Personality and Social Psychology Bulletin
2018, Vol. 44(3) 359–383

DOI: 10.1177/0146167217741312
journals.sagepub.com/home/pspb
SAGE

# Positive Feedback From Male Authority Figures Boosts Women's Math Outcomes

**Lora E. Park[1], Cheryl L. Kondrak[1], Deborah E. Ward[1], and Lindsey Streamer[2]**

**Abstract**

People often search for cues in the environment to determine whether or not they will be judged or treated negatively based on their social identities. Accordingly, feedback from gatekeepers—members of majority groups who hold authority and power in a field—may be an especially important cue for those at risk of experiencing social identity threat, such as women in math settings. Across a series of studies, women who received positive ("Good job!") versus objective (score only) feedback from a male (vs. female) authority figure in math reported greater confidence; belonging; self-efficacy; more favorable Science, Technology, Engineering, and Mathematics (STEM) attitudes/identification/interest; and greater implicit identification with math. Men were affected only by the type of math feedback they received, not by the source of feedback. A meta-analysis across studies confirmed results. Together, these findings suggest that positive feedback from gatekeepers is an important situational cue that can improve the outcomes of negatively stereotyped groups.

**Keywords**

women and math, STEM, feedback, social identity threat, gender stereotypes

Received March 3, 2017; revision accepted October 12, 2017

> Not even the math professor who supervised my senior thesis urged me to go on for a Ph.D. I had spent nine months missing parties, skipping dinners and losing sleep, trying to figure out why waves—of sound, of light, of anything—travel in a spherical shell. When at last I found the answer, I knocked triumphantly at my adviser's door. Yet I don't remember him praising me in any way. I was dying to ask if my ability to solve the problem meant that I was good enough to make it as a theoretical physicist. But I knew that if I needed to ask, I wasn't.
>
> —Eileen Pollack (October 2013), *New York Times Magazine*

In the article, "Why Are There Still So Few Women in Science?" author Eileen Pollack discusses the feelings of uncertainty and self-doubt that women often experience in the traditionally male-dominated fields of math and science. Drawing upon her experience, Pollack reveals how concerns about her ability to do well in these domains—coupled with a lack of perceived support and encouragement to pursue these fields—ultimately led her to switch career paths and abandon her dreams of becoming a theoretical physicist.

Pollack's doubts about her ability resonate with broader cultural stereotypes, in which women in U.S. society are judged to be less quantitatively skilled than men, and math and science are viewed as stereotypically "masculine" domains (Eccles, Jacobs, & Harold, 1990; Nosek, Banaji, & Greenwald, 2002; Park, Cook, & Greenwald, 2002; Schmader, Johns, & Forbes, 2008; Spencer, Steele, & Quinn, 1999; Steele, Spencer, & Aronson, 2002). To combat such beliefs, social psychologists have recommended exposure to successful ingroup peers and role models (Dasgupta, 2011; Dasgupta, McManus Scircle, & Hunsinger, 2015; Marx & Roman, 2002; McIntyre, Paulson, & Lord, 2003; Stout, Dasgupta, Hunsinger, & McManus, 2011; Young, Rudman, Buettner, & McLean, 2013). Indeed, female undergraduate students show stronger implicit associations between women and math when they see female math and science professors (Dasgupta & Asgari, 2004). They also report higher self-efficacy; more favorable attitudes toward science, technology, engineering, and mathematics (STEM); and greater motivation to pursue degrees and careers in these fields when they see successful female peers and experts in STEM (Dasgupta, 2011; Stout et al., 2011). These and other studies suggest that women's educational and career aspirations are strongly shaped by same-sex role models (Lockwood, 2006).

[1]University at Buffalo, The State University of New York, NY, USA
[2]MediaScience

**Corresponding Author:**
Lora E. Park, University at Buffalo, The State University of New York, 344 Park Hall, Buffalo, NY 14260, USA.
Email: lorapark@buffalo.edu

Although exposure to successful ingroup members is beneficial, focusing on only one segment of society overlooks the potentially positive impact of interacting with outgroup members (Akcinar, Carr, & Walton, 2011; Cheryan, Drury, & Vichayapai, 2013; Cheryan, Siy, Vichayapai, Drury, & Kim, 2011; Drury, Siy, & Cheryan, 2011; Mendoza-Denton & Page-Gould, 2008). Given that women are numerically underrepresented in STEM fields, they are likely to encounter outgroup, majority group members (i.e., men) while taking STEM courses and beyond, if they choose to pursue careers in these fields. Majority outgroup members are influential because they are thought to serve as “symbolic gatekeepers,” determining through their attitudes and behaviors who belongs in a field and who does not (Akcinar et al., 2011). The feedback that women receive from male gatekeepers in math may therefore be especially influential in shaping their self-evaluations and attitudes toward, identification with, and interest in STEM.

In sum, a more comprehensive approach to understanding and improving women’s outcomes in STEM fields is to consider how contact *across* group boundaries may benefit those who are at risk of being judged negatively based on their social identity, such as women in math settings. To this end, the present research examined intergroup contact in the context of a common experience that students face during their academic careers: receiving feedback from authority figures.

## Situational Cues Can Invoke Social Identity Threat

Cues in the environment can elicit concerns about the value of one’s social identity and expectations of being evaluated negatively in light of this identity (Abrams & Hogg, 1999; Goffman, 1963; Purdie-Vaughns, Steele, Davies, Ditlmann, & Crosby, 2008; Steele et al., 2002). Indeed, individuals often view themselves in terms of the social identity that is most stereotyped or stigmatized in a given context (Branscombe, Ellemers, Spears, & Doosje, 1999). Given that women in North American culture are viewed as having poor mathematical skills (Marx & Roman, 2002; Nosek et al., 2002; O’Brien & Crandall, 2003; Schmader et al., 2008; Spencer et al., 1999), they may be vigilant for noticing cues in math-related settings that convey whether or not they will be judged or treated negatively based on their social identity.

Cues in the environment can range from the numerical representation of individuals who share one’s identity (e.g., Murphy, Steele, & Gross, 2007), to physical features of the environment that signal who belongs and who does not (e.g., Cheryan, Plaut, Davies, & Steele, 2009), to interactions with members of the outgroup (e.g., Logel et al., 2009). Indeed, heightened attention—to the possibility and probability of experiencing identity threat—is most likely to occur in settings where outgroup (vs. ingroup) members are present (Wout, Shih, Jackson, & Sellers, 2009). Along these lines, women performed worse on a math test and reported lower performance self-esteem when they took a math test in the presence of an outgroup versus ingroup member (i.e., a male vs. female experimenter; Marx & Roman, 2002). Similarly, Inzlicht and Ben-Zeev (2000) found that women performed worse on a difficult math (but not verbal) test when they were the only woman in a room with all male test takers. Sekaquaptewa and Thompson (2003) further showed that women who experienced solo status and stereotype threat simultaneously performed poorly on a math test, whereas men’s performance was unaffected by the gender composition of the test-taking situation. In another set of studies, Cheryan and colleagues (2009) found that women reported less belonging and interest in computer science when they saw objects in computer science environments that conveyed masculine stereotypes associated with the outgroup (e.g., Star Trek posters, video games).

Beyond mere exposure, identity threat can occur in the context of actual interactions with outgroup members. For example, Logel and colleagues (2009) demonstrated that female engineering students who interacted with a male engineering student who displayed subtly sexist behaviors (i.e., acted in a dominant, flirtatious manner), or with a confederate who was trained to behave in a sexist way, performed worse on an engineering test than women who interacted with men who behaved in a nonsexist way. Indeed, the mere suggestion that a male instructor is sexist leads women, but not men, to report a less positive instructional experience and to perform worse on a standardized logic test than women who are not exposed to a suggestion of sexism (Adams, Garcia, Purdie-Vaughns, & Steele, 2006).

If cues in the environment can undermine interest and achievement for members of certain groups, then changing such cues could create environments that are safe for those who are at risk of experiencing identity threat. In identity-safe environments, individuals do not feel at risk of being judged or treated negatively based on their social identity (Davies, Spencer, & Steele, 2005; Markus, Steele, & Steele, 2002; Purdie-Vaughns et al., 2008). For example, telling participants that a test is nondiagnostic, or unable to assess group-based differences in ability, reduces the relevance of the stereotype from the setting and diminishes concerns about being perceived in a negative light (Blascovich, Spencer, Quinn, & Steele, 2001; Spencer et al., 1999; Steele & Aronson, 1995).

Even when stigmatized identities and corresponding stereotypes are activated, assuring individuals that they will not be viewed through the lens of their identity “clears the air” and prevents declines in motivation and aspirations. For example, whereas women who saw gender-stereotypic TV commercials reported less interest in quantitative domains and in pursuing traditionally masculine careers and roles (Davies, Spencer, Quinn, & Gerhardstein, 2002; Davies et al., 2005), women who were told that no gender differences exist in leadership and problem-solving ability did not show diminished interest in leadership positions (Davies et al., 2005).

## Feedback as a Situational Cue to Reduce Social Identity Threat

Feedback is defined as "information provided by an agent . . . regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). Research on *wise feedback* suggests that for members of underrepresented or negatively stereotyped groups, identity safety can be created through the type of feedback given. In an early demonstration of this idea, Cohen, Steele, and Ross (1999) found that when mentoring minority students in identity threatening settings, such as Black college students being evaluated by a White evaluator, feedback was most effective when it invoked high standards and assured students of their ability to meet those standards. Black students who received unbuffered critical feedback on their essay from a White evaluator showed less motivation and effort than White students. However, when the feedback conveyed both high standards and the belief that students could reach those standards, Black students responded just as favorably as White students.

More recently, Yeager and colleagues (2014) showed that wise feedback increased minority adolescent students' likelihood of submitting a revised essay and improved the quality of their final drafts. These effects were especially strong among minority students who were more mistrusting of school, suggesting that wise feedback may be most effective for those at risk of disengaging from academics. Overall, these studies support the idea that subtle cues in the environment—from how a test is portrayed to the type of feedback given—can create environments that foster interest and motivation, rather than distrust and disengagement.

In the present research, we propose that even brief, positive feedback may benefit individuals in situations where they could be negatively evaluated based on their social identity, such as women in math settings. In such settings, women may be motivated to determine whether or not they will be judged by others, especially by outgroup members, based on their identity (Schmader & Croft, 2011; Steele et al., 2002; Walton & Cohen, 2007; Wout et al., 2009). To the extent that certain contexts activate concerns about being evaluated based on one's identity, receiving positive feedback from gatekeepers could be a promising antidote to alleviating such concerns.

In fact, a common reason why college students say they drop out of STEM majors and switch to other majors is because of the highly competitive, impersonal, and intimidating environment in STEM courses (Seymour & Hewitt, 1997; Vogt, Hocevar, & Hagedorn, 2007). For example, the more engineering students perceive faculty distance (i.e., feel inferior to professors, feel uneasy asking professors for help), the lower their self-efficacy, academic confidence, and grade point average (GPA; Vogt, 2008). Students also exhibit less effort and adopt fewer learning strategies when they are in the presence of "distant" faculty or "chilly" classroom environments (Vogt et al., 2007). Receiving positive feedback could, therefore, challenge the "chilly climate" that students often experience in STEM settings. Along these lines, research suggests that positive classroom dynamics and student–faculty interactions can enhance students' learning and engagement (Umbach & Wawrzynski, 2005). For example, college students in an engineering course showed greater self-efficacy, effort, and critical thinking in relation to their course material the more they perceived their instructors to be supportive, approachable, and accessible (Vogt, 2008). A key element of motivation, then, is self-efficacy, or beliefs about one's ability to achieve an outcome (Bandura, 1982).

Self-efficacy is based not only on one's performance in a domain, but also on aspects of the environment, such as the feedback one receives. Positive feedback about one's ability or performance should boost confidence and perceptions of competence in a domain. Consistent with this notion, students who received positive feedback during a puzzle task spent more time on the same task during a free-choice session than those who received no feedback (Deci, 1971), and this same pattern of results has been replicated in subsequent studies (Harackiewicz, 1979; Vallerand & Reid, 1998).

Cues in the environment can also shape feelings of belonging, especially for those who are at risk of social identity threat. For example, although female STEM majors reported lower sense of belonging and less interest in participating in a STEM conference when they saw a conference video portraying an unbalanced ratio of male to female participants, this was not the case when women saw a video of a gender-balanced ratio of participants (Murphy et al., 2007). Similarly, Cheryan and colleagues (2009) found that a small tweak in the environment—replacing objects in a stereotypically masculine computer science classroom with neutral items, such as nature posters and a phone book—was sufficient to increase women's interest in computer science to a level comparable with men's.

If women show increased vigilance in situations where they are susceptible to identity threat, then positive feedback could serve as an important contextual cue that creates identity safety and increases self-efficacy, belonging, and attitudes/identification/interest toward math and related fields. Specifically, the *type, source*, and *domain* of feedback may affect women's math-related outcomes. Feedback may be most beneficial when it (a) is positive versus objective (i.e., receiving one's test score with the written comment "Good job!" vs. receiving one's test score only), (b) comes from a perceived gatekeeper in the field (i.e., a male vs. female authority figure in math), and (c) is given in a domain in which individuals could be judged negatively in light of their social identity (i.e., women in math vs. verbal settings).

**Table 1.** Results of Paired *t* Tests and Descriptive Statistics for Pilot Study.

| | *M* | *SD* | *t* | *df* | *p* | Cohen's *d* |
|---|---|---|---|---|---|---|
| Field of math | | | | | | |
| Approximately what percentage of college professors in math are male? | 5.78 | (1.53) | 12.33 | 119 | *** | 1.89 |
| Approximately what percentage of college professors in math are female? | 2.98 | (1.43) | | | | |
| Approximately what percentage of college students in math are male? | 5.56 | (1.44) | 8.76 | 119 | *** | 1.35 |
| Approximately what percentage of college students in math are female? | 3.51 | (1.59) | | | | |
| Of the most influential (i.e., powerful/competent/capable) professionals in math-related occupations, what percentage are likely to be male? | 5.76 | (1.59) | 10.68 | 119 | *** | 1.68 |
| Of the most influential (i.e., powerful/competent/capable) professionals in math-related occupations, what percentage are likely to be female? | 3.20 | (1.45) | | | | |
| Field of nursing | | | | | | |
| Approximately what percentage of college professors in nursing are male? | 2.38 | (1.66) | −16.11 | 119 | *** | 2.34 |
| Approximately what percentage of college professors in nursing are female? | 6.42 | (1.79) | | | | |
| Approximately what percentage of college students in nursing are male? | 2.38 | (1.74) | −14.69 | 119 | *** | 2.33 |
| Approximately what percentage of college students in nursing are female? | 6.56 | (1.85) | | | | |
| Of the most influential (i.e., powerful/competent/capable) professionals in nursing occupations, what percentage are likely to be male? | 2.91 | (1.90) | −11.58 | 119 | *** | 1.74 |
| Of the most influential (i.e., powerful/competent/capable) professionals in nursing occupations, what percentage are likely to be female? | 6.12 | (1.78) | | | | |

*Note.* Response scales were 0 = 0%-10%, 1 = 11%-20%, 2 = 21%-30%, 3 = 31%-40%, 4 = 41%-50%, 5 = 51%-60%, 6 = 61%-70%, 7 = 71%-80%, 8 = 81%-90%, 9 = 91%-100%.
***$p < .001$.

## Preliminary Study

To provide initial evidence for the idea that men are perceived as gatekeepers in math and related fields, we conducted a study with 120 undergraduate students in the introductory psychology subject pool at a large Eastern public university (60 women, $M_{age}$ = 18.87 years, *SD* = 1.02 years). The sample was 49% White, 32% Asian, 7% Black, and 12% Other ethnicities. Participants responded to the following items: (a) "In general, approximately what percentage of college professors in math are male (female)?" (b) "In general, approximately what percentage of college students in math are male (female)?" and (c) "Out of the most influential (i.e., powerful/competent/capable) professionals in math-related occupations, what percentage is likely to be male?" using scales 0 = 0% to 10%, 1 = 11% to 20%, 2 = 21% to 30%, 3 = 31% to 40%, 4 = 41% to 50%, 5 = 51% to 60%, 6 = 61% to 70%, 7 = 71% to 80%, 8 = 81% to 90%, 9 = 91% to 100%.

To examine whether these perceptions were unique to math, we used the same question format and response scales to ask about perceptions of men and women in the field of nursing. We selected nursing as a comparison category because it is a traditionally feminine career in which women are well represented (Diekman, Brown, Johnston, & Clark, 2010; U.S. Department of Labor, 2009). Finally, participants were asked to list up to three subject areas in which men tend to have greater authority, expertise, and credibility than women. They then listed up to three subject areas in which they thought women have greater authority, expertise, and credibility than men.

Descriptive statistics and results of paired *t* tests are shown in Table 1. As predicted, participants perceived more professors, college students, and professionals in math to be male than female. In the field of nursing, the findings were reversed: Participants perceived more professors, college students, and professionals in nursing to be female than male. For the open-ended responses, we first coded the domains based on the most frequently mentioned items, then conducted a frequency count of the number of times participants mentioned the top three domains that were categorized as the most commonly listed subject areas. Men were perceived to have greater authority, expertise, and credibility than women in STEM fields (mentioned 73 times), business (27 times), and sports (14 times). Women were perceived to have greater authority and expertise in the arts/humanities (mentioned 36 times), nursing (35 times), and teaching/education (33 times).

Overall, these preliminary findings suggest that men are perceived as the majority in STEM fields, but not in other fields where women are viewed as the majority, such as nursing. Furthermore, men are perceived as having greater power, authority, and expertise in fields where they are viewed as being the majority (i.e., STEM), but not in fields in which they are the minority (i.e., nursing). Based on these results, we expect that in contexts in which men are perceived as gatekeepers (i.e., math), women—who are at risk of identity threat in math settings—may be affected by the type of feedback they receive from male authority figures in particular.

## Overview of Current Research

We conducted a series of experiments to see how receiving feedback in math affects women's self-evaluations and math-related outcomes. Across studies, *type* and *source* of feedback were always manipulated to be either positive or objective and came from a perceived gatekeeper (i.e., a male vs. female authority figure in math). We conceptualized the male authority figure as a gatekeeper, rather than a role model or ally, who provided feedback to participants regarding their test performance. Given that the preliminary study revealed that men are viewed as having more authority and expertise in STEM fields than women, we expected their feedback in math to be an especially influential cue in shaping women's math outcomes.

The first set of studies examined feelings of confidence and belonging (Studies 1a and 1b) in response to receiving imagined feedback on a math test from a male or female professor. Study 1b further examined perceptions of the implications of feedback, to determine whether women perceived the feedback differently when it came from a male versus female math professor. The next set of studies were conducted with male and female experimenters as authority figures who gave positive or objective feedback to participants on an actual math test. These studies examined women's math self-efficacy (Study 2), attitudes/identification/interest in math and related fields (Studies 3 and 4), and implicit identification with math versus the arts (Study 4).

To examine whether the effects of feedback from gatekeepers were specific to negatively stereotyped domains, we also conducted a study that varied the test domain, by comparing the experience of receiving feedback on a math test versus a verbal test (Study 3). Study 5 explored men's responses to receiving positive (vs. objective) math test feedback. Finally, we conducted a meta-analysis across studies to provide a more comprehensive depiction of the results.

## Study 1a

As an initial investigation of the effects of feedback, Study 1a examined feelings of confidence and belonging in response to receiving imagined feedback on a math test from a male or female authority figure. Specifically, we predicted that women who imagined receiving positive feedback (i.e., their supposed score + "Good job!" feedback) from a gatekeeper in math (i.e., a male professor) would report greater confidence and belonging than receiving objective feedback (i.e., their score only), or receiving positive feedback from a female professor.

Sample sizes for the current studies were based on previous research examining the effects of exposure to male versus female math role models (Marx & Roman, 2002) and the effects of feedback on minority students' motivation and career identification (Cohen et al., 1999). These studies typically included 15 to 30 participants per cell, so we sought to recruit a similar number by posting open enrollment for a predetermined number of days and then analyzing the data upon completion of each study. In some studies, we were able to recruit more than this minimum based on availability of lab space, research assistants, or when the data were collected (i.e., beginning of the semester when there are more participants available vs. at the end of the semester).

Participants were recruited from introductory psychology courses, which consist of mostly freshmen and sophomore students. Across the studies that collected data on participants' major (Studies 1b, 2, 3, and 4), there were a total of 19.5% STEM majors, 17.6% business majors, 14.5% health science and service majors, 12.2% social science majors, 8.3% undecided, and 3.0% arts/humanities majors. Classification of majors was based on the 2013 National Center for Education Statistics. In terms of preexisting math ability, 36.2% of participants across studies reported a math SAT score between 601 and 700, which reflected the median response; 30.8% scored between 501 and 600; 12.2% scored between 701 and 800; and 5.5% scored below 500.

### *Participants and Procedure*

Seventy-nine female undergraduates ($M = 19.13$, $SD = 2.35$) from the introductory psychology subject pool participated in the study for partial course credit. The sample was 61% White, 16% Black, 13% Asian, 10% Other ethnicities. The study was a 2 (professor sex: male vs. female) × 2 (feedback type: score + "Good Job!" vs. score only) between-subjects design.

Participants came to the lab in groups of up to six and were seated at private computer cubicles. They were randomly assigned to read a scenario in which they envisioned taking an upper level math course at the university and received objective or positive feedback from a male or female professor on a math test:

> Imagine that you took a difficult math test in class last week. Today, your professor announces that he (she) graded the tests and will be handing them back at the end of class. When he (she) hands back your test, you see that you got a score of +9/12 (+9/12 with the comment, "Good job!")

To measure feelings of confidence, participants completed the state version of the Authentic Pride Scale (Tracy & Robins, 2007), which includes items such as, "Right now, I feel . . . 'confident,' 'accomplished,' 'successful'" (seven items, $\alpha = .93$) on a scale from 1 (*not at all*) to 5 (*extremely*). Using the same response scale, participants reported their current feelings of belonging: "Right now, I feel accepted" and "Right now, I feel like I belong" (two items, $r = .66$, $p < .001$). They then completed demographics and were debriefed and dismissed.

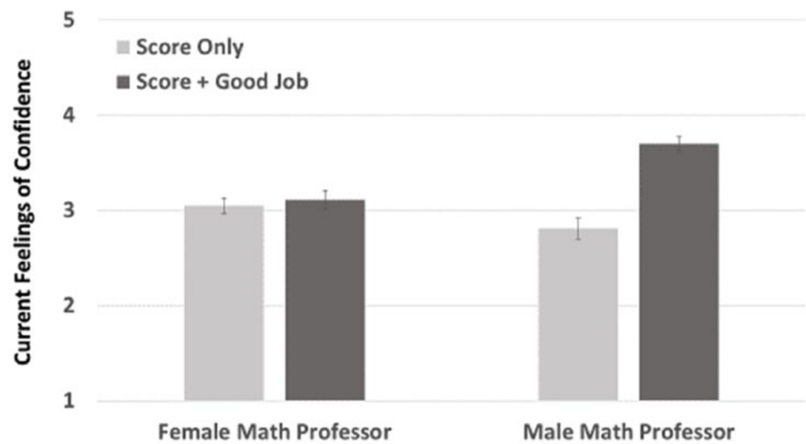


**Figure 1.** Study 1a. Mean scores for current feelings of confidence as a function of professor's sex and feedback type.
*Note.* Error bars in all figures represent ±1 *SE* of the mean.

## Results and Discussion

*Current feelings of confidence.* Results of a 2 × 2 ANOVA predicting current feelings of confidence revealed a significant main effect of feedback type, $F(1, 75) = 6.61$, $p = .01$, $\eta_p^2 = .08$, qualified by a significant professor sex × feedback type interaction, $F(1, 75) = 5.01$, $p = .03$, $\eta_p^2 = .06$ (see Figure 1). Women who imagined receiving feedback from a male professor experienced a significant boost in confidence when the feedback was positive ($M = 3.70$, $SD = 0.66$) versus objective ($M = 2.81$, $SD = 0.96$), $F(1, 75) = 10.15$, $p = .002$, $\eta_p^2 = .12$, 95% confidence interval (CI) = [0.33, 1.44]. Women who imagined receiving positive feedback also felt more confident when they received this feedback from a male versus female professor ($M = 3.11$, $SD = 0.88$), $F(1, 75) = 5.14$, $p = .03$, $\eta_p^2 = .06$, 95% CI = [0.07, 1.11]. No other simple effects were significant, all *p*s > .37.

*Current feelings of belonging.* A 2 × 2 ANOVA predicting current feelings of belonging revealed a marginally significant professor sex × feedback type interaction, $F(1, 75) = 3.34$, $p = .07$, $\eta_p^2 = .04$ (see Figure 2); no other effects approached significance. Women who imagined receiving feedback from a male professor tended to report greater feelings of belonging when they received positive ($M = 3.64$ $SD = 0.94$) versus objective ($M = 3.13$, $SD = 1.18$) feedback, although this difference did not reach significance $F(1, 75) = 2.33$, $p = .13$, $\eta_p^2 = .03$, 95% CI = [−1.18, 0.16]. Women who imagined receiving positive feedback also felt marginally greater belonging when they received this feedback from a male versus female professor ($M = 3.10$, $SD = 0.94$), $F(1, 75) = 2.99$, $p = .09$, $\eta_p^2 = .04$, 95% CI = [−1.17, 0.08]. No other simple effects approached significance, all *p*s > .31.

Consistent with predictions, Study 1a revealed that women who imagined receiving feedback from a male professor felt more confident when they received positive feedback (i.e., score + "Good job!" feedback) versus objective feedback (i.e., score only). In addition, women who received positive feedback felt more confident when the feedback came from a male versus female professor. The pattern of results for belonging was not as robust, but was still in the expected direction: Women who received positive feedback on a math test tended to feel greater belonging when this feedback came from a male versus female professor. Although in the predicted direction, we did not find strong evidence that women reported greater feelings of belonging when they received positive versus objective feedback on the math test from a male professor. A potential reason for the lack of findings is that we asked participants about their feelings of belonging in general, and not related specifically to belonging in math.

Overall, these initial findings suggest that, for women, receiving positive feedback in math seems to matter most when it comes from a perceived gatekeeper in that context. But why? Study 1b sought to address this question by examining how women interpret the meaning of the feedback they receive from gatekeepers in math and their subsequent feelings of belonging in math, in particular.

## Study 1b

Study 1b examined three possibilities regarding how women might perceive the implications of the feedback they receive on a math test. Receiving "Good Job!" feedback might convey to women that (a) they have the ability to succeed in math, (b) they belong in the field of math, and/or (c) challenges the stereotype that gender differences exist in math ability. We further examined whether women would report greater feelings of belonging *in math* when they received positive (vs. objective) feedback on a math test from a male

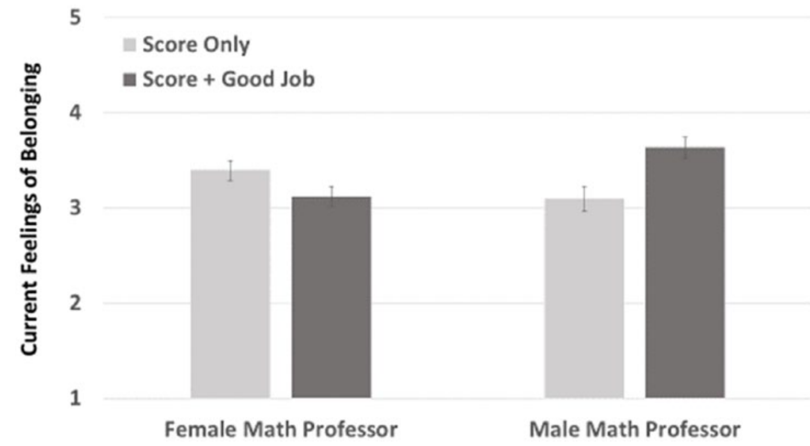


**Figure 2.** Study 1a. Mean scores for current feelings of belonging as a function of professor's sex and feedback type.

professor, given our earlier findings indicating that men are perceived to be the gatekeepers in math.

An alternative explanation for the predicted findings is that participants might respond to the feedback differently because male and female professors are perceived differently. For example, given the gender stereotype that women are kind and nurturing, receiving positive feedback from female professors may simply be viewed in light of the stereotype that women are supposed to be caring. To explore this possibility, participants reported their perceptions of the professor's competence, caring, and trustworthiness. Finally, to enhance participants' investment in the scenario, we added more information to the instructions in the current study to make the situation more vivid and engaging.

## *Participants and Procedure*

Two hundred four female undergraduates ($M$ = 19.14, $SD$ = 1.74) from the introductory psychology subject pool participated in the study for partial course credit. The sample was 46% White, 38% Asian, 8% Black, 8% Other ethnicities and was a 2 (professor sex: male vs. female) × 2 (feedback type: score + "Good Job!" vs. score only) between-subjects design. Participants came to the lab in groups of up to six and were seated at private cubicles. As in Study 1a, they were randomly assigned to read a scenario in which they imagined taking an upper level college math course and received objective or positive feedback from a male or female professor on a math test:

> Imagine for a moment that you are taking an upper-level math course at UB. It is important that you do well in this course, given that your performance in this class will ultimately affect your overall GPA. Last week, you took an exam in this math course; the test was challenging, and you are eager to find out how you did. In class today, your math professor announces that (he/ she) graded your tests and will be handing them back to you at the end of class. When (he/she) comes around and hands back your exam, you see that (he/she) has (written +9/12; OR, +9/12 with the comment "Good job!" written) at the top of the exam.

After reading the scenario, participants reported their perceptions of the implications of the feedback in terms of how much it conveyed their ability to succeed in math, their belonging in math, and the degree to which the feedback challenged the gender stereotype that women and men differ in their math abilities. Specifically, they were asked: "How much do you think the test feedback that you received from the professor . . . " (a) "Conveys that you could succeed in a math-related career," (b) "Indicates that you belong in the field of math," and (c) "Conveys that men and women are equally capable of doing well in math" on scales from 1 (*not at all*) to 7 (*very much*).

Next, participants completed a Source Credibility Scale (McCroskey & Teven, 1999) that was adapted to assess perceptions of the professor's *competence* (e.g., "Intelligent vs. Unintelligent," six items, α = .87), *caring* (e.g., "Has my interests at heart vs. Doesn't have my interests at heart," six items, α = .82), and *trustworthiness* (e.g., "Untrustworthy vs. Trustworthy," six items, α = .88). Items were presented as bipolar traits on scales anchored from 1 to 7 (e.g., 1 = *unintelligent* to 7 = *intelligent*) and relevant items were later reverse scored so that higher numbers indicate more favorable perceptions. Next, participants reported their current feelings of belonging in math by completing the phrase, "Right now, I feel . . . " "accepted by others in the field of math" and "connected with others in the field of math ($r$ = .78, $p$ < .001) on a scale from 1 (*not at all*) to 5 (*extremely*). They then reported demographic information and were debriefed and dismissed.

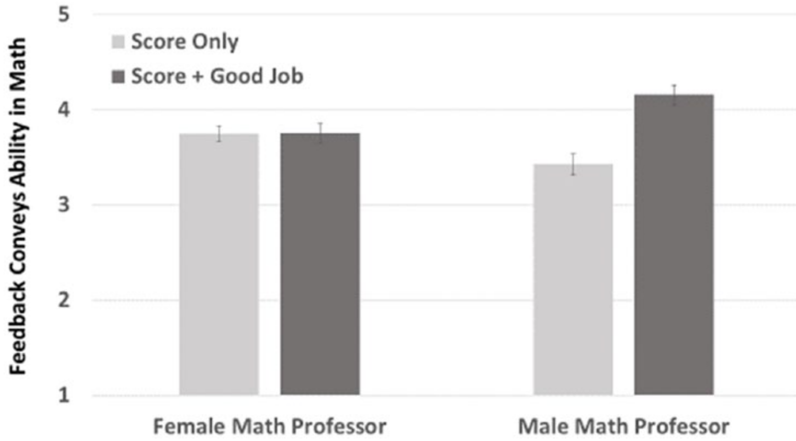


**Figure 3.** Study 1b. Mean scores of perceptions of the implications of feedback conveying ability in math as a function of professor's sex and feedback type.

## Results and Discussion

*Perceptions of the implications of feedback.* We first conducted 2 × 2 ANOVAs to test whether women perceived the implications of the feedback differently as a function of the professor's gender and type of feedback.

*Feedback conveys ability in math.* There was a marginal main effect of feedback, $F(1, 200) = 3.44$, $p = .06$, $\eta_p^2 = .02$, qualified by a professor sex × feedback type interaction, $F(1, 200) = 3.52$, $p = .07$, $\eta_p^2 = .02$ (see Figure 3). Women who imagined receiving feedback on a math test from a male professor felt that the feedback conveyed that they had ability in math when they received positive ($M = 4.16$, $SD = 1.44$) versus objective feedback ($M = 3.43$, $SD = 1.43$), $F(1, 200) = 6.94$, $p = .009$, $\eta_p^2 = .03$, 95% CI = [0.18, 1.23]. No other simple effects were significant, all *p*s > .15.

*Feedback conveys belonging in math.* Results revealed a marginal main effect of feedback, $F(1, 200) = 3.10$, $p = .08$, $\eta_p^2 = .02$, qualified by a professor sex × feedback type interaction, $F(1, 200) = 2.87$, $p = .09$, $\eta_p^2 = .01$ (see Figure 4). Women who imagined receiving feedback on a math test from a male professor felt that the feedback conveyed greater belonging in math when they received positive ($M = 3.84$, $SD = 1.35$) versus objective feedback ($M = 3.16$, $SD = 1.33$), $F(1, 200) = 6.20$, $p = .014$, $\eta_p^2 = .03$, 95% CI = [0.14, 1.22]. The only other finding that approached significance is that women who received their score only tended to think this feedback conveyed less belonging in math when it came from a male versus female professor ($M = 3.67$, $SD = 1.53$), $F(1, 200) = 3.14$, $p = .08$, $\eta_p^2 = .02$, 95% CI = [−1.06, 0.06]. No other effects approached significance, all *p*s > .55.

*Feedback challenges the gender math stereotype.* There were no significant main effects of professor sex, feedback type, or their interaction in predicting the perception that men and women are equally capable of doing well in math, all *p*s > .15.

*Current feelings of belonging in math.* For current feelings of belonging in math, there was a significant main effect of feedback type, $F(1, 200) = 3.83$, $p = .05$, $\eta_p^2 = .02$, qualified by a marginal professor sex × feedback type interaction, $F(1, 200) = 2.85$, $p = .09$, $\eta_p^2 = .01$ (see Figure 5). Women who imagined receiving feedback on a math test from a male professor reported greater feelings of belonging in math when the feedback was positive ($M = 3.01$, $SD = 0.96$) versus objective ($M = 2.51$, $SD = 1.04$), $F(1, 200) = 6.90$, $p = .009$, $\eta_p^2 = .04$, 95% CI = [0.12, 1.87]. Women who imagined receiving positive feedback on a math test also felt greater belonging in math when the feedback came from a male versus female professor ($M = 2.62$, $SD = 1.00$), $F(1, 200) = 4.24$, $p = .04$, $\eta_p^2 = .02$, 95% CI = [0.02, 0.76]. No other simple effects were significant, all *p*s > .71.

*Perceptions of professor*

*Competence.* There were no significant main effects of professor sex, feedback type, or their interaction in predicting perceptions of the professor's competence, all *p*s > .39.

*Caring.* There was a significant main effect of feedback type, $F(1, 200) = 38.71$, $p < .001$, $\eta_p^2 = .16$, 95% CI = [0.52, 1.01], such that participants rated professors who gave them positive feedback as being more caring ($M = 5.25$, $SD = 0.87$) than those who received objective feedback ($M = 4.48$, $SD = 0.88$). No other effects or interaction were significant, all *p*s > .40.

*Trustworthiness.* There was a significant main effect of feedback type, $F(1, 200) = 5.92$, $p = .02$, $\eta_p^2 = .03$, 95% CI = [0.06, 0.59], such that participants rated professors who gave

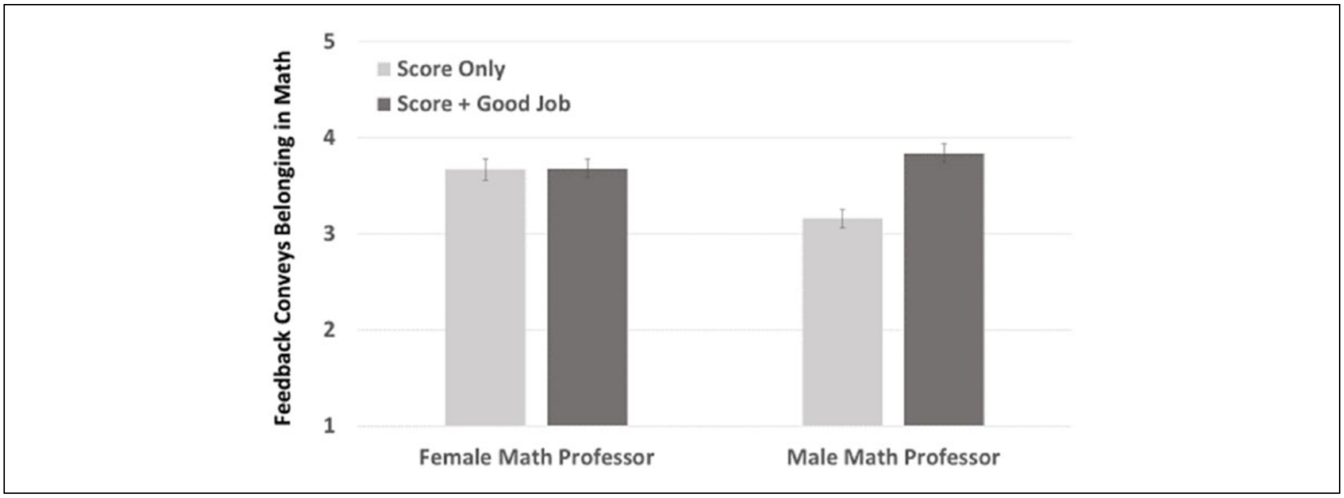


**Figure 4.** Study 1b. Mean scores of perceptions of the implications of feedback conveying belonging in math as a function of professor's sex and feedback type.


**Figure 5.** Study 1b. Mean scores for current feelings of belonging in math as a function of professor's sex and feedback type.

them positive feedback as more trustworthy ($M = 5.29$, $SD = 0.92$) than those who received objective feedback ($M = 4.97$, $SD = 0.96$). No other effects or interaction were significant, all $p$s > .79.

Overall, Study 1b suggests that women's responses to feedback depend on the source and type of feedback given. The pattern of results showed that women who imagined receiving "Good job!" feedback (vs. receiving their score only) on a math test from a male (vs. female) professor perceived this feedback as conveying that they had the ability to succeed in math and belonged in math. However, receiving positive feedback did not challenge gender stereotypes regarding men's and women's ability in math. Such findings are consistent with previous work, which also found that women showed greater implicit identification with math in response to situational cues, even while their gender stereotypes about STEM remained intact (Stout et al., 2011).

It is also plausible that women who received positive feedback on the math test subtyped themselves as disconfirming members of the broader stereotype that women are bad at math. According to theories of subtyping, stereotypes are unlikely to change if perceivers group together disconfirming members into a subtype and view them as an exception, rather than representative of the group as a whole (Richards & Hewstone, 2001; Weber & Crocker, 1983). Along these lines, women who received positive math test feedback may have subtyped their performance in this domain as reflective of their personal ability, rather than perceiving this feedback to be typical of their superordinate group (i.e., women in general). In this way, gender stereotypes about STEM persisted despite receiving positive feedback on a math test.

Women also reported greater belonging in math when they received positive feedback on a math test from a male professor, whereas this was not the case for those who received feedback from a female professor. Such findings are consistent with the idea that feedback may be especially influential when it comes from perceived gatekeepers in a field. Professors were rated as being more caring and trustworthy when they gave positive versus objective feedback, regardless of the professor's gender. Such findings weaken the alternative possibility that women merely responded differently to feedback from male (vs. female) professors because of descriptive or prescriptive gender stereotypes about women being caring or trustworthy. Finally, perceptions of competence were unaffected by the professor's gender or type of feedback; this could be due to the fact that perceptions of general competence were assessed in the current study, whereas the preliminary study found that males were viewed as having greater authority, expertise, and credibility in STEM fields, in particular.

## Study 2

A limitation of the studies thus far is that they involved hypothetical scenarios, which may not fully capture the experience of receiving actual feedback on a test from an authority figure. Accordingly, Study 2 examined the effects of receiving feedback on a math test from a male or female authority figure in the lab on women's self-efficacy in math. We predicted that women who received positive feedback on a math test from a male authority figure (i.e., experimenter) would report greater math self-efficacy than receiving their score only, or receiving positive feedback from a female experimenter.

### *Participants and Procedure*

One hundred twenty-one undergraduate women from the introductory psychology subject pool participated in the study for partial course credit. Participants were excluded from analyses if they failed the manipulation check (i.e., incorrectly recalled the type of feedback they received; $n$ = 2) or reported suspicion about the test results ($n$ = 11). The final sample consisted of 108 women ($M_{age}$ = 18.75 years, $SD$ = 1.13 years) that was 57% White, 25% Asian, 7% Black, and 11% Other ethnicities.[1] The study was a 2 (experimenter sex: female vs. male) × 2 (feedback type: score + "Good Job!" vs. score only) between-subjects design. All studies involved multiple male and female experimenters.

Participants came to the lab in groups of up to six and completed the study in private cubicles. Participants were given a scantron and a test booklet that contained 12 challenging items that were taken from a quantitative section of the Graduate Records Exam (GRE); they were given 12 min to complete the exam and record their answers on a scantron form. To heighten the evaluative nature of the situation, participants were told that performance on the test predicted academic and career success and that students who score high on the test have higher college GPAs and are likely to be accepted to top-ranked graduate and professional programs (see Methodology File in the online appendix for instructions and question items).

Afterward, the experimenter collected the exams and went to an adjacent room to "grade" them. For each participant, the experimenter used a red pen to mark three preselected items as incorrect on their scantron form. To manipulate feedback type, the experimenter provided written feedback on a small slip of paper that was attached to their exam with a paper clip. Depending on experimental condition, the experimenter wrote (in red ink) at the top of the slip of paper either: "+9/12" or "+9/12" with the comment "Good Job!" handwritten on the slip. The experimenter then handed back the test materials and feedback slip for participants to look over.

Next, participants completed a measure of math self-efficacy (adapted from Stout et al., 2011) that asked, "How confident do you feel about your math performance?" on a scale from 1 (*not at all*) to 7 (*very*); "If you were to take an introductory level college math class right now, how well would you do in that class?" and "If you were to take an intermediate level college math class right now, how well would you do in that class?" with response scales corresponding to letter grades (C, C+, B−, B, B+, A−, or A). The letter grades were later converted into numerical values ranging from 1 to 7, so that a common metric could be created across items, which were standardized and averaged to create a math self-efficacy score (α = .83).

Participants then completed demographic information and responded to the following question: "Was there anything odd or suspicious about today's experiment? If yes, what?" Responses that expressed doubt or disbelief about their test scores or feedback were coded as suspicious. Participants also responded to the question, "What type of feedback did you receive from the experimenter about your performance on the test you took earlier in the study?" with the following options: (a) The experimenter wrote my score and that is it, (b) the experimenter wrote my score and said that I did a good job, or (c) other. Participants were excluded from analyses if they failed the manipulation check (i.e., incorrectly recalled the type of feedback they received; $n$ = 2) or reported suspicion about the test results ($n$ = 11) (see Note 1).

### *Results and Discussion*

A 2 (experimenter sex) × 2 (feedback type) ANOVA showed no significant main effects, but did reveal the expected experimenter sex × feedback type interaction, $F(1, 104) = 5.42$, $p = .02$, $\eta_p^2 = .05$ (see Figure 6). Women who received feedback from a male experimenter reported higher math self-efficacy when they received positive ($M$ = 0.24, $SD$ = 0.77) versus objective feedback ($M$ = −0.26, $SD$ = 0.75), $F(1, 104) = 5.10$, $p = .03$, $\eta_p^2 = .05$, 95% CI = [0.06, 0.95]. Women who received positive feedback also tended to report higher

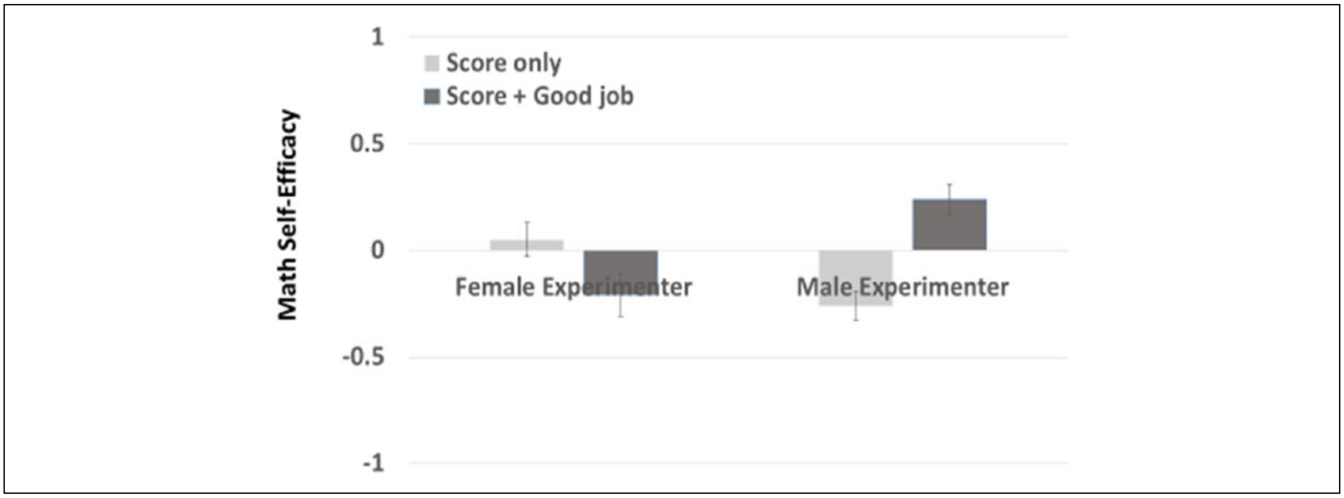


**Figure 6.** Study 2. Mean scores for math self-efficacy (standardized scores) as a function of experimenter's sex and feedback type.

math self-efficacy when they received this feedback from a male ($M$ = 0.24, $SD$ = 0.77) versus female experimenter ($M$ = −0.21, $SD$ = 1.04), $F(1, 104) = 3.63$, $p = .06$, $\eta_p^2 = .03$, 95% CI = [−0.02, 0.92]. No other effects approached significance, all $ps > .17$.[2]

Consistent with predictions, Study 2 revealed that receiving feedback boosted women's math self-efficacy when it was positive (i.e., their supposed score + "Good job!" feedback) versus objective in nature (i.e., their score only) and came from a male authority figure. Furthermore, women reported greater self-efficacy in math when they received feedback from a male (but not female) experimenter on a math test. These findings parallel the results of Studies 1a and 1b and suggest that the benefits of receiving positive feedback from perceived gatekeepers occur in both imagined and in real-life settings.

## Study 3

The findings thus far suggest that women benefit from receiving positive (vs. objective) feedback from male authority figures in particular. A limitation of these studies, however, is that they focused only on feedback in one domain: math. We focused on math because feedback should be most impactful when it comes from gatekeepers in a field in which one is susceptible to social identity threat, such as women in math settings. Given that women are viewed as having poor quantitative abilities, feedback from gatekeepers may be more influential when it is given in math versus a domain in which women are not negatively stereotyped, such as verbal ability.

To test this idea, Study 3 examined whether the domain of feedback mattered in shaping women's STEM-related attitudes/identification/interest. We expected to find a three-way interaction between feedback type, feedback source, and test domain, such that women would report more favorable attitudes toward, identification with, and interest in STEM fields when they received positive (vs. objective) feedback from a gatekeeper in a domain in which they are negatively stereotyped (i.e., math vs. verbal domain). To provide further evidence that these particular variables affect math-related outcomes above and beyond effects of other variables, we controlled for participants' initial interest in STEM fields in the current study.

### *Participants and Procedure*

One hundred ninety-four undergraduate women from the introductory psychology subject pool participated in the study for partial course credit. Participants were excluded from analyses if they incorrectly recalled the type of feedback they received ($n$ = 13), reported suspicion about the test results ($n$ = 12), or both ($n$ = 1). The final sample consisted of 168 women ($M_{age}$ = 19.02 years, $SD$ = 1.20 years) who were 50% White, 32% Asian, 11% Black, and 7% Other ethnicities.[3] The study was a 2 (experimenter sex: female vs. male) × 2 (feedback type: score + "Good Job!" vs. score only) × 2 (test domain: math vs. verbal) between-subjects design. As part of an initial online survey, participants reported their interest in pursuing a degree or career in math and science on a scale from 1 (*not at all*) to 7 (*very much; M* = 4.72, *SD* = 2.24).[4] This item was included as a covariate in the analyses to control for baseline differences in STEM interest.

A few weeks later, participants came to the lab and were randomly assigned to complete a math or verbal test using the same materials, instructions, and feedback manipulations as in Study 2.[5] After receiving feedback, participants reported their attitudes toward, identification with, and interest in STEM (e.g., "How much do you like Math and Science—e.g., Math, Chemistry, Physics, Computer Science, Engineering, Technology, etc."; "How much do you identify with Math and Science?"; "How interested are you in Math

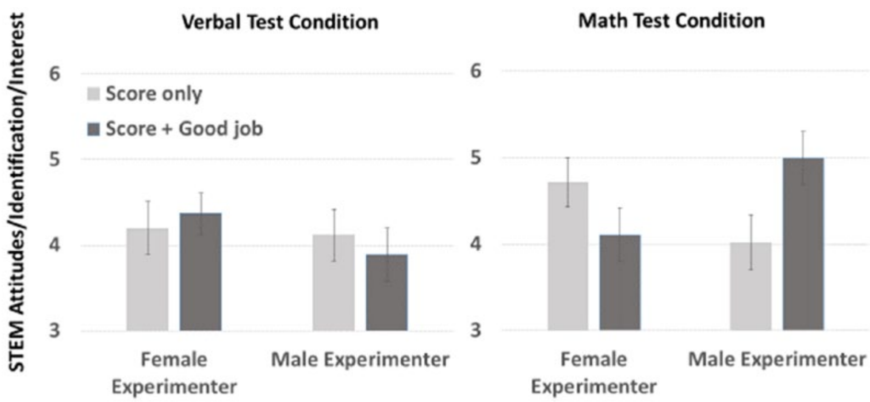


**Figure 7.** Study 3. Mean scores for attitudes/identification/interest in STEM as a function of experimenter's sex, feedback type, and test type, adjusted for initial interest in STEM.
*Note.* STEM = science, technology, engineering, and mathematics.

and Science?") on scales from 1 (*not at all*) to 7 (*very much*; six items, α = .96; adapted from Park, Young, Troisi, & Pinkus, 2011). Past research examined attitudes toward STEM, identification with STEM (e.g., Dasgupta, 2011; Nosek et al., 2002; Stout et al., 2011), and interest in STEM as outcomes (e.g., Cheryan et al., 2013; Davies et al., 2002). Although often examined as separate variables, for brevity's sake, we present the results as a composite. Indeed, a principal components factor analysis on all the attitudes, identification with, and interest in STEM items in the current study revealed a single factor with an eigenvalue of 4.99 that explained 83.25% of the variance.

Using the same response scale, participants reported how they felt about the fields of arts and humanities (e.g., "How much do you like the Arts and Humanities—e.g., Visual Arts, Performing Arts, English, Literature, Philosophy, Foreign Languages?" "How much do you identify with the Arts and Humanities?" "How interested are you in the Arts and Humanities?" six items, α = .93). Participants then completed the same demographic items, suspicion, and manipulation check items as in Study 2, and were then debriefed and dismissed.

## Results and Discussion

*STEM attitudes/identification/interest.* A 2 (experimenter sex) × 2 (feedback type) × 2 (test type) ANCOVA controlling for initial interest in STEM revealed a significant main effect of the covariate, $F(1, 153) = 125.16$, $p < .001$, $\eta_p^2 = .45$, qualified by a significant experimenter sex × feedback type × test domain interaction, $F(1, 153) = 5.25$, $p = .02$, $\eta_p^2 = .03$ (see Figure 7). No other main effects or interactions were significant, all *p*s > .17.[6]

As predicted, among women who took a math test, those who received positive feedback from a male experimenter reported more favorable attitudes/identification/interest in STEM ($M_{adj} = 5.01$, $SE = 0.30$) compared with when they received their score only from a male experimenter on a math test ($M_{adj} = 4.03$, $SE = 0.32$), $F(1, 153) = 4.96$, $p = .03$, $\eta_p^2 = .03$, 95% CI = [0.11, 1.86], or when they received positive feedback on a math test from a female experimenter ($M_{adj} = 4.17$, $SE = 0.30$), $F(1, 153) = 3.80$, $p = .053$, $\eta_p^2 = .02$, 95% CI = [−0.01, 1.69]. Furthermore, women who received positive feedback from a male experimenter reported more favorable attitudes/identification/interest in STEM when they received feedback on a math versus verbal test ($M_{adj} = 3.90$, $SE = 0.31$), $F(1, 153) = 6.59$, $p = .01$, $\eta_p^2 = .04$, 95% CI = [0.26, 1.97]. No other simple effects were significant, all *p*s > .10.

*Arts attitudes/identification/interest.* Results of an ANCOVA revealed only a significant main effect of initial interest in STEM, $F(1, 153) = 21.65$, $p < .001$, $\eta_p^2 = .12$; no other main effects or interactions were significant, all *p*s > .37.

Study 3 established a key boundary condition of the effects of feedback. Specifically, the benefits of receiving positive (vs. objective) feedback from a gatekeeper (i.e., a male vs. female authority figure) were only associated with domains in which women are negatively stereotyped (i.e., math ability). In domains in which women are not negatively stereotyped (i.e., verbal ability), the type of feedback they received did not matter. Indeed, receiving positive (vs. objective) feedback on a verbal test did not affect women's attitudes/identification/interest in STEM or in the arts/humanities. Thus, women do not appear to be responding indiscriminately to positive feedback from male authority figures. Rather, the results of the studies thus far suggest that they are especially attuned to the meaning of the feedback they receive, the domain in which they receive feedback, and then shift their attitudes/identification/interest accordingly.

## Study 4

Study 4 sought to replicate and extend the findings of the previous studies by examining whether the effects of receiving positive (vs. objective) feedback from a male (vs. female) authority figure influence not only women's explicit, self-reported feelings toward STEM, but also their implicit, automatic associations between themselves and math. Previous research suggests that women, on average, show less implicit identification with math and science than men (Nosek et al., 2002; Park et al., 2002). If receiving positive feedback from a gatekeeper changes how women think about themselves in relation to negatively stereotyped domains, then receiving positive feedback from male authority figures in math contexts might even change women's implicit cognitive associations between themselves and math.

People's implicit self-concept is thought to be malleable to situational cues; specific self-traits become more cognitively accessible than other traits in response to situational manipulations. Supporting this idea, Markus and Kunda (1986) found that cues in the environment led to shifts in people's working self-concept and self-beliefs, particularly when these responses were measured indirectly. In terms of changes to implicit math self-concept, research has demonstrated that women who were initially low in math identification showed increased implicit identification with this domain when they were trained to approach math using an approach-avoid joystick task (Kawakami, Steele, Sifa, Phills, & Dovidio, 2008). It therefore seems plausible that receiving feedback from a gatekeeper in math could alter the way women think about themselves in relation to math, and implicit measures could be a useful way to capture subtle changes that might otherwise be difficult to detect or report (Greenwald & Banaji, 1995).

### Participants and Procedure

One hundred twelve undergraduate women from the introductory psychology subject pool participated in the study for partial course credit. Participants were excluded from analyses if they incorrectly recalled the type of feedback they received ($n$ = 6) or were suspicious of the test results ($n$ = 12). The final sample consisted of 94 women ($M_{\text{age}}$ = 19.47 years, $SD$ = 2.91 years) who were 43% White, 38% Asian, 10% Black, and 9% Other ethnicities.[7]

The procedure of this study was identical to Study 3, except all participants in the current study completed the same math test as in Study 2 and an Implicit Association Test (IAT). Specifically, after receiving positive or objective feedback on a math test, participants completed an IAT to assess their implicit identification with math. Participants were instructed to pair words pertaining to the category "Mathematics" (e.g., calculus, equations) versus "The Arts" (e.g., poetry, literature) with "I" words (e.g., me, myself) versus "They" words (e.g., they, them; Nosek et al., 2002).[8]

Scores on the IAT were calculated using the D score algorithm recommended by Greenwald, Nosek, and Banaji (2003). The mean response times to arts words (where self was paired with arts words) were subtracted from the mean response times to math words (where self was paired with math words). This difference score was then divided by the standard deviation of all correct response time in the trials that were paired with the self to adjust for the effect of response variability on scores. Lower scores reflect quicker reaction times to associate the self with math words relative to arts words. Three participants' data were excluded from analyses because more than 10% of their trials exceeded the recommended latency of less than 300 ms (Greenwald et al., 2003). Afterward, all participants responded to the same items as in Study 3 to assess their attitudes/identification/interest in STEM ($\alpha$ = .96) and in the arts/humanities ($\alpha$ = .94). They then completed demographic information and the same suspicion and manipulation check items as before.

### Results and Discussion

*Implicit math identification.* A 2 (experimenter sex) × 2 (feedback type) ANOVA showed no significant main effects, but did reveal the expected experimenter sex × feedback type interaction, $F(1, 87) = 4.78$, $p = .03$, $\eta_p^2 = .05$ (see Figure 8). Women who received feedback on a math test from a male experimenter were quicker to associate themselves with math (vs. arts) when they received positive ($M$ = 0.28, $SD$ = 0.44) versus objective feedback ($M$ = 0.61, $SD$ = 0.51), $F(1, 87) = 4.48$, $p = .04$, $\eta_p^2 = .05$, 95% CI = [−0.63, −0.02]. Women also tended to more quickly associate themselves with math (vs. arts) when they received positive feedback on a math test from a male versus female experimenter ($M$ = 0.56, $SD$ = 0.56), $F(1, 87) = 3.44$, $p = .07$, $\eta_p^2 = .04$, 95% CI = [−0.56, 0.02]. No other simple effects were significant, all $p$s > .21.

*STEM attitudes/identification/interest.* A 2 (experimenter sex) × 2 (feedback type) ANOVA revealed a marginally significant main effect of experimenter sex, $F(1, 90) = 3.72$, $p = .06$, $\eta_p^2 = .04$, qualified by a significant experimenter sex × feedback type interaction, $F(1, 90) = 3.99$, $p = .049$, $\eta_p^2 = .04$ (see Figure 9). Replicating Study 3, women who received feedback on a math test from a male experimenter reported more favorable attitudes/identification/interest in STEM when they received positive ($M$ = 4.94, $SD$ = 1.60) versus objective feedback ($M$ = 3.89, $SD$ = 1.67), $F(1, 90) = 4.14$, $p = .045$, $\eta_p^2 = .04$, 95% CI = [0.02, 2.08], and compared with when they received positive feedback on a math test from a female experimenter ($M$ = 3.55, $SD$ = 1.83), $F(1, 90) = 7.63$, $p = .007$, $\eta_p^2 = .08$, 95% CI = [0.39, 2.39]. No other simple effects were significant, all $p$s > .45 (see Tables S3 and S4 in Supplemental Materials for results of analyses in Studies 3 and 4 examining effects of the experimental manipulations


**Figure 8.** Study 4. Mean reaction time scores on Math Identification Implicit Association Test as a function of experimenter's sex and feedback type.
*Note.* Lower values reflect quicker average reaction times to associate "Me" words with "Mathematics" relative to "The Arts."


**Figure 9.** Study 4. Mean scores for attitudes/identification/interest in STEM as a function of experimenter's sex and feedback type.
*Note.* STEM = science, technology, engineering, and mathematics.

on STEM attitudes, identification, and interest as separate factors).

*Arts attitudes/identification/interest.* Results of a two-way ANOVA showed no significant main effects or interaction in predicting attitudes/identification/interest in the arts/humanities, all $p$s > .24.

In sum, women who received positive feedback from a male experimenter on a math test not only reported more favorable explicit attitudes/identification/interest in STEM, consistent with Study 3, but also showed stronger implicit associations between themselves and math relative to the arts. Although disassociations have previously been documented between implicit and explicit measures (Bosson, Swann, &

Pennebaker, 2000; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald & Banaji, 1995), implicit and explicit measures have been shown to be interrelated, as well (Cunningham, Nezlek, & Banaji, 2004; Cunningham, Preacher, & Banaji, 2001). For example, more positive implicit math attitudes and identity are associated with more positive explicit math attitudes and identity (Nosek et al., 2002).

In addition, Stout and colleagues (2011) found that women who were exposed to ingroup STEM role models showed greater subjective identification with these role models, which in turn, increased their implicit identification with STEM, self-reported self-efficacy, and commitment to pursue future career goals in STEM. Given that we were not testing the effects of role models in the current research, we did not expect feedback from a female authority figure to shift women's implicit math identification or explicit self-reports related to math. Rather, women who received positive feedback on a math test from a male authority figure—a gatekeeper in the field—showed greater implicit identification with math and reported more favorable explicit attitudes/identification/interest in STEM.

Thus, whereas previous research found that women showed less identification with math and science relative to arts and languages on both explicit and implicit measures (Nosek et al., 2002; Nosek & Smyth, 2011), the current study qualifies such findings by demonstrating that implicit associations between the self and math can be shifted when women receive positive (vs. objective) feedback from a male authority figure in an area in which they are negatively stereotyped, such as math.

## Study 5

In all the studies so far, we focused on women's responses to receiving feedback on a math test, given that women are susceptible to social identity threat in math settings. To explore the generalizability of the findings, Study 5 investigated men's responses to receiving positive or objective feedback on a math test from a male or female authority figure (i.e., an experimenter). Unlike women, men are not negatively stereotyped in math[9] (Spencer et al., 1999); thus, we did not expect men to respond differently to receiving feedback on a math test as a function of whether the feedback came from a male or female experimenter. However, it seems plausible that men might benefit from receiving positive versus objective feedback, given past work showing that people exhibit greater motivation and engagement in tasks when they receive positive versus neutral feedback (Harackiewicz, 1979; Vallerand & Reid, 1998).

### Participants and Procedure

One hundred forty undergraduate men from the introductory psychology subject pool participated in the study for partial course credit. Participants were excluded from analyses if they were suspicious of the test feedback ($n$ = 30).[10] All participants correctly recalled the type of feedback they received. The final sample consisted of 110 men ($M_{age}$ = 19.62 years, $SD$ = 1.46 years) who were 57% White, 26% Asian, 12% Black, and 5% Other ethnicities.

As in the previous studies, participants took the same math test as described in Study 2 and then received their score + "Good job!" feedback, or their score only from a male or female experimenter. Afterward, they completed the following measures: (a) their *perceptions of the implications of feedback*, as in Study 1b (i.e., "How much do you think the test feedback that you received from the experimenter . . ." "conveys that you could succeed in a math-related career," "indicates that you belong in the field of math," and "conveys that men and women are equally capable of doing well in math") on scales from 1 (*not at all*) to 7 (*very much*); (b) their current feelings of *confidence* using the same scale as in Study 1a (e.g., "Right now I feel . . . " "confident," "accomplished," seven items, $\alpha$ = .95; Tracy & Robins, 2007) from 1 (*not at all*) to 5 (*extremely*); (c) their current feelings of *belonging in math* using the same items as in Study 1b, in which participants completed the phrase, "Right now, I feel . . ." "accepted by others in the field of math" and "connected with others in the field of math" ($r$ = .86, $p$ < .001) from 1 (*not at all*) to 5 (*extremely*); and (d) their *math self-efficacy* using the same items as in Study 2, which were standardized and averaged together (three items, $\alpha$ = .86; adapted from Stout et al., 2011). They then completed demographics and the same suspicion and manipulation check items as in the previous studies.

### Results and Discussion

*Perceptions of the implications of feedback.* For all analyses, we conducted 2 × 2 ANOVAs, in which experimenter sex and feedback type were the independent variables (IVs).

*Feedback conveys ability in math.* Results revealed only a significant main effect of feedback condition, $F(1, 106)$ = 7.13, $p$ = .009, $\eta_p^2$ = .06, 95% CI = [0.23, 1.58]. Men who received positive feedback on a math test perceived this feedback to convey that they had greater ability in math ($M$ = 3.83, $SD$ = 1.95) than men who received objective feedback ($M$ = 2.90, $SD$ = 1.63); no other effects or interaction were significant, all $p$s > .10.

*Feedback conveys belonging in math.* There were no significant main effects or interaction predicting perceptions of the feedback challenging the gender stereotype about math, all $p$s > .16.

*Feedback challenges the gender math stereotype.* There were no significant main effects or interaction predicting perceptions of the feedback challenging the gender stereotype about math, all $p$s > .14.

*Current feelings of confidence.* For feelings of confidence, there was only a significant main effect of feedback type, $F(1, 106) = 7.19$, $p = .008$, $\eta_p^2 = .06$, 95% CI = [0.13, 0.88]. Men felt more confident when the feedback they received was positive ($M = 3.39$, $SD = 0.91$) versus objective ($M = 2.87$, $SD = 1.03$).

*Current feelings of belonging in math.* For belonging in math, there was only a significant main effect of feedback type, $F(1, 106) = 4.92$, $p = .03$, $\eta_p^2 = .04$, 95% CI = [0.05, 0.89]. Men reported greater belonging in math when they received positive ($M = 2.65$, $SD = 1.11$) versus objective ($M = 2.19$, $SD = 1.08$) feedback.

*Math self-efficacy.* Results showed a marginally significant main effect of feedback type, $F(1, 106) = 3.38$, $p = .07$, $\eta_p^2 = .03$, 95% CI = [−0.02, 0.64]; no other findings approached significance. Men tended to report greater math self-efficacy when they received positive ($M = 0.18$, $SD = 0.86$) versus objective ($M = −0.13$, $SD = 0.88$) math test feedback.

Overall, Study 5 showed that men who received positive feedback on a math test perceived this feedback to convey that they had ability in math, compared with those who received objective feedback. Men also reported greater confidence, belonging in math, and marginally greater math self-efficacy when they received positive (vs. objective) feedback on a math test, regardless of whether the feedback came from a male or female authority figure.

Although women were not part of the current study, examining the findings collectively across studies suggests that women are affected by whether the feedback they receive in math comes from a gatekeeper or not. However, men who receive feedback on a math test do not appear to be affected by the source of the feedback, but by whether the feedback is positive or objective in nature. Thus, whereas women in our earlier studies distinguished between the type of feedback they received depending on whether it came from a male (vs. female) authority figure, men's responses in the current study did not differ based on who the feedback came from.

A potential explanation for these findings is that, in the current study, only feedback on a math test was examined. Members of nonstereotyped groups do not have to contend with the threat of being negatively stereotyped, so they should be less affected by cues in the environment that indicate whether members of an outgroup (vs. ingroup) are evaluating them (Wout et al., 2009). In the current study, the majority of the sample was White or Asian males who are unlikely to be negatively stereotyped in the math domain. Thus, men may be more influenced by the source of feedback when they perform in an area in which they are negatively stereotyped or feel uncertain about their abilities. For example, men often underperform women on tasks of social sensitivity (Leyens, De´sert, Croizet, & Darcis, 2000), so this could be a domain in which feedback from women (vs. men) might affect men's self-evaluations and related outcomes.

## Meta-Analysis Across Studies

The present studies suggest that women who receive feedback on a math test show better math-related outcomes when they receive positive (vs. objective) feedback from a male (vs. female) authority figure. Given that some of the interactions and simple effects were not always significant, however, we sought to provide a more comprehensive depiction of the results and a more precise estimate of the effects by conducting a meta-analysis across studies. Furthermore, our use of multimethods and meta-analytic techniques adequately address issues of power, given that our meta-analysis weights the studies based on the different sample sizes used across studies (see Braver, Thoemmes, & Rosenthal, 2014).

### *Meta-Analytic Strategy*

We examined the overall effect of the experimental manipulations on the same, or conceptually similar, dependent variables (DVs; i.e., confidence/self-efficacy in Studies 1a and 2, belonging in Studies 1a and 1b, STEM attitudes/identification/interest in Studies 3 and 4). For each analysis, we first conducted two meta-analyses across the selected studies: one for the male authority figure conditions and one for the female authority figure conditions. We calculated the effect size *d* for the difference between the objective and positive math test feedback conditions for the relevant DVs and then input these *d*s into the meta-analytic spreadsheet tool provided by Braver et al. (2014).

Next, we examined the meta-analytic effect of receiving positive versus objective feedback as a function of the gender of the authority figure. To do this, we conducted two meta-analyses: one for the positive feedback conditions and one for the objective feedback conditions. We then calculated the effect size *d* for the difference between the male and female authority figure conditions for the relevant DVs and input these *d*s into the meta-analytic spreadsheet tool provided by Braver et al. (2014). Fixed and random effects meta-analyses for all DVs produced identical results.

*Confidence/self-efficacy.* Table 2 summarizes the simple effects of type and source of feedback for each DV. For confidence/self-efficacy, we first standardized and averaged together the confidence items in Study 1a to be consistent with the way self-efficacy scores were computed in Study 2. As predicted, the meta-analysis revealed that when women received feedback on a math test from a male authority figure, they felt greater confidence/self-efficacy when they received positive feedback compared to when they received objective feedback. Furthermore, women felt greater confidence/

**Table 2.** Meta-Analysis of Simple Effects for Women Across Studies.

| | Simple effect of feedback type (objective vs. positive feedback) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Feedback from female authority figure | | | | Feedback from male authority figure | | | |
| DV | *M d* | *SE* | *z* | 95% CI | *M d* | *SE* | *z* | 95% CI |
| Confidence/self-efficacy | 0.10 | 0.21 | 0.48 | [−0.31, 0.51] | 0.83*** | 0.20 | 4.17 | [−1.21, −0.44] |
| Feelings of belonging | 0.05 | 0.17 | 0.33 | [−0.38, 0.27] | 0.50** | 0.16 | 3.06 | [−0.82, −0.18] |
| STEM attitudes/identification/interest | 0.30 | 0.22 | 1.05 | [−0.20, 0.65] | 0.69** | 0.21 | 3.20 | [−1.11, −0.27] |
| | Simple effect of authority figure (female vs. male authority figure) | | | | | | | |
| | Objective feedback condition | | | | Positive feedback condition | | | |
| DV | *M d* | *SE* | *z* | 95% CI | *M d* | *SE* | *z* | 95% CI |
| Confidence/self-efficacy | 0.34 | 0.21 | 1.62 | [−0.07, 0.76] | 0.72** | 0.22 | 3.28 | [0.29, 1.15] |
| Feelings of belonging | 0.13 | 0.17 | 0.74 | [−0.21, 0.47] | 0.45** | 0.16 | 2.77 | [−0.76, −0.13] |
| STEM attitudes/identification/interest | 0.23 | 0.22 | 1.05 | [−0.20, 0.65] | 0.73** | 0.21 | 3.39 | [−1.14, −0.31] |

*Note.* DV = dependent variable; 95% CI = 95% confidence interval; STEM = = science, technology, engineering, and mathematics.
**$p < .01$. ***$p < .001$.

self-efficacy when they received positive feedback from a male (vs. female) authority figure.

*Feelings of belonging.* Results of the meta-analysis revealed that women who received feedback from a male authority figure felt greater sense of belonging when they received positive (vs. objective) feedback on a math test (see Table 2). Women also felt greater sense of belonging when they received positive feedback from a male (vs. female) authority figure.

*STEM attitudes/identification/interest.* Women who received feedback from a male authority figure reported more favorable attitudes/identification/interest in STEM when they received positive (vs. objective) feedback on the math test (see Table 2). Women who received positive feedback on the math test also reported more favorable attitudes/identification/interest in STEM when they received this feedback from a male (vs. female) authority figure.

Overall, the findings of the meta-analyses confirm that women who received feedback (a) felt greater confidence/self-efficacy, (b) felt belonging, and (c) expressed more favorable attitudes/identification/interest in STEM fields when they received positive (vs. objective) feedback on a math test from a male (vs. female) authority figure. When women received feedback on a math test from a female authority figure, it did not matter whether the feedback was positive or objective in nature. In addition, when women received objective feedback on a math test, it did not matter whether the authority figure giving them this feedback was male or female. Thus, consistent with predictions, receiving *positive* feedback in math from a *male* authority figure, in particular, led women to feel greater confidence/self-efficacy, belonging, and to report more favorable attitudes/identification/interest in STEM.

## General Discussion

When individuals are in situations where they could be judged based on their social identity, they become vigilant for signs of whether they will be valued or devalued in that context (Steele et al., 2002). Given that negative stereotypes exist about women's quantitative abilities, women may be especially attentive to, and influenced by, cues in math environments that signal whether or not they belong and have the ability to succeed in this domain. Consistent with these ideas, women who received positive ("Good job!") feedback on a math test from a male authority figure in math reported greater feelings of confidence (Study 1a), belonging (Study 1a, Study 1b), self-efficacy (Study 2), and expressed more favorable explicit attitudes/identification/interest in STEM (Studies 3 and 4), and greater implicit identification with math (Study 4) than women who received objective feedback, or received positive feedback on a math test from a female authority figure. Importantly, women showed more favorable attitudes/identification/interest in STEM when they received positive (vs. objective) feedback from a male authority figure on a math test, but not on a verbal test (Study 3). These findings underscore the benefits of receiving positive feedback in domains in which individuals are susceptible to social identity threat and corresponding stereotypes.

Men were affected only by the type of feedback they received, not by whether the feedback came from a gatekeeper. Specifically, men who received positive feedback on a math test interpreted this feedback to convey that they had ability in math and reported greater confidence, belonging in math, and marginally greater self-efficacy in math (Study 5). Overall, the findings across studies and the meta-analyses suggest that women are highly attuned to both the type and source of feedback they receive in math settings, whereas men focus more on whether the feedback they receive in math is positive or not, irrespective of the feedback source.

### *Why Is Positive Feedback Effective?*

Women who received positive feedback in math from a male gatekeeper experienced boosts in their math-related outcomes. Why might this be? Positive feedback may have been effective by (a) conveying that the authority figure cared about the participant, (b) giving participants a lens through which to interpret their performance, (c) reducing uncertainty about one's ability and belonging in math, and/or (d) leading participants to identify more with math on an implicit level, which may have led them to feel more favorably toward math and related fields on an explicit level.

First, positive feedback may have been effective by conveying caring and support. Indeed, professors who gave positive feedback in Study 1b were rated as being more caring and trustworthy than professors who gave objective feedback. However, these effects emerged regardless of the gender of the professor. Thus, although authority figures who give positive feedback are viewed as more caring, understanding, and trustworthy, these perceptions do not explain why women benefitted from receiving positive feedback from male gatekeepers in particular.

A second possibility is that positive feedback helped participants interpret their ambiguous math performance. Although the math test was always described as being evaluative and diagnostic of future outcomes, no information was given regarding how other students performed on the test, or any other comparative information that would provide a benchmark as to how participants did on the test. Thus, receiving one's supposed score plus "Good job!" feedback may have helped participants construe the score they received in a more favorable light. Recall, however, that all participants—including men in Study 5—always received the same score of +9/12, regardless of their actual math performance or the feedback condition they were assigned to. Because all participants received the same score, it seems unlikely that receiving "Good job!" feedback from male authority figures only helped women interpret their performance, but not men.

A third possibility is that positive feedback helped to reduce uncertainty about one's ability and belonging in math. In math settings, women are susceptible to identity threat; they are likely to be concerned about being negatively evaluated based on their gender and are, therefore, attentive to cues in the environment that signal whether or not they are capable and "fit in" with the domain. Indeed, female students report less confidence and ability in math than males, even after accounting for actual observed ability in math (Perez-Felkner, Nix, & Thomas, 2017). In the current research, women perceived positive feedback in math from a male authority figure to imply that they had the ability to succeed in math and that they belonged in this field (Study 1b). When women received feedback on a math test from a female authority figure, the type of feedback did not matter in shaping their perceptions of the implications of the feedback. Men perceived positive feedback in math to imply that they had ability in this area, but it did not matter whether this feedback came from a gatekeeper or not.

Results of mediated moderation analyses (using PROCESS) in Study 1b, however, revealed that all the 95% CIs of the indirect effect of the IV (i.e., professor sex × feedback type interaction) on the DV (i.e., belonging in math) through the mediators (i.e., perceptions of the feedback conveying that one belonged in math and had the ability to succeed in this field) included zero, indicating nonsignificant indirect effects. Thus, it does not seem to be the case that women's perceptions of the implications of feedback affected their feelings of belonging in math.

A fourth possibility is that women who received positive feedback from a male authority figure came to identify themselves more with math on an implicit level, which then increased their explicit attitudes/identification/interest in math and related fields. Supporting this idea, results of mediated moderation analyses was significant. Specifically, the (a) IV (i.e., the experimenter sex × feedback type interaction) significantly predicted the DV (i.e., explicit attitudes/identification/interest in STEM; $B = 0.71$, $p = .049$), (b) the IV significantly predicted the mediator (i.e., implicit math identification; $B = -0.23$, $p = .031$), and (c) the mediator significantly predicted the DV (i.e., explicit attitudes/identification/interest in STEM; $B = -1.20$, $p = .001$). We then used bootstrapping analyses (Preacher & Hayes, 2008) based on 5,000 resamples to test the 95% CI of the indirect effect of the IV (i.e., experimenter sex × feedback type interaction) on the DV (i.e., explicit attitudes/identification/interest in STEM) through the mediator (i.e., implicit math identification). The bias-corrected confidence interval for the size of the indirect effect on the DV excluded zero, indicating a significant indirect effect of implicit math identification on explicit attitudes/identification/interest in STEM, $B = 0.49$, 95% CI = [0.08, 1.20]. Results also indicated that the direct effect of the experimenter sex × feedback type interaction on attitudes/identification/interest in STEM became nonsignificant ($B = 0.53$, $p = .13$) when implicit math identification was included in the model, suggesting full mediation.

Although we did not make this prediction a priori, the link between shifts in implicit measures and explicit responses is consistent with past research. For example, Stout and colleagues (2011) found that women who saw female experts in

STEM showed stronger identification with STEM fields at an implicit level, and reported greater self-efficacy and commitment to STEM careers at an explicit level. Such findings suggest that subtle shifts in women's self-concept—in the form of greater implicit identification with math—may be an important variable to assess when examining women's STEM outcomes.

### *Effects of Exposure to Outgroup (vs. Ingroup) Members*

A growing body of research suggests that contact between ingroup and outgroup members can lead to positive outcomes for minorities in academic settings (Mendoza-Denton & Page-Gould, 2008; Page-Gould, Mendoza-Denton, Mendes, 2014). For example, African American students who had greater exposure to White students in high school showed better adjustment when they entered predominantly White college settings later in life (Graham, Baker, & Wapner, 1984). Even populations at risk of perceiving negativity from outgroups, such as minority students who are sensitive to race-based rejection, benefit from having cross-group friendships with majority group members. Indeed, both experimental and longitudinal evidence reveals that cross-group friendships can buffer minority students from feelings of low belonging and dissatisfaction with their university (Mendoza-Denton & Page-Gould, 2008) and protect against increases in anxious mood that might otherwise occur among minorities in academic settings (Page-Gould, Mendoza-Denton, & Tropp, 2008). Cross-group friendships do not have the same benefits for majority students; this makes sense, given that White students have less reason to doubt their ability or acceptance in academic institutions, where they have historically been represented and viewed as belonging (Mendoza-Denton & Page-Gould, 2008; Page-Gould et al., 2008).

Of relevance to women in STEM, Walton, Logel, Peach, Spencer, and Zanna (2015) found that increasing belonging helped to mitigate the chilly climate and social marginalization that women often experience in engineering. Specifically, first-year female engineering students who were assigned to a social-belonging intervention—in which feelings of belonging uncertainty were normalized and the similarities between men and women in engineering were emphasized—showed more positive academic attitudes, more stable self-esteem, and better academic performance compared with students who participated in a study skills intervention. Cheryan, Siy, Vichayapai, Drury, and Kim (2011) also found that women felt more similar to a both male and female non-stereotypical (vs. stereotypical) role models in computer science, which then increased their beliefs that they would be successful in STEM. Contributing to this literature, we found that positive feedback from outgroup members in math is an important form of intergroup contact and contextual cue that benefits women in situations where they are likely to experience identity threat, such as in math settings.

In particular, receiving positive feedback on a math test from an outgroup member may have helped to reduce the experience of "own-reputation threat" from the outgroup. According to Shapiro and Neuberg's (2007) Multi-Threat Framework, own-reputation threat occurs when individuals are concerned about being judged or treated negatively by outgroup members due to the possibility of being negatively stereotyped. This occurs when individuals (a) recognize that they belong to a negatively stereotyped group; (b) think that others could recognize that they belong to the negatively stereotyped group; (c) believe that their stereotype-relevant behavior is, or could be, publicly visible to others; and (d) believe that others might think the negative stereotype could be true of them. In the current research, women were in a context where they were taking a math test in the presence of an outgroup or ingroup member and their performance was publicly visible to the evaluator. If women were concerned about being evaluated by outgroup members through the lens of their gender identity, then it makes sense that positive feedback from outgroup members (i.e., male authority figures) would be particularly effective at creating a sense of identity safety for women in this setting.

Finally, the current findings may seem at odds with previous work on the benefits of ingroup role models. However, there are important distinctions to make between the two streams of research that may help to account for the differences. First, previous research suggests that ingroup role models are effective when they (a) are perceived to be similar to the participant, (b) their success is viewed as attainable, and (c) individuals are highly identified with the domain in which the role model excels (Cheryan & Plaut, 2010; Lockwood & Kunda, 1997; Marx & Roman, 2002; Stout et al., 2011). Individuals are thought to feel a sense of connection with ingroup role models, and this subjective identification increases the salience of a "possible self" that suggests how to attain similar outcomes as the ingroup member in the future (Markus & Nurius, 1986).

In the current studies, the focus was not on the effects of role models, but on the effects of feedback from perceived gatekeepers in STEM. We, therefore, did not assess perceptions of similarity, attainability, or self-relevance of the authority figure's competence in math; women were not expected to seek out the female authority figure as a standard of comparison, but to respond based on the type of feedback and from whom they received this feedback. Thus, our focus—besides the type of feedback participants received—was on the feedback provider's gatekeeping status in STEM fields. As the preliminary study revealed, men were viewed as the numerical majority in math and as having more authority and expertise in STEM than women. We did not give additional information to participants conveying that the female authority figure was exceptionally competent or interested in STEM fields, as past role model studies have often done (Marx & Roman, 2002; McIntyre et al., 2003; Stout et al., 2011). Rather, the authority figure simply graded the tests and gave participants positive or objective feedback about their performance.

Another difference between the current work and previous research is that we examined effects of math test feedback among students taking introductory psychology courses who varied in their initial interest in STEM. According to the stereotype inoculation model (Dasgupta, 2011), exposure to ingroup members in high-stakes achievement settings is most effective when individuals are highly invested in and care about doing well in the domain. For example, women who were highly identified with math were protected from the negative effects of stereotype threat on their math performance when they saw a competent female role model in math (Marx & Roman, 2002; McIntyre et al., 2003). Other studies have shown that female STEM majors who saw same-sex STEM role models expressed greater self-efficacy in STEM and more positive implicit attitudes and identification with STEM (Stout et al., 2011). First-year female engineering students also felt less anxious and participated more in female-majority and sex-parity groups than women in female-minority groups (Dasgupta et al., 2015).

In contrast, participants in the current studies were introductory psychology students, the majority of whom were non-STEM majors (70% non-STEM majors in Study 1b, 77.8% in Study 2, 67.9% in Study 3, 90.4% in Study 4). Although we did not have enough STEM majors to test whether identification with STEM moderates the effects found in the current studies, future research could examine the effects of feedback at different stages of the STEM pipeline. For example, whereas identification with ingroup role models might be effective in the retention of women who are already invested in STEM (Dasgupta, 2011; Drury et al., 2011), positive feedback from male professors and authority figures could be beneficial in recruiting women into STEM. Positive feedback may also be beneficial to students who lack confidence about whether they have the skills to succeed in the domain. If so, then targeting students when they are most likely to doubt their abilities, such as the transition from high school to college, may be a key time point to intervene with positive feedback, especially coming from gatekeepers in a field.

### Positive Feedback Versus Other Types of Feedback

The current research differs from previous research using *wise feedback* in a number of ways. First, although studies using wise feedback also examined the effects of positive feedback, this was typically done in conjunction with the delivery of critical feedback; positive feedback was added as a buffer against critical feedback (Cohen et al., 1999). In contrast, the present studies gave positive feedback only in the context of the supposed score that participants received, rather than being paired with critical feedback or acting as a buffer against negative feedback.

A second difference is that the current studies always involved giving feedback that framed one's objective performance on a multiple-choice exam, whereas studies on wise feedback typically involved giving feedback about performance that was more subjective in nature, such as an essay or presentation (Cohen et al., 1999). It, therefore, remains an empirical question as to how effective positive feedback would be in situations that are less objective, such as grading the work that goes into solving a math problem.

Third, wise feedback invokes high standards and reassures the student that he or she has the ability to meet those standards. Students in the wise feedback condition in previous studies received written comments such as, "Remember, I wouldn't go to the trouble of giving you this feedback if I didn't think, based on what I've read in your letter, that you are capable of meeting the higher standard I mentioned" (Cohen et al., 1999, p. 1307), or "I'm giving you these comments because I have very high expectations and I know that you can reach them" (Yeager et al., 2014, p. 809). Whereas participants in the wise feedback studies were made aware of the reasons underlying the feedback they received, participants in the positive feedback conditions in the current studies only received the written comment, "Good job!" along with their supposed test score; there was no mention of why participants received the feedback they received. In short, whereas wise feedback involves giving critical feedback accompanied by a rationale for why the feedback was given, the feedback in the present research was always brief, positive, and did not elaborate on the reasons for giving the feedback.

### Future Directions

The current studies examined the effects of positive versus objective feedback while keeping participants' supposed score on the test constant across feedback conditions. Future research would benefit from examining the conditions under which positive feedback is helpful versus not. For example, "Good job!" may be helpful feedback when one's objective score is ambiguous (e.g., +9/12), but not when one receives a perfect score (e.g., +/12/12). Such feedback could even be viewed as condescending and produce negative outcomes if it is accompanied by a very low score (e.g., +3/12). Similarly, "Good job!" feedback might need to be perceived as authentic and grounded in legitimate achievement. If feedback is interpreted as inaccurate or attributable to external sources, it may do more harm than good.

Research could also examine whether there is something unique about positive feedback, or whether providing any type of feedback matters. For example, a feedback condition could be added in which participants receive their test score plus a written comment that does not directly convey encouragement (e.g., "here is your score, +9/12), yet still involves time and effort on the part of the feedback provider. Although it remains an open question, theoretically, it seems unlikely that receiving neutral feedback would function in the same way as positive feedback in conveying that one belongs or has the ability to succeed in the domain.

Although we did not collect data on how participants perceived their score of +9/12, we did have information on how participants actually performed on the math test. In the studies where women took an actual math test, their scores ranged from +0 to +12, with average scores ranging between +5.12 and +5.60 and standard deviations between 2.10 and 2.57. Given such scores, it seems unlikely that participants construed the +9/12 score in a negative light. Rather, for most participants, receiving a +9/12 on the math test may have been viewed as a higher score than they expected. Importantly, though, it was not the score itself (i.e., objective feedback) that led to boosts in women's math-related outcomes. Rather, it was receiving their supposed score plus positive feedback from a male gatekeeper that increased women's math-related outcomes.

In the present studies, feedback was always given by an ingroup or outgroup member who was in a position of authority (i.e., a professor, experimenter). Thus, it remains to be seen whether feedback providers need to be in a position of power to elicit these responses, or whether they can just be a member of the majority outgroup to produce these effects. In addition, although we did not directly compare receiving feedback from a professor versus an experimenter in the current studies, our conceptualization of feedback is that it can come from any agent (Hattie & Timperley, 2007). Indeed, the experimenter did many of the things a professor might do, such as giving instructions, grading the exams, giving feedback about students' performance, and assigning credit for their participation. However, because professors have more experience and training and ultimately determine students' grades, they may be viewed as having more authority and gatekeeper status than an undergraduate research assistant, which could make their feedback more impactful than an experimenter's.

Another direction for future research is to examine how a broader range of groups respond to feedback. Study 5 found that men showed better math-related outcomes when they received positive versus objective feedback, but this did not depend on the gatekeeper status of the authority figure. Although we were unable to directly compare men's and women's responses with feedback in this study, the results suggest a potential boundary condition of the effects of feedback. Studies could examine situations in which men might also be affected by the source of feedback. For example, men may be influenced more by feedback from female gatekeepers in the domain of social sensitivity—an area in which men typically underperform relative to women (Leyens et al., 2000). White men may also experience identity threat—and thus, be more attentive to the source of feedback—if they are told that Asians outperform Whites in math (Aronson, Lustina, Good, Keough, & Steele, 1999).

## Conclusion

The present research shows that individuals who are at risk of being negatively evaluated by others based on their social identity, such as women in math settings, benefit from receiving positive feedback from gatekeepers. Specifically, women who received positive (vs. objective) feedback on a math test showed greater feelings of confidence, self-efficacy, belonging, more favorable attitudes/identification/interest in STEM fields, and greater implicit identification with math when they received this feedback from a gatekeeper in math (i.e., a male vs. female professor or experimenter). Meta-analyses across studies confirmed these results. Future research would benefit from examining further the conditions under which the type, source, and domain of feedback differentially shape minority and majority group members' self-evaluations and achievement-related outcomes.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### Notes

1. When we ran the analyses in Study 2 without excluding any participants, the same pattern of results emerged. The experimenter sex × feedback type interaction remained significant, $F(1, 117) = 6.18$, $p = .01$, $\eta_p^2 = .05$; women who received feedback from a male experimenter felt greater math self-efficacy when they received positive ($M = 0.26$, $SD = 0.78$) versus objective feedback ($M = -0.25$, $SD = 0.75$), $F(1, 117) = 5.77$, $p = .02$, $\eta_p^2 = .05$, 95% confidence interval (CI) = [0.09, 0.93]. Women who received positive feedback also tended to show greater math self-efficacy when they received this feedback from a male versus female experimenter ($M = -0.13$, $SD = 0.99$), $F(1, 117) = 3.29$, $p = .07$, $\eta_p^2 = .03$, 95% CI = [−0.04, 0.82]. When women received objective feedback, they tended to report lower math self-efficacy when this feedback came from a male versus female experimenter ($M = 0.12$, $SD = 0.86$), $F(1, 117) = 2.89$, $p = .09$, $\eta_p^2 = .02$, 95% CI = [−0.80, 0.06]. No other simple effects approached significance, all $p$s > .26.
2. When we controlled for actual math test performance in Studies 2 to 5, the pattern of results was the same except for math self-efficacy in Study 5, in which the marginal main effect of feedback type became nonsignificant, $p = .24$.
3. When we ran the analyses in Study 3 without excluding any participants, the same pattern of results emerged. The experimenter sex × feedback type × test domain interaction remained significant, $F(1, 178) = 5.41$, $p = .02$, $\eta_p^2 = .03$: Women who took a math test and received positive feedback from a male experimenter tended to report more favorable attitudes/identification/interest in STEM ($M_{adj} = 4.80$, $SE = 0.29$) versus when they received objective feedback from a male experimenter on

a math test ($M_{adj}$ = 4.06, *SE* = 0.31), *F*(1, 178) = 2.97, *p* = .09, $\eta_p^2$ = .02, 95% CI = [−0.11, 1.58], or when they received positive feedback on a math test from a female experimenter ($M_{adj}$ = 4.10, *SE* = 0.29), *F*(1, 178) = 3.44, *p* = .065, $\eta_p^2$ = .02, 95% CI = [−1.52, 0.05]. Women who received positive feedback from a male experimenter also tended to report more favorable attitudes/identification/interest in STEM when they received feedback on a math test versus a verbal test ($M_{adj}$ = 4.07, *SE* = 0.31), *F*(1, 178) = 2.94 *p* = .09, $\eta_p^2$ = .02, 95% CI = [−0.11, 1.56].

4. Eight participants did not complete the initial interest in STEM item, which was collected prior to the lab session as a baseline measure. However, given that they completed all of the lab portion of the study, their data were retained in the analyses.
5. The average score on the math test was *M* = 5.86 (*SD* = 2.82) out of 12 points; for the verbal test, it was *M* = 5.55 (*SD* = 2.70) out of 12 points. There was no significant difference across the tests, *t*(166) = 0.76, *p* = .45, suggesting that both tests were equally difficult. Also, across Studies 3 to 5, there were three to four male research assistants and three to five female research assistants who served as experimenters in each study. All the experimenters were undergraduate research assistants who assumed a position of authority in the lab, given that they were in charge of running the study, grading the exams, giving feedback, and, therefore, being knowledgeable about the correct answers.
6. We controlled for baseline differences in initial interest in STEM because this variable was positively correlated with interest in STEM as assessed in the lab (*r* = .61, *p* < .001). When the analyses were run without this covariate, the experimenter sex × feedback type × test domain interaction predicting STEM attitudes/identification/interest became nonsignificant, *F*(1, 160) = 1.86, *p* = .17, $\eta_p^2$ = .01. The only significant simple effect was that women who received feedback from a male experimenter reported more favorable attitudes/identification/interest in STEM when they received positive feedback on a math test ($M_{adj}$ = 4.87, *SE* = 0.39) versus a verbal test ($M_{adj}$ = 3.72, *SE* = 0.42), *F*(1, 160) = 6.59, *p* = .047, $\eta_p^2$ = .02, 95% CI = [0.01, 2.28].
7. When we ran the analyses without excluding any participants in Study 4, the same pattern of results emerged for implicit identification with math, although the experimenter sex × feedback type interaction became marginally significant, *F*(1, 105) = 3.63, *p* = .06, $\eta_p^2$ = .03. Women who received feedback on a math test from a male experimenter tended to more quickly associate themselves with math (vs. the arts) when they received positive (*M* = 0.36, *SD* = 0.45) versus objective feedback (*M* = 0.61, *SD* = 0.51), *F*(1, 105) = 3.54, *p* = .06, $\eta_p^2$ = .03, 95% CI = [−0.52, 0.01]. Women also tended to more quickly associate themselves with math (vs. arts) when they received positive feedback on a math test from a male versus female experimenter (*M* = 0.56, *SD* = 0.54), although this effect became nonsignificant, *F*(1, 105) = 2.51, *p* = .12, $\eta_p^2$ = .02, 95% CI = [−0.46, 0.05]. No other simple effects approached significance, all *p*s > .26. When we ran the analyses without excluding any participants to predict explicit STEM attitudes/identification/interest, there was only a significant main effect of experimenter sex, *F*(1, 108) = 5.73, *p* = .02, $\eta_p^2$ = .05; no other main effects or interactions were significant, all *p*s > .22.
8. Due to a computer malfunction, we were unable to collect Implicit Association Test (IAT) data from three participants. However, their data were retained in the analyses because they completed all other measures. In addition, whereas Study 3 controlled for initial interest in STEM, this item was not part of mass testing in Study 4 and, therefore, was not included as a covariate.
9. It is important to note, however, that not all men are immune to negative stereotypes about their math ability. Men of color, for example (i.e., being Black, Hispanic) may be negatively stereotyped based on their intellectual aptitude, which likely includes quantitative ability (Steele & Aronson, 1995; Steele et al., 2002).
10. When we ran the analyses without excluding any participants in Study 5, the results were similar, except for perceptions of the implications of the feedback conveying a challenge to the gender math stereotype and for feelings of confidence. For the former, there was a significant main effect of feedback type, *F*(1, 136) = 4.08, *p* = .045, $\eta_p^2$ = .03, 95% CI = [0.01, 1.55], such that men who received positive feedback perceived this feedback to more strongly convey that men and women are equally capable of doing well at math (*M* = 3.86, *SD* = 2.23) than receiving objective feedback (*M* = 3.06, *SD* = 2.29). For feelings of confidence, there were no significant main effects or interaction, all *p*s > .19. The remaining results were similar to the ones reported earlier: For perceptions of the feedback conveying ability in math, the main effect of feedback type remained significant, *F*(1, 136) = 6.13, *p* = .02, $\eta_p^2$ = .04, 95% CI = [0.15, 1.39]: Men who received positive feedback perceived this feedback to more strongly convey their ability in math (*M* = 3.86, *SD* = 1.96) than receiving objective feedback (*M* = 3.01, *SD* = 1.73). For perceptions of the feedback conveying belonging in math, there were no significant main effects or interaction, all *p*s > .35. For current belonging in math, the main effect of feedback type remained significant, *F*(1, 136) = 3.93, *p* = .049, $\eta_p^2$ = .03, 95% CI = [0.00, 0.79]: Men reported greater belonging in math when they received positive (*M* = 2.59, *SD* = 1.18) versus objective feedback (*M* = 2.22, *SD* = 1.15).

## Supplemental Material

Supplementary material is available online with this article.

## References

Abrams, D., & Hogg, M. A. (Eds.). (1999). *Social identity and social cognition*. Oxford, UK: Blackwell.

Adams, G., Garcia, D., Purdie-Vaughns, V., & Steele, C. (2006). The detrimental effects of a suggestion of sexism in an instruction situation. *Journal of Experimental Social Psychology*, *42*, 602-615.

Akcinar, E. N., Carr, P. B., & Walton, G. M. (2011). Interactions with men and whites matter too. *Psychological Inquiry*, *22*, 247-251.

Aronson, J., Lustina, M. J., Good, C., Keough, K., & Steele, C. M. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35*, 29-46.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122-147.

Blascovich, J., Spencer, S. J., Quinn, D. M., & Steele, C. M. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, *12*, 225-229.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631-643.

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 35-58). Oxford, UK: Blackwell.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science*, *9*, 333-342.

Cheryan, S., Drury, B. J., & Vichayapai, M. (2013). Enduring influence of stereotypical computer science role models on women's academic aspirations. *Psychology of Women Quarterly*, *37*, 72-29.

Cheryan, S., & Plaut, V. C. (2010). Explaining underrepresentation: A theory of precluded interest. *Sex Roles*, *63*, 475-488.

Cheryan, S., Plaut, V. C., Davies, P., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045-1060.

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*, 656-664.

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, *25*, 1302-1318.

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (2004). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, *30*, 1332-1346.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *12*, 163-170.

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry*, *22*, 231-246.

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*, 642-658.

Dasgupta, N., McManus Scircle, M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academic of Sciences*, *112*, 4988-4993.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*, 1615-1628.

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*, 276-287.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*, 105-115.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between roles and goals: A new look at why women opt out of STEM careers. *Psychological Science*, *21*, 1051-1057.

Drury, B. J., Siy, J. O., & Cheryan, S. (2011). When do female role models benefit women? The importance of differentiating recruitment from retention in STEM. *Psychological Inquiry*, *22*, 265-269.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*, 183-201.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013-1027.

Goffman, E. (1963). *Stigma: Notes on the management of a spoiled identity*. New York, NY: Touchstone.

Graham, C., Baker, R. W., & Wapner, S. (1984). Prior interracial experience and Black student transition into predominantly White colleges. *Journal of Personality and Social Psychology*, *47*, 1146-1154.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *10*, 24-27.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.

Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, *37*, 1352-1363.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81-112.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, *11*, 365-371.

Kawakami, K., Steele, J. R., Sifa, C., Phills, C. E., & Dovidio, J. F. (2008). Approaching math increases math=me and math=pleasant. *Journal of Experimental Social Psychology*, *44*, 818-825.

Leyens, J., De´sert, M., Croizet, J., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin*, *26*, 1189-1199.

Lockwood, P. (2006). Someone like me can be successful: Do college students need same- gender role models? *Psychology of Women Quarterly*, *30*, 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*, 91-103.

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology*, *96*, 1089-1103.

Markus, H. R., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, *51*, 858-866.

Markus, H. R., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954-969.

Markus, H. R., Steele, C. M., & Steele, D. M. (2002). Colorblindness as a barrier to inclusion: Assimilation and nonimmigrant minorities. In R. Shweder, M. Minow, & H. R. Markus (Eds.), *Engaging cultural differences: The multicultural challenge in liberal democracies* (pp. 453-472). New York, NY: Russell Sage Foundation.

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

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, *66*, 90-103.

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*, 83-90.

Mendoza-Denton, R., & Page-Gould, E. (2008). Can cross-group friendships influence minority students' well-being at historically white universities? *Psychological Science*, *19*, 933-939.

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, *18*, 879-885.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math=male, me=female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44-59.

Nosek, B. A., & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal*, *48*, 1124-1154.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 782-789.

Page-Gould, E., Mendoza-Denton, R., & Mendes, W. B. (2014). Stress and coping in interracial contexts: The influence of race-based rejection sensitivity and cross-group friendship in daily experiences of health. *Journal of Social Issues*, *70*, 256-278.

Page-Gould, E., Mendoza-Denton, R., & Tropp, L. (2008). With a little help from my cross-group friends: Reducing intergroup anxiety through cross-group friendship. *Journal of Personality and Social Psychology*, *95*, 1080-1094.

Park, L. E., Cook, K. E., & Greenwald, A. G. (2002). Implicit indicators of women's persistence in math, science, and engineering. *Psi Chi Journal of Undergraduate Research*, *6*, 145-152.

Park, L. E., Young, A. F., Troisi, J. D., & Pinkus, R. T. (2011). Effects of everyday romantic goal pursuit on women's attitudes toward math and science. *Personality and Social Psychology Bulletin*, *37*, 1259-1273.

Perez-Felkner, L., Nix, S., & Thomas, K. (2017). Gendered pathways: How mathematics ability beliefs shape secondary and postsecondary course and degree field choices. *Frontiers in Psychology*, *8*, 1-11.

Pollack, E. (2013, October 3). Why are there still so few women in science? *The New York Times Magazine*. Retrieved from http://www.nytimes.com/2013/10/06/magazine/why-are-there-still-so-few-women-in-science.html?pagewanted=1&_r=3&ref=magazine&

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, A. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*, 615-630.

Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5*, 52-73.

Schmader, T., & Croft, A. (2011). How stereotypes stifle performance. *Personality and Social Psychology Compass*, *5*, 792-806.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336-356.

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, *39*, 68-74.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, *11*, 107-130.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4-28.

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

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379-440). San Diego, CA: Academic Press.

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*, 255-270.

Tracy, J. L., & Robins, R. W. (2007). The psychological structure of pride: A tale of two facets. *Journal of Personality and Social Psychology*, *92*, 506-525.

Umbach, P., & Wawrzynski, M. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education*, *46*, 153-184.

U.S. Department of Labor. (2009). *Women in the labor force: A databook (2009)*. Washington, DC: Bureau of Labor Statistics, U.S. Department of Labor.

Vallerand, R. J., & Reid, G. (1998). On the relative effects of positive and negative verbal feedback on males' and females' intrinsic motivation. *Canadian Journal of Behavioral Science*, *20*, 239-250.

Vogt, C. M. (2008). Faculty as a critical juncture in student retention and performance in engineering programs. *The Research Journal for Engineering Education*, *97*, 27-36.

Vogt, C. M., Hocevar, D., & Hagedorn, L. (2007). A social cognitive construct validation: Determining women's and men's success in engineering programs. *Journal of Higher Education*, *78*, 337-364.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*, 82-96.

Walton, G. M., Logel, C., Peach, J., Spencer, S. J., & Zanna, M. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*, 468-485.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961-977.

Wout, D. A., Shih, M. J., Jackson, J. S., & Sellers, R. M. (2009). Targets as perceivers: How people determine when they will be negatively stereotyped. *Journal of Personality and Social Psychology*, *96*, 349-362.

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., . . . Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, *143*, 804-824.

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The influence of female role models on women's implicit science cognitions. *Psychology of Women Quarterly*, *37*, 283-292.