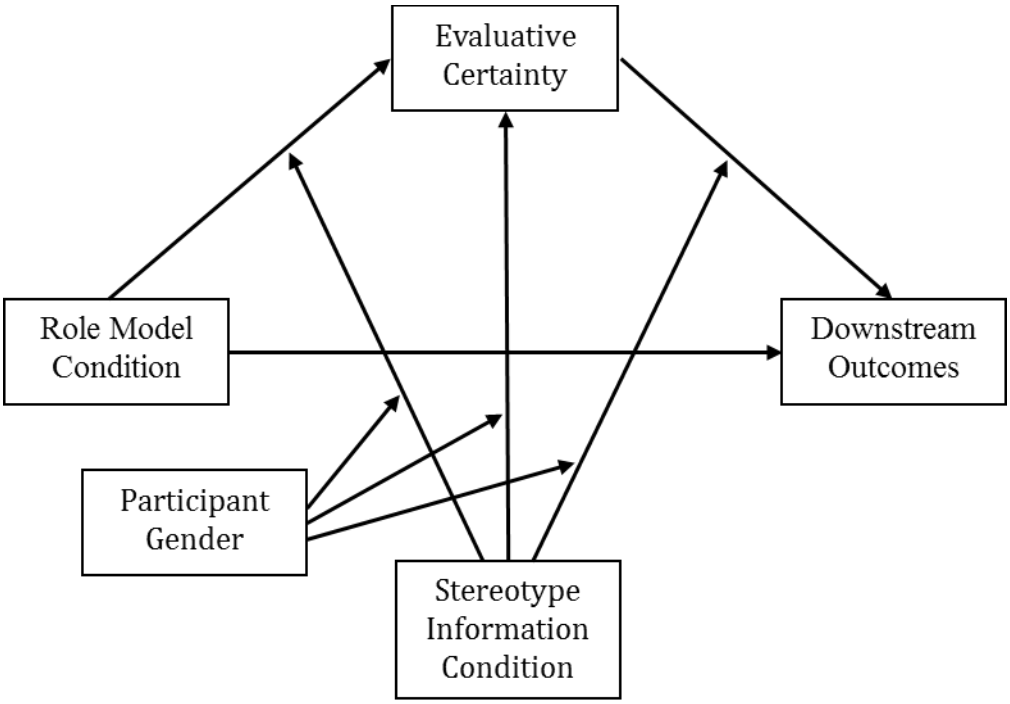

Figure 6. PROCESS model 73 used in Study 3.

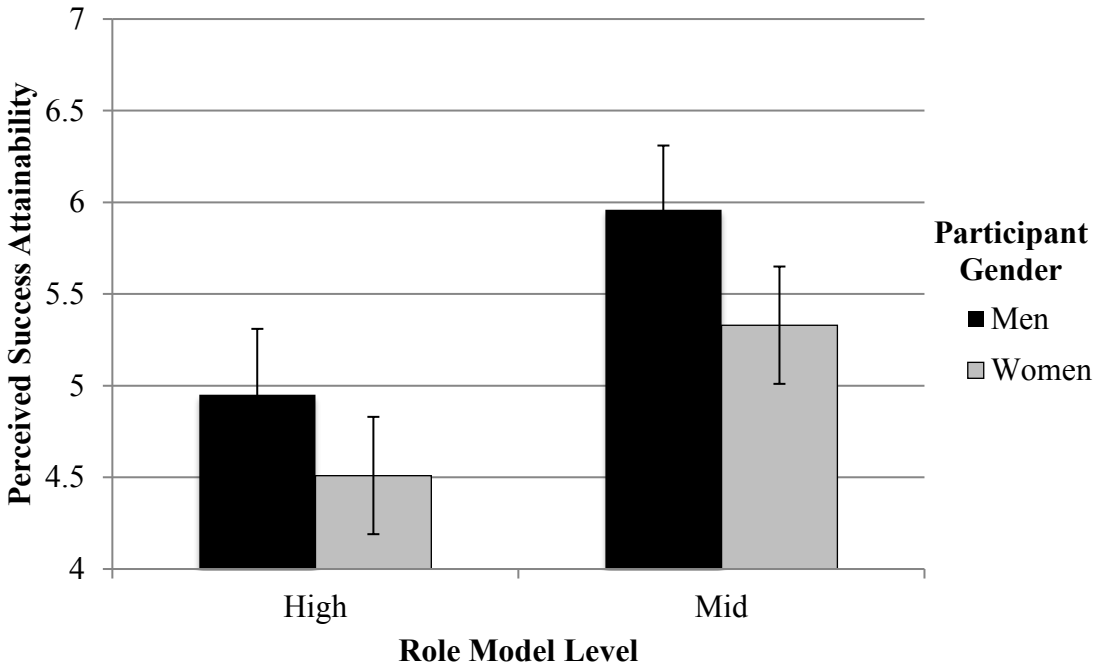

Figure 7. Adjusted mean perceived success attainability as a function of role model condition and participant gender in Study 4. Error bars represent standard errors.

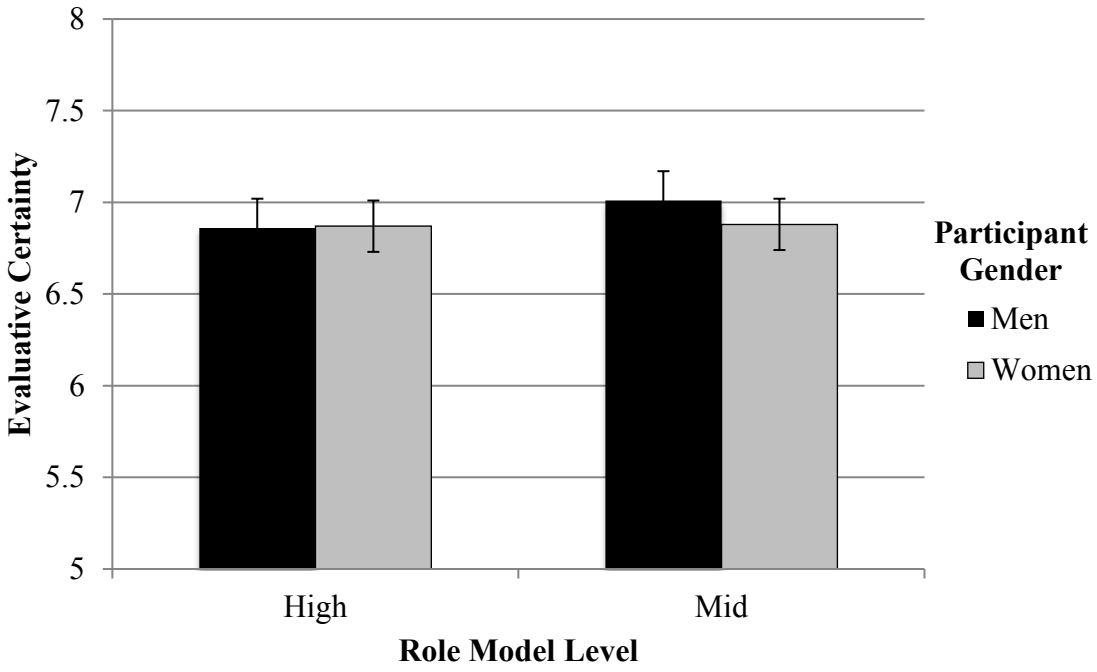

Figure 8. Adjusted mean evaluative certainty as a function of role model condition and participant gender in Study 4. Error bars represent standard errors.

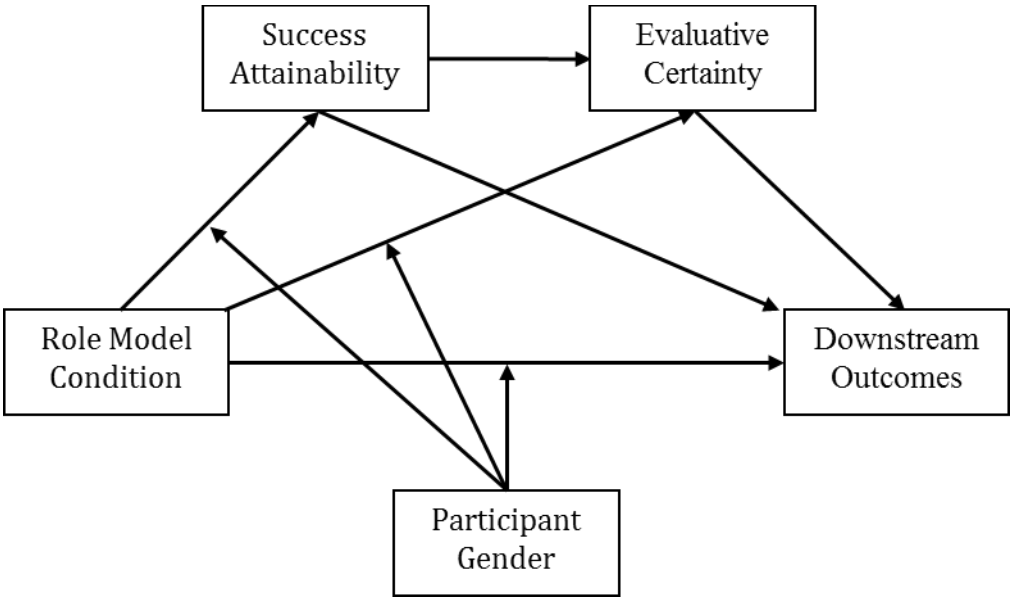

Figure 9. Modified PROCESS model 6 used in Study 4.

REFERENCES

Asgari, S., Dasgupta, N., & Stout, J. G. (2012). When Do Counterstereotypic Ingroup Members Inspire Versus Deflate? The Effect of Successful Professional Women on Young Women's Leadership Self-Concept. *Personality and Social Psychology Bulletin*, *38*(3), 370–383.

Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology*, *35*(3), 125–139.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*(4), 637–653.

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, *3*(6), 738–746.

Bizer, G. Y., Tormala, Z. L., Rucker, D. D., & Petty, R. E. (2006). Memory-based versus on-line processing: Implications for attitude strength. *Journal of Experimental Social Psychology*, *42*(5), 646–653.

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, *76*(3), 420–430.

Brewer, M. B., & Weber, J. G. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, *66*(2), 268–275.

Brickman, P., & Bulman, R. J. (1977). Pleasure and pain in social comparison. In J. M. Suls & R. L. Miller (Eds.), *Social Comparison Processes: Theoretical and Empirical Perspectives* (pp. 149–186). Washington, D.C.: Hemisphere.

Chandler, J., Paolacci, G., Peer, E., Mueller, P., & Ratliff, K. A. (2015). Using nonnaive participants can reduce effect sizes. *Psychological Science*, *26*(7), 1131–1139.

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, *2*(6), 656–664.

Clark, J. K., & Thiem, K. C. (under review). Stereotype validation and intellectual performance: Negative implications for future achievement.

Clark, J. K., Thiem, K. C., Barden, J., Stuart, J. O., & Evans, A. T. (2015). Stereotype validation: The effects of activating negative stereotypes after intellectual performance. *Journal of Personality and Social Psychology*, *108*(4), 531–552.

Clark, J. K., Thiem, K. C., Hoover, A. E., & Habashi, M. M. (2017). Gender stereotypes and intellectual performance: Stigma consicousness as a buffer against stereotype validation. *Journal of Experimental Social Psychology*, *68*, 185–191.

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*(1), 51–69.

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, *40*(5), 642–658.

Dijksterhuis, A., & Van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology*, *74*(4), 865–877.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Psychology Press.

Fazio, R. H., & Zanna, M. P. (1978). Attitudinal qualities relating to the strength of the attitude-behavior relationship. *Journal of Experimental Social Psychology*, *14*, 398–408.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Friend, R. M., & Gilbert, J. (1973). Threat and fear of negative evalution as determinants of locus of social comparison. *Journal of Personality*, *41*(3), 328–340.

Gonzales, P., Guzmán, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., … Commissioner, R. L. (2004). *Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003*. Washington, D.C.

Hayes, A. F. (2014). PROCESS (Version 2.12). Retrieved from http://www.afhayes.com/introduction-to-mediation-moderation-and-conditional-process-analysis.html

Hoyt, C. L., Burnette, J. L., & Innella, A. N. (2012). I can do that: The impact of implicit theories on leadership role model effectiveness. *Personality and Social Psychology Bulletin*, *38*(2), 257–268.

Hoyt, C. L., & Simon, S. (2011). Female leaders: Injurious or inspiring role models for women? *Psychology of Women Quarterly*, *35*(1), 143–157.

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*(1), 16–32.

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, *80*(6), 942–958.

Latu, I. M., Mast, M. S., Lammers, J., & Bombari, D. (2013). Successful female leaders empower women's behavior in leadership tasks. *Journal of Experimental Social Psychology,* (49), 444-448.

Lockwood, P. (2006). "Someone like me can be successful": Do college students need same-gender role models? *Psychology of Women Quarterly*, *30*(1), 36–46.

Lockwood, P., & Kunda, Z. (1997). Superstars and Me: Predicting the Impact of Role Models on the Self. *Journal of Personality and Social Psychology*, *73*(1), 91–103.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*(1), 83–104.

Major, B., & Forcey, B. (1985). Social comparisons and pay evaluations: Preferences for same-sex and same-job wage comparisons. *Journal of Experimental Social Psychology*, *21*, 393–405.

Major, B., Sciacchitano, A. M., & Crocker, J. (1993). In-group versus out-group comparisons and self-esteem. *Personality and Social Psychology Bulletin*, *19*(6), 711–721.

Major, B., Testa, M., & Blysma, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: contemporary theory and research* (pp. 237–260). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Marx, D. M., & Ko, S. J. (2012). Superstars "like" me: The effect of role model similarity on performance under threat. *European Journal of Social Psychology*, *42*(7), 807–812.

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183–1193.

McIntyre, R. B., Lord, C. G., Gresky, D. M., Eyck, L. L. T., Frye, G. D. J., & Bond, C. F. (2005). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. *Current Research in Social Psychology*, *10*(9), 116–136.

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*(1), 83–90.

McIntyre, R. B., Paulson, R. M., Taylor, C. A., Morin, A. L., & Lord, C. G. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology*, *41*(3), 301–311.

Meichenbaum, D. H. (1971). Examination of model characteristics in reducing avoidance behavior. *Journal of Personality and Social Psychology*, *17*(3), 298–307.

Nadler, A., & Fisher, J. D. (1986). The role of threat to self-esteem and perceived control in recipient reaction to help: Theory development and empirical validation. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (pp. 81–121). Orlando, FL: Academic Press.

Nosanchuk, T. A., & Erickson, B. H. (1985). How high is up? Calibrating social comparison in the real world. *Journal of Personality and Social Psychology*, *48*(3), 624–634.

Parks-Stamm, E. J., Heilman, M. E., & Hearns, K. A. (2008). Motivated to penalize: Women's strategic rejection of successful women. *Personality and Social Psychology Bulletin*, *34*(2), 237–247.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*(1), 114–128.

Pleban, R., & Tesser, A. (1981). The effects of relevance and quality of another's performance on interpersonal closeness. *Social Psychology Quarterly*, *44*(3), 278–285.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, *5*(6), 359–371.

Rudman, L. a., & Phelan, J. E. (2010). The Effect of Priming Gender Roles on Women's Implicit Gender Beliefs and Career Aspirations. *Social Psychology*, *41*(3), 192–202.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008.

Samuel, W. (1973). On clarifying some interpretations of social comparison theory. *Journal of Experimental Social Psychology*, *9*, 450–465.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, *77*(3), 313–322.

Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology*, *42*(2), 281–291.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, *95*(6), 323–332.

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*(2), 255–270.

Suls, J. M., & Tesch, F. (1978). Students' preferences for information about their test performance: A social comparison study. *Journal of Applied Social Psychology*, *8*(2), 189–197.

Taylor, C. A., Lord, C. G., McIntyre, R. B., & Paulson, R. M. (2011). The Hillary Clinton effect: When the same role model inspires or fails to inspire improved performance under stereotype threat. *Group Processes & Intergroup Relations*, *14*(4), 447–459.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (pp. 181–227). San Diego, CA: Academic Press.

Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty. *Journal of Personality and Social Psychology*, *83*(6), 1298–1313.

Tormala, Z. L., & Rucker, D. D. (2007). Attitude certainty: A review of past findings and emerging perspectives. *Social and Personality Psychology Compass*, *1*(1), 469–492.

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

von Hippel, C., Sekaquaptewa, D., & McFarlane, M. (2015). Stereotype threat among women in finance: Negative effects on identity, workplace well-being, and recruiting. *Psychology of Women Quarterly*, *39*(3), 405–414.

Wheeler, L. (1966). Toward a theory of behavioral contagion. *Psychological Review*, *73*(2), 179–192.

Wheeler, L., & Koestner, R. (1984). Performance evaluation: On choosing to know the related attributes of others when we know their performance. *Journal of Experimental Social Psychology*, *20*, 263–271.

Wheeler, L., Martin, R., & Suls, J. (1997). The proxy model of social comparison for self-assessment of ability. *Personality and Social Psychology Review*, *1*(1), 54–61.

Wheeler, L., Shaver, K. G., Jones, R. A., Goethals, G. R., Robinson, J. E., Gruder, C. L., & Butzine, K. W. (1969). Factors determining choice of a comparison other. *Journal of Experimental Social Psychology*, *5*, 219–232.

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*(2), 231–248.

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504.