Psychology of Women Quarterly
2016, Vol. 40(3) 378-397

DOI: 10.1177/0361684316630965
pwq.sagepub.com
SAGE

STEM Stereotypic Attribution Bias Among Women in an Unwelcoming Science Setting

Jennifer LaCosse[1], Denise Sekaquaptewa[2], and Jill Bennett[2]

Abstract

Science, technology, engineering, and math (STEM) stereotypic attribution bias (SSAB) is the tendency to spontaneously generate external attributions for men's setbacks in STEM fields and to spontaneously make internal attributions for women's setbacks in STEM fields. Among samples of undergraduate STEM students, STEM settings perceived as unwelcoming to women through self-report (Study 1) and a manipulation (Study 2) were shown to predict SSAB. Among undergraduate women, experiencing the negative treatment of other women in a science setting predicted SSAB, which was negatively correlated with feelings of belonging in STEM (Study 1) and with intentions to continue in STEM after graduation (Studies 1 and 2). Research materials (e.g, data, measures, materials, etc.) used in both studies will be made available upon request to either of the first two authors. Those interested in increasing retention of women in STEM majors should develop strategies designed to reduce internal attributions for women's setbacks among women facing negative STEM environments and should cultivate a more positive climate for women in STEM fields. *Online slides for instructors who want to use this article for teaching are available on* PWQ's *website at http://pwq.sagepub.com/supplemental*

Keywords

stigma, psychology of women, stereotyped behavior, attribution, sex discrimination

It is well-documented that women are underrepresented in science, technology, engineering, and mathematics (STEM) fields, particularly at higher education levels such as postdoctoral positions and tenured/tenure track professorships (e.g., Shaw & Stanton, 2012). One factor that diminishes women's intentions to continue in STEM is the negative environment that they may experience (e.g., Beller & Gafni, 1996; Catsambis, 1995; Eccles & Blumenfeld, 1985). Increasing knowledge about the consequences of perceived negative environments for women is an important step towards understanding the underrepresentation of women in STEM. In the current work, we hypothesized that one cognitive consequence of an unwelcoming STEM environment may be the generation of attributions biased by stereotypes about men's and women's abilities in STEM.

Drawing from research in both the academic achievement attribution literature and work by Sekaquaptewa and colleagues on *stereotypic explanatory bias* (Sekaquaptewa & Espinoza, 2004; Sekaquaptewa, Espinoza, Thompson, Vargas, & von Hippel, 2003), we investigated the relation between perceived negative environment for women and students' attributions for men's and women's STEM-related behaviors, and their relation to intentions to continue in STEM. We predicted that women's STEM setbacks (which are stereotype-consistent behaviors) would be spontaneously attributed to internal causes (e.g., inherent ability) whereas men's STEM setbacks (which are stereotype-inconsistent behaviors) would be attributed to external causes (e.g., unfortunate circumstances). We present a novel method to measure this bias and extend previous research on stereotypic attribution research (e.g., Sekaquaptewa et al., 2003; Sekaquaptewa & Espinoza, 2004) to the domain of gender-STEM stereotypes. Finally, we examine the effects of biased attribution patterns on STEM women's feelings of belonging in their field as well as their intentions to continue in STEM.

Gender-STEM Stereotypes and Achievement Attribution

One factor proposed to influence women's intentions to participate in STEM is gender-STEM stereotypes. Gender-STEM stereotypes reflect the belief that women have inferior ability in these fields compared to men; previous work shows

[1] Department of Psychology, Florida State University, Tallahassee, FL, USA
[2] Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Corresponding Author:
Jennifer LaCosse, Department of Psychology, Florida State University, 1107 W Call Street, Tallahassee, FL 32306, USA.
Email: jlacosse@fsu.edu

the negative influence of these stereotypes on women's performance, persistence, and belonging in STEM fields (e.g., Murphy, Steele, & Gross, 2007; Spencer, Steele, & Quinn, 1999). Gender-STEM stereotypes create behavioral expectancies that may influence causal explanations for women's behavior in STEM. Past research supports this claim and indicates that one way gender-STEM stereotypes affect women's STEM outcomes is through exerting influence on achievement attribution styles (Reyna, 2000).

According to attribution theory, people may engage in positive or negative attribution styles when explaining their own successes and failures (Weiner, 1979). Individuals evincing a positive attribution style attribute their success to internal factors (e.g., high ability) and attribute their failures to external influences (e.g., bad luck or a weak effort). On the other hand, individuals evincing a negative attribution style show the opposite pattern, making external attributions for successes and internal attributions for failures. Achievement attribution predicts academic motivation, persistence, and performance (Henry, Martinko, & Pierce, 1993; Hong, Dweck, Lin, & Wan, 1999; Weiner, 1979). Positive attribution styles predict more positive academic outcomes and negative attribution styles predict more negative academic outcomes. For example, those who view their success as being due to skill, as opposed to chance, have a higher expectancy for future success, whereas those who view their failures as due to lack of skill as opposed to bad luck have a more pessimistic view of their future success (Kelley & Michela, 1980; Martinez & Sewell, 2000; Peterson & Barrett, 1987).

Gender-STEM stereotypes can influence whether people engage in either positive or negative patterns of achievement attribution by creating positive or negative expectations for the performance of men and women (Reyna, 2000). For example, in STEM domains, people tend to attribute men's successes to competence, and women's successes to good luck, while attributing men's failures to situational factors, but women's failures to a lack of competence (Swim & Sanna, 1996). This pattern also holds for women's attributions for their own performance (Koch, Müller, & Sieverding, 2008), with important consequences. Kiefer and Shih (2006) found that when women received feedback about their performance on a math and a verbal test, they were more sensitive to negative feedback concerning the math test, and attributed their failure to low ability. Men, who are stereotyped as having less verbal ability than women, were more sensitive to feedback about the verbal test. Kiefer and Shih (2006) explain that women, compared to men, are more likely to attribute failure on the math test to internal and stable factors resulting in lower perseverance in math, which offers a potential explanation for why some women decide not to continue in math-related fields after graduation.

STEM Stereotypic Attribution Bias (SSAB)

Because expectations about people can develop from stereotypes (Hamilton, Sherman, & Ruvolo, 1990), an attribution bias may emerge regarding stereotype-consistent versus inconsistent behaviors. Sekaquaptewa and colleagues (Sekaquaptewa et al., 2003; Sekaquaptewa & Espinoza, 2004) propose that stereotypic explanatory bias is the tendency to spontaneously provide explanations for stereotype-inconsistent over stereotype-consistent behaviors. The tendency to explain away stereotype disconfirmation serves to maintain the stereotype, as behaviors defying the stereotype are discounted by explanations that may reduce the potential of the behavior to change the stereotype. Stereotypic explanatory bias can promote discrimination, as those who engage in stereotypic explanatory bias may selectively discount the counter-stereotypic behaviors of stereotyped targets, and thereby behave towards that target in a manner that is primarily driven by stereotypes. Consistent with this idea, previous work demonstrated that White individuals who showed this bias regarding race stereotypes had more negative social interactions with Black partners (Sekaquaptewa et al., 2003).

In the current work, we predicted that due to gender-STEM stereotypes, the gender of the actor would influence the type of attributions generated (internal vs. external) for STEM outcomes. We define SSAB as the spontaneous attribution of women's STEM setbacks to internal causes (e.g., inherent ability) and men's STEM setbacks to external causes (e.g., unfortunate circumstances), and the attribution of women's STEM successes to external causes and men's STEM successes to internal causes. For example, if a male student performs poorly on a math test, this is inconsistent with gender-STEM stereotypes. A person showing SSAB may therefore spontaneously attribute a man's poor performance to an external cause, such as the test being unrealistically difficult. On the other hand, a woman's failure on a math test is stereotype consistent and may therefore be spontaneously attributed to an internal factor, such as her low ability in math. Thus, the same behavior may be attributed differently depending on the gender of the actor, particularly when stereotypes about the ability of men and women in the performance domain are activated and influence information processing (Vargas, Sekaquaptewa, & von Hippel, 2004). By engaging in this attribution bias, people may be able to resolve inconsistencies between their stereotypic expectations and the events they witness by attributing expectancy-inconsistent events to external or fleeting circumstances, while also supporting their stereotypic expectancies by attributing expectancy-consistent events to stable and enduring dispositions (Sekaquaptewa & Espinoza, 2004).

Past research on spontaneous attributional processing supports our hypothesis; spontaneous attributional processing is most likely to emerge when one finds a behavior to be unexpected or surprising, given one's expectations for the actor (Hastie, 1984). For example, if one expects an actor, Tom, to be an excellent student, one may be surprised to hear that "Tom received the lowest test score in the class." Learning about this event can lead to the spontaneous attribution, "Tom was sick that day," an attribution that helps maintain one's expectations of Tom. Our assessment of SSAB assumes a spontaneous, initial reaction to gender-STEM stereotype relevant events may reflect a less controlled stage of information processing, in which inconsistency between one's expectations and the actor's behavior is resolved through the use of an attribution that supports one's expectations. This is in contrast to attribution assessments that directly prompt people to provide attributions for stereotype-relevant events.

Given past research on achievement attribution (Peterson, 1992; Weiner, 1979), we might expect that stereotypic attribution patterns would emerge equally for successes and failures. However, responses to failure in STEM experiences may be particularly influenced by gender-STEM stereotypes. Research on naive causal theories (Ybarra, 2002) indicates that negative behaviors are seen as particularly diagnostic of dispositions, whereas positive behaviors are seen as being driven by social norms. These naive theories make non-normative negative behaviors more salient and more meaningful when people make judgments about an individual (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Fiske, 1980). Therefore, although stereotypically biased attributions might emerge equally in regards to STEM successes and setbacks, the work on naive causal theories suggests the possibility that SSAB may emerge more strongly when STEM setbacks occur compared with STEM successes. To the extent that negative behaviors tend to be more diagnostic of dispositions, traits, and abilities, SSAB based on STEM failures may be particularly meaningful as a predictor of STEM intentions.

Negative Environments, Belonging in STEM, and SSAB

Research has demonstrated that classroom climate influences students' motivation, learning, and academic performance (e.g., Ames, 1992; Church, Eliot, & Gable, 2001). For example, Wong and Fraser (1996) found that individuals who experienced a "chillier" climate in a chemistry lab had more negative attitudes toward the field of chemistry. The negative repercussions of experiencing a "chilly" environment may influence the retention of female students in STEM. Perceived gender-STEM stereotype relevance can be increased via environmental cues, including the absence of women in the setting (Dasgupta, 2011; Murphy et al., 2007; Sekaquaptewa & Thompson, 2002, 2003), the presence of objects associated with masculinity (Cheryan, Plaut, Davies, & Steele, 2009), and the perception that women are disfavored in the setting (Beller & Gafni, 1996; Catsambis, 1995; Eccles & Blumenfeld, 1985).

It is important to understand the mechanisms underlying the effects of environmental cues. Research on the attrition of women from STEM indicates that chilly climates lead to decreased interest in STEM by influencing women's sense of belonging in the field (e.g., Cheryan et al., 2009; Murphy et al., 2007; Thoman, Smith, Brown, Chase, & Lee, 2013; Walton & Cohen, 2007). According to Murphy and Taylor (2012), environmental cues can trigger social identity threat (Steele, Spencer, & Aronson, 2002), which signals to women that they do not belong in STEM fields and that their contributions to STEM may not be valued. Recent research suggests that low levels of belonging in STEM, combined with competing experiences of belonging in non-STEM fields, predict less interest in STEM, particularly among women motivated to preserve their self-esteem (Thoman, Arizaga, Smith, Story, & Soncuya, 2014).

The Current Study

In two studies, we examined the influence of perceiving a negative environment for women on SSAB and women's STEM outcomes by looking at both self-reports of negative environments (Study 1) and manipulating an STEM environment to be positive or negative for women (Study 2). If gender-STEM stereotypes are activated by being in an environment perceived to be negative for women, and these stereotypes exert a subtle influence on attributional processing of stereotype-relevant events, SSAB may be predicted by negative environments for women.

Because these studies introduce a new measure of SSAB, we also included an assessment of academic attributional style that has been used in previous research, the Academic Attributional Style Questionnaire (AASQ; Peterson & Barrett, 1987). The SSAB measure differs from the AASQ in several ways. First, the SSAB measure is specific to STEM events happening to men and women (see Table 1 for the SSAB measure), whereas the AASQ assesses general, non-domain specific academic outcomes happening to one's self (e.g., "You are placed on academic probation"). Second, the SSAB measure does not directly ask participants to provide an attribution for events, whereas the AASQ does, making the SSAB measure a more subtle and less deliberative measure. Asking perceivers directly to explain why an event happened may produce more controlled attributional processing, which may allow people to become aware of the implication of their explanations for stereotype-relevant behaviors, and choose explanations that cast themselves in a positive light (Gilbert, Krull, & Pelham, 1988; Gilbert, Pelham, & Krull, 1988; Taylor, Wright, Moghaddam, & Lalonde, 1990). In light of these differences, we predicted that perceived negative environments for women would be more strongly predictive of SSAB than the measure of academic attributional style.

Because engaging in SSAB may serve to support and maintain gender-STEM stereotypes that target women as poor performers, we also predicted that SSAB would emerge as a stronger predictor of STEM outcomes for women compared to men. Finally, SSAB was assessed in response to men's and women's STEM successes as well as setbacks, in order to test the possibility that SSAB specific to negative STEM outcomes is particularly meaningful.

Table 1. STEM Stereotypic Attribution Bias Measure.

	Sentence Beginning	Sentence Type
1.	Kendra took notes during the physics lecture …	Filler
2.	Adam(Aisha) was accepted into the engineering competition …	Success
3.	Jill(Justin) "choked" during an oral presentation of her(his) science project …	Setback
4.	Aisha(Adam) designed the winning science experiment …	Success
5.	Tom(Helen) submitted his(her) chemical reactions paper for publication …	Success
6.	Justin(Jill) dropped out of his math study group …	Setback
7.	Matthew(Kelly) did not make the cutoff score for the math competition …	Setback
8.	Helen(Tom) was the only student to solve the math problem …	Success
9.	Denzel highlighted several sections in the readings …	Filler
10.	Kelly(Matthew) studied the wrong chapters for the physics test …	Setback
11.	Deborah(Brian) was named captain of the engineering competition team	Success
12.	Brian(Deborah) fixed the nagging problem with the experiment …	Success
13.	Ethan(Brittany) was cut in the first round of the science competition …	Setback
14.	Brittany(Ethan) was told her(his) research project was a failure …	Setback
15.	Jade logged on to the course website …	Filler
16.	Ron(Laura) received the highest math SAT score in the class	Success
17.	Laura(Ron) was chosen for the science scholarship …	Success
18.	Lisa(Steve) accidentally deleted the critical data spreadsheets …	Setback
19.	Steve(Lisa) was "let go" from his(her) research assistant job …	Setback
20.	Omar sharpened his pencil …	Filler

Note. Participants completed one of the two actor name-counterbalanced orders. The instructions were "Complete the sentence: In this section you will see the beginnings of sentences. Your job is to add words to each sentence beginning to form a complete sentence. You can add any words you like to the end of each sentence beginning, as long as it forms a grammatically correct sentence. Please type in the FIRST response that comes to your mind. Do not think of a lot of possibilities, then choose one of them to type in. Instead, be sure to type in the FIRST response that comes to you."

Hypotheses

We tested four hypotheses across the two studies:

> **Hypothesis 1:** Undergraduate STEM majors will evince SSAB by spontaneously attributing women's STEM setbacks to internal causes (e.g., inherent ability) and men's STEM setbacks to external causes (e.g., unfortunate circumstances), whereas women's STEM successes will be attributed to external causes and men's STEM successes to internal causes. SSAB will be primarily reflected in patterns of attributions for STEM setbacks, as opposed to STEM successes, due to the diagnostic nature of negative behaviors (Studies 1 and 2).

> **Hypothesis 2:** Perceived negative environments for women as self-reported (Study 1), or manipulated (Study 2), will predict SSAB only among women (who are the targets of negative gender-STEM stereotypes). Whether perceived negative environments predicted academic attributional style will also be explored.

> **Hypothesis 3:** Among women, but not men, SSAB will be associated with both less perceived belonging in STEM and weaker postgraduate intentions to stay in STEM (Studies 1 and 2).

> **Hypothesis 4:** SSAB will mediate the detrimental effect of experiencing a negative environment for women on women's performance, perceived belonging, and postgraduate intentions in STEM (Studies 1 and 2).

Study 1

Study 1 was designed to assess SSAB among a sample of male and female science undergraduates, testing Hypotheses 1–4 regarding perceived negative environment for women and SSAB. The study employed a longitudinal design, assessing SSAB, academic attributional style, and perceived negative environment for women at two time points during an academic semester. We used a longitudinal design to assess changes in SSAB over time as correlates of changes in two STEM outcomes: perceived belonging in STEM and postgraduate intentions to stay in STEM. In addition, the longitudinal design tested whether environments that are initially perceived as negative for women predict subsequent SSAB.

Method

Participants

We used fliers and e-mails to recruit students majoring in STEM fields. The sample included 85 undergraduates (46 women) enrolled in math, engineering, physics, informatics, and chemistry courses. Of the full sample, 71 participated at both time points (43 women) and were included in the main analyses. The majority of the final sample identified as White/European American (86%) or Asian/Asian American (7%) with the remainder identifying with another racial/ethnic group(s). Participants' ages ranged from 18 to 24 years (M = 18.78, SD = 0.84). Participants were paid US$20 to participate in this two-part study. Given that we targeted an underrepresented group, sample size was determined by the number of participants we were able to recruit during one academic semester. In addition, post-hoc linear bivariate regression power analyses were conducted in G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) to ensure that we had adequate power to detect the hypothesized gender difference in the effect of negative environments for women on SSAB. Power was computed using a two-tailed test, with an error probability of .05, and the relevant standard deviations and *B*s from the results are reported below. The power analysis confirmed that we had adequate power to detect our effects (power > .71).

Procedure

Participants were told that the purpose of the research was to examine the academic experiences and attitudes of college students over the course of a semester. The first time point of the study was initiated in the 8th week of the term, after students had acclimated to their classes. Participants were e-mailed a link to an electronic survey where they provided consent and then completed all of the measures described below. The order in which participants viewed the questionnaires was counterbalanced so that half the participants completed the SSAB measure before the rest of the measures and half the participants completed it after the rest of the measures. Order of presentation did not interact with other variables in analyses, all Fs < 1.40, ps > .20. The second time point of the study was initiated in the 11th week of classes, 2 weeks before final exams. Time 2 was administered in a second electronic survey that was identical to the one used in Time 1. The order in which participants viewed the measures was counterbalanced in the same manner as Time 1: Each participant completed the measures in the same order as they did initially. After completing Time 2 of the study, participants were fully debriefed and compensated.

Measures

Adapted from methodology used in previous work on stereotypic explanatory bias (e.g., Sekaquaptewa et al., 2003), SSAB was measured using a sentence fragment completion task in which respondents were asked to provide grammatically correct endings to sentences that depict male and female actors experiencing STEM setbacks and successes (see Table 1). Respondents were asked to complete 20 sentence fragments with the first grammatically correct ending that comes to mind, and their responses were coded for attribution type. Sixteen sentences described either a STEM success or STEM setback, paired with either a female- or a male-identified name, so that there were four sentences depicting each type of name/behavior combination. The remaining four sentences described stereotype-neutral behaviors (fillers) with two sentences having a male-identified name and two sentences having a female-identified name. SSAB behaviors and names were pre-tested with an independent sample ($n = 36$) to identify behaviors that were rated as significantly more representative of STEM success and failure compared to neutral items, and names that were clearly male-identified versus female-identified, all ts(35) > 2.24, all ps < .04. Two versions of the SSAB were used; a given behavior was paired with a female name in one version and paired with a male name in the other. Examples of SSAB items are "Laura was chosen for the science scholarship ..." (a STEM success performed by a woman) and "Ethan was cut in the first round of the science competition ..." (a STEM setback performed by a man). In order to ensure that spontaneous attributions were being measured (which most accurately reflects the concept of SSAB as a spontaneous attribution bias influenced by gender-STEM stereotypes), participants were instructed to write their first response and to avoid thinking about various ways the sentence could be completed and choosing to write one of them. Past research indicates that attributions for behaviors follow a two-stage process, with initial, "gut reactions" likely to be dispositional or internal attributions, which in a second stage may be adjusted to account for situational influences (Gilbert, Krull, & Pelham, 1988; Weary, Vaughn, Stewart, & Edwards, 2006). Spontaneous, unprompted attributions may reflect this initial, less controlled stage of attributional processing that is typically associated with more implicit measures (Vargas, Sekaquaptewa, & von Hippel, 2007). Asking perceivers directly to explain why an event happened may reflect the influence of the second more controlled stage of attributional processing, which may allow for socially desirable responding (Gilbert, Pelham, & Krull, 1988).

In a second independent sample ($n = 42$), pilot testing for the instructions of the measure revealed that when participants completed the measure with these "first response" instructions they spent significantly fewer seconds, $t(41) = -2.91$, $p = .007$, $d = -.83$, completing the sentence stems ($M = 15.96$, $SD = 5.45$) than participants who didn't receive the first response instructions ($M = 26.78$, $SD = 16.82$). Participants with the first response instructions ($M = 5.33$, $SD = 1.82$) also took significantly fewer minutes, $t(41) = 26.67$, $p = .02$, $d = -.73$, to complete the entire measure, compared to participants who did not receive the instructions ($M = 8.65$, $SD = 5.92$). This was an average difference of about 11 seconds faster per sentence, and an average 3.32 minutes faster for the entire measure, providing support responses to the SSAB are more spontaneous. Levene's test for equality of variances was violated in the present analysis, Fs > 19.97, ps < .001, and the t statistic computed did not assume homogeneity of variance.

Because participants were not explicitly asked to explain the behaviors described in the beginnings of the SSAB sentences, their sentence completions could reflect an attribution for the behavior or simply a continuation of the sentence. Each SSAB sentence completion was coded by three independent raters (one male and two female) as to whether it provided an internal explanation (e.g., "Jill(Justin) 'choked' during an oral presentation of her(his) science project ... because she(he) wasn't good at science"), an external

Table 2. SSAB Subscale and *L* (Lambda) Scoring Procedure.

Internal Success Difference Subscore	External Success Difference Subscore	Internal Setback Difference Subscore	External Setback Difference Subscore
The # of internal explanations given for women' STEM successes MINUS the # of internal explanations for men's STEM successes	The # of external explanations given for women's STEM successes MINUS the # of external explanations for men's STEM successes	The # of internal explanations given for women's STEM setbacks MINUS the # of internal explanations for men's STEM setbacks	The # of external explanations given for women's STEM setbacks MINUS the # of external explanations for men's STEM setbacks
Contrast weight −1	Contrast weight +1	Contrast weight +1	Contrast weight −1

Note. The four SSAB subscores are multiplied by their contrast weights and summed to form an *L* score (Furr & Rosenthal, 2003). Higher *L* scores indicate greater SSAB across setbacks and successes. STEM = science, technology, engineering, and math; SSAB = STEM stereotypic attribution bias.

explanation (e.g., "... because she(he) didn't get enough sleep the night before"), or no explanation (e.g., "... and then went home") for the behavior in the sentence stem. The criterion for coding an attribution was that a participant's sentence completion began with a term such as "because" or "in order to," which indicated explanation of the event in the sentence beginning. Raters were trained in the characteristics of internal (reflective of dispositional traits or character, level of ability, and personality) and external attributions (referring to situational factors, circumstances, and luck) as described in attribution theory research (e.g., Weiner, 1979). Interrater reliability analyses indicated significant agreement across judges, all $\kappa > .63_{\text{Time 1}}$ and $\kappa > .44_{\text{Time 2}}$, all *p*s < .001; therefore, attribution ratings were averaged across all three judges. The average number of attributions across all 20 sentences (internal or external) at Time 1 was 5.30 ($SD = 3.71$) and at Time 2 was 5.99 ($SD = 3.94$). Three participants at Time 1 (4.2%) and two participants at Time 2 (2.8%) did not provide any attributions at all across the SSAB items; 95.8% of participants provided at least one attribution. Across the 16 stereotype-relevant items, $17\%_{\text{Time 1}}$ and $21\%_{\text{Time 2}}$ were completed with internal attributions, $20\%_{\text{Time 1}}$ and $19\%_{\text{Time 2}}$ were completed with external attributions, with the remainder being completed with no attribution, $63\%_{\text{Time 1}}$ and $60\%_{\text{Time 2}}$. The fewer attributions compared to non-attributions, the participants provided likely emerged because participants were not asked to explain the behavior, only to complete the sentence in any grammatically correct way, in order to capture spontaneous attributions.

In order to test the predicted relation between perceived negative environment for women and SSAB, we computed a single SSAB index for each participant. We first computed difference scores, reflecting the degree to which a participant engaged in stereotype-biased attribution. We then computed a single index of the degree to which a participant's difference scores followed the predicted SSAB pattern, as described in the section on *L* (*lambda*) scores (Furr & Rosenthal, 2003).

Four difference subscores were calculated from the responses coded as attributions. Because there were 4 SSAB items of each type (men's successes, men's setbacks, women's successes, and women's setbacks), the total number of attributions given within each of the 4 item types could range from 0 to 4 (responses that were not judged as attributions were assigned scores of 0). We subtracted the total number of internal (external) attributions given for each of the 4 SSAB item types when they were performed by men, from the total when the behaviors were performed by women (see Table 2). This scoring procedure held constant behavior type (success or setback) and attribution type (internal or external), controlling for the possibility that more attributions may be generated for success or setbacks overall or that more internal or external attributions may be generated overall. Results of the omnibus analysis of variance (ANOVA) revealed a marginally significant main effect of attribution type, $F(1, 69) = 3.25$, $p = .076$, in addition to the reported main effect of behavior type reported below. No main effect of actor gender emerged, $F(1, 69) = 0.193$, $p = .662$. This scoring procedure allowed us to focus on the influence of actor gender and the STEM stereotypes associated with gender.

The first of these four difference subscores, the internal success difference, was obtained by subtracting the number of internal attributions given for STEM successes of men from the number of internal attributions given for STEM successes of women. Internal success difference scores less than 0 would indicate SSAB, or attributing STEM successes to internal (ability) causes when they are performed by men, more than when successful behaviors are performed by women.

The second score, external success difference, was the number of external attributions given for STEM successes of men subtracted from the number of external attributions given for STEM successes of women. External success difference scores greater than 0 would indicate SSAB, or attributing STEM successful behaviors to external (situational) causes when they are performed by women more than when the successful behaviors are performed by men.

The third score, named internal setback difference, was obtained by subtracting the number of internal attributions

Table 3. Study 1 SSAB Actor Gender Effects by Time.

		Female Actor	Male Actor	Simple Effects Contrast	
Time	Attribution Provided	*M* (*SD*)	*M* (*SD*)	$F(1, 69)$	*p*
Time 1	Internal attributions for STEM successes	.61 (0.84)	.70 (0.93)	0.65	.424
	External attributions for STEM successes	.49 (0.59)	.47 (0.59)	0.04	.839
	Internal attributions for STEM setbacks	.76 (0.76)	.62 (0.76)	2.23	.140
	External attributions for STEM setbacks	.71 (0.76)	.86 (1.01)	1.12	.294
Time 2	Internal attributions for STEM successes	.98 (1.01)	.88 (0.93)	0.76	.386
	External attributions for STEM successes	.46 (0.51)	.41 (0.51)	0.76	.386
	Internal attributions for STEM setbacks	.64 (0.67)	.83 (0.76)	3.99	.050
	External attributions for STEM setbacks	.84 (0.93)	.88 (0.93)	0.08	.785

Note. STEM = science, technology, engineering, and math; SSAB = STEM stereotypic attribution bias.

given for STEM setbacks of men from the number of internal attributions for STEM setbacks of women. Internal setback difference scores greater than 0 would indicate SSAB, or attributing STEM setbacks to internal (ability) causes when the behaviors are performed by women more than when they are performed by men.

The fourth score, external setback difference, was obtained by subtracting the number of external attributions given for STEM setbacks of men from the number of external attributions for STEM setbacks of women. External setback difference scores less than 0 would indicate SSAB, or attributing STEM setbacks to external (situational) causes when the behaviors are performed by men more than when they are performed by women.

SSAB difference subscores reflect the magnitude of the difference in number of internal or external attributions provided for men's and women's STEM successes and setbacks, independent of one's individual tendency to provide many or few attributions. For example, a participant who provided no external attributions for a woman's setbacks but provided two external attributions for a man's setbacks would receive an external setback difference of −2, as would a participant who provided two external attributions for a woman's setbacks but provided four external attributions for a man's setbacks. In either case, the participants are providing two more external attributions for men's STEM setbacks than they are for women's STEM setbacks.

Next, we computed a single SSAB index for each participant, reflecting the degree to which the participant's four SSAB subscores fit the predicted SSAB pattern. To accomplish this, *L* (*lambda*) scores were computed from contrast weights, as recommended by Furr and Rosenthal (2003). Following their procedure, each participant's SSAB difference scores were assigned weights reflecting the SSAB pattern (see Table 3). Multiplying the assigned contrast weights by the participant's SSAB differences scores, and summing the products, produced an *L* score. A positive *L* score indicated the degree to which participants showed the SSAB pattern of attribution bias. Negative *L* scores indicated that the participant showed the opposite of the SSAB pattern.

The AASQ (Peterson & Barrett, 1987) presented respondents with 12 negative academic scenarios (e.g., "You get a D in a course required for your major") and asked them to provide an explanation (If this situation were to happen to you, what do you think would have caused it?). No positive academic scenarios are included in this measure. Participants then rated each provided explanation on the dimensions internal/external, global/specific, stable/unstable, using 7-point scales for each of the three dimensions, for example, internal/external: "Was the cause due to something about other people or circumstances (1) or something about you (7)?" Scores on each dimension were averaged to form one total academic attributional style score ($M_{\text{Time 1}} = 4.36$, $SD_{\text{Time 1}} = .58$; $M_{\text{Time 2}} = 4.33$, $SD_{\text{Time 2}} = .57$). Peterson and Barrett (1987) reported a reliability of $\alpha = .84$ in their initial testing of the measure; in the current work, the reliability was somewhat lower but acceptable, $\alpha_{\text{Time 1}} = .79$, $\alpha_{\text{Time 2}} = .70$. Reliabilities for the subscales were $\alpha_{\text{Time 1}} = .62$, $\alpha_{\text{Time 2}} = .62$ for internality; $\alpha_{\text{Time 1}} = .76$, $\alpha_{\text{Time 2}} = .78$ for the stability subscale; and $\alpha_{\text{Time 1}} = .75$, $\alpha_{\text{Time 2}} = .77$ for the global subscale.

Participants' perceptions of a negative environment for women in their field of study were measured using the average of 4 items rated on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), $M_{\text{Time 1}} = 2.22$, $SD_{\text{Time 1}} = 1.18$; $M_{\text{Time 2}} = 4.33$, $SD_{\text{Time 2}} = .57$. The 4 items used were developed by the authors and included, "In my field of study, hearing derogatory comments about women is common;" "In general, my field of study is not welcoming to women;" "During class periods in my field of study courses, I have heard derogatory comments about women;" and "During group assignments in my field of study courses, I have heard derogatory comments about women," $\alpha_{\text{Time 1}} = .76$, $\alpha_{\text{Time 2}} = .88$.

Participants' feelings of belonging in STEM were assessed using five statements adapted from Walton and Cohen's (2007) 17-item academic fit measure ($\alpha = .89$). Each statement was rated on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), $M_{\text{Time 1}} = 5.42$, $SD_{\text{Time 1}} = .98$; $M_{\text{Time 2}} = 4.45$, $SD_{\text{Time 2}} = .98$. The four statements were

"I feel accepted by other students in my field of study," "I feel accepted by the instructors in my field of study," "I feel like I have a lot in common with other students in my field of study," and "Being a student in my field of study is an important part of who I am"; $\alpha_{\text{Time 1}} = .78$, $\alpha_{\text{Time 2}} = .80$.

Participants' intentions after graduation were assessed with 8 items rated on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree)*, $M_{\text{Time 1}} = 5.41$, $SD_{\text{Time 1}} = 1.22$; $M_{\text{Time 2}} = 5.39$, $SD_{\text{Time 2}} = 1.23$. The 8 items were developed by the authors and included, "It is very likely that I will work in my field of study after graduation," "I intend to graduate with a degree in my field of study," "I plan to have a career in my field of study," "I don't think I will ever lose interest in my field of study," "I am pretty sure I will soon change my field of study (e.g., change to another major)" (reversed), "I regret choosing my field of study" (reversed), "I will probably choose to work in a different area than my field of study after graduation" (reversed), and "I will probably have to work in a different area than my field of study after graduation" (reversed), $\alpha_{\text{Time 1}} = .91$, $\alpha_{\text{Time 2}} = .91$.

Results

SSAB Analyses

The number of attributions (averaged across the three judges) for each SSAB sentence type (female actor/success, female actor/setback, male actor/success, and male actor/setback) at Time 1 and Time 2 were analyzed in a 2 (Time Point) $\times$ 2 (Actor Gender: Female/Male Actor) $\times$ 2 (Behavior Type: STEM Success/Setback) $\times$ 2 (Attribution Type: Internal/External) $\times$ 2 (Participant Gender) mixed model ANOVA, with the first four factors within subjects and the fifth factor between subjects. Support for SSAB would be indicated in the pattern of attribution means that emerges from a significant Actor Gender $\times$ Behavior Type $\times$ Attribution Type interaction, and higher order interactions would indicate whether SSAB differed at Time 1 versus Time 2 and/or by participant gender. Results of this omnibus ANOVA revealed several lower order main effects and interactions: a main effect of behavior type, $F(1, 69) = 12.84$, $p = .001$, $\eta^2 = .157$; a Time $\times$ Actor Gender $\times$ Behavior Type interaction, $F(1, 69) = 4.44$, $p = .039$, $\eta^2 = .060$; a Time $\times$ Actor Gender $\times$ Behavior Type $\times$ Participant Gender interaction, $F(1, 69) = 4.52$, $p = .037$, $\eta^2 = .061$; a Behavior Type $\times$ Attribution Type interaction, $F(1, 69) = 26.52$, $p = .000$, $\eta^2 = .278$; and a Time $\times$ Behavior Type $\times$ Attribution Type interaction, $F(1, 69) = 7.31$, $p = .009$, $\eta^2 = .096$. These lower order effects were qualified by a Time $\times$ Actor Gender $\times$ Behavior Type $\times$ Attribution Type interaction, $F(1, 69) = 7.113$, $p = .01$, $\eta^2 = .093$. This four-way interaction was not qualified by participant gender, $F(1, 69) = 1.12$, $p = .294$. No other significant effects emerged, all Fs < 3.35, all ps > .075. Means are presented in Table 3.

Simple interaction analyses indicated that the Actor Gender $\times$ Behavior Type $\times$ Attribution Type interaction was significant at Time 1, $F(1, 69) = 3.867$, $p = .050$, $\eta^2 = .053$, but not at Time 2, $F(1, 69) = 1.19$, $p = .279$. Inspection of the means indicated that attributions appeared to follow the predicted SSAB pattern at Time 1. Although the Time 1 SSAB attribution means follow this pattern, simple effects contrast analyses showed that these mean differences were not statistically significant (see Table 3). In addition, our analysis of SSAB using L scores (a single index computed from all four subscales) demonstrated a similar pattern of results. The overall mean of Time 1 L scores was positive, $M = 0.41$, $SD = 1.66$, and differed significantly from 0, $t(71) = 2.06$, $p = .043$, $d = .49$, providing additional evidence that participants showed the SSAB pattern of attribution at Time 1. The overall mean of Time 2 L scores was negative, $M = -0.13$, $SD = 1.40$, and did not differ significantly from 0, $t(71) = -.77$, $p = .44$, indicating that participants did not show the SSAB pattern of attribution at Time 2, consistent with the results of the omnibus ANOVA.

In order to test Hypothesis 1, that SSAB would be primarily reflected in attributions for STEM setbacks as opposed to STEM successes, L scores were computed and analyzed separately by behavior type (setbacks vs. successes). Specifically, L scores were computed for science setbacks by multiplying the contrast weights for the internal and external setback difference scores by each participant's internal and external setback differences scores, and summing the products; L scores were computed for science successes by multiplying the contrast weights for the internal and external success difference scores by each participant's internal and external success differences scores, and summing the products. SSAB L scores derived from science setbacks at Time 1 and Time 2 were correlated, $r(71) = .503$, $p = .000$, as were SSAB L scores derived from science successes at Time 1 and Time 2, $r(71) = .402$, $p = .001$.

The mean for L scores derived from science setbacks at Time 1 was positive, $M = 0.30$, $SD = 1.12$, and differed significantly from 0, $t(70) = 2.30$, $p = .02$, $d = .55$, indicating that participants showed SSAB in regard to science setbacks. In contrast, the means for L scores derived from science successes at Time 1 ($M = 0.10$, $SD = 1.26$) and from Time 2 setbacks ($M = -0.12$, $SD = 1.37$) and successes ($M = 0.004$, $SD = 1.12$) did not differ significantly from 0, Time 1 SSAB-successes, $t(70) = .69$, $p = .49$; Time 2 SSAB-setbacks, $t(70) = -.77$, $p = .45$; and Time 2 SSAB-successes, $t(70) = -.04$, $p = .97$. These results provided support for our first hypothesis that SSAB may emerge primarily in response to STEM setbacks. Therefore, ensuing analysis focused only on SSAB as assessed in response to the STEM setback items, not the STEM success items.

Although SSAB did not emerge at Time 2, the effect of time (i.e., comparisons between Time 1 and Time 2 variables) was retained in subsequent analyses to assess the effect of increases or decreases in these variables over time.

Table 4. Study 1 Bivariate Correlations Between Variable Change Scores.

	1	2	3	4	5
1. SSAB change	—	−.21	−.05	.09	−.06
2. Academic attributional style change	−.15	—	.39**	−.44**	−.18
3. Perceived negative environment for women change	.14	.14	—	−.08	−.03
4. Belonging in STEM change	−.37***	.06	.05	—	.05
5. Postgraduate intentions change	−.25*	−.09	.05	.30**	—

Note. Women's data are presented below and men's data presented above the diagonal. Variables represent Time 2 − Time 1 difference scores. STEM = science, technology, engineering, and math; SSAB = STEM stereotypic attribution bias.
*$p \leq .10$. **$p \leq .05$. ***$p \leq .01$.

Perceived Negative Environment for Women as a Predictor of SSAB

We predicted that perceived negative environment for women would predict SSAB among women (Hypothesis 2). Participant gender, perceived negative environment for women, and their interaction were tested as predictors of Time 1 and Time 2 SSAB, using regression analyses. Participant gender was dummy coded as 0 = *female*, 1 = *male*, and the continuous variable perceived negative environment for women was mean centered (i.e., standardized) prior to computing the interaction term. In analyses of variables collected at Time 1, participant gender, $B = -.04$, $t = -0.37$, $p = .71$, and Time 1 perceived negative environment for women, $B = .04$, $t = 0.31$, $p = .76$, did not emerge as significant predictors, whereas the Participant Gender × Perceived Negative Environment for Women interaction term emerged as a marginally significant predictor of Time 1 SSAB, $B = -.28$, $t = -1.84$, $p = .07$, *semi-partial* $r = -.22$. Simple slopes analyses revealed no significant effects; however, perceived negative environment for women was a relatively stronger predictor of Time 1 SSAB scores among women, $B = .49$, $t = 1.37$, $p = .17$, than among men, $B = .05$, $t = 0.31$, $p = .76$. In analyses of variables collected at Time 2, the interaction term also emerged as a marginally significant predictor of Time 2 SSAB, $B = -.28$, $t = -1.72$, $p = .091$, *semi-partial* $r = -.21$, whereas participant gender, $B = -.05$, $t = -0.44$, $p = .66$, and Time 2 perceived negative environment for women, $B = .06$, $t = 0.47$, $p = .64$, did not. Finally, Time 1 negative environment for women was tested as a predictor of Time 2 SSAB (controlling for Time 1 SSAB entered as the first step). Time 1 SSAB emerged as a significant predictor of Time 2 SSAB, $B = .480$, $t = 4.45$, $p = .001$, *semi-partial* $r = .48$, whereas participant gender, $B = -.06$, $t = -0.60$, $p = .55$, perceived negative environment for women, $B = .20$, $t = 1.145$, $p = .15$, and the participant gender by perceived negative environment for women interaction term did not, $B = -.18$, $t = -1.331$, $p = .19$. This suggested that the strong relation between Time 1 and Time 2 SSAB may have overshadowed any relation between perceived negative environment for women at Time 1 on SSAB at Time 2.

Perceived Negative Environment for Women as a Predictor of AASQ

In analyses similar to those of SSAB, participant gender, perceived negative environment for women, and their interaction did not emerge as a significant predictor of overall academic attributional style, all *B*s < .24, all *t*s < 1.54, all *p*s > .12, nor of the AASQ subscales (internal subscale, all *B*s < .16, all *t*s < .1.36, all *p*s > .18; stable subscale, all *B*s < .17, all *t*s < 1.01, all *p*s > .31; and global subscale, all *B*s < .16, all *t*s < .99, all *p*s > .32). These results provided support for Hypothesis 3 that perceived negative environment for women would predict SSAB but not academic attributional style as measured by AASQ.

Testing Effect of Changes in SSAB Over Time on STEM Outcomes

Change scores were computed for perceived negative environment for women, academic attributional style, SSAB *L* scores, and the two STEM outcomes (belonging in STEM and postgraduate intentions), by subtracting Time 1 scores from Time 2 scores. Positive scores indicated increases and negative scores indicated decreases in the variable from Time 1 to Time 2. In support of our fourth hypothesis, among women, correlational analyses showed that SSAB change scores were negatively correlated with changes in belonging in STEM, and marginally negatively correlated with postgraduate intentions, such that increases in SSAB scores were associated with decreases in the two STEM outcomes from Time 1 to Time 2. Among men, the correlations between SSAB change scores and changes in belonging in STEM and postgraduate intentions were not significant (see Table 4). SSAB change scores were not significantly correlated with perceived negative environment for women change scores, for either women or men, as would be predicted if perceived negative environments contribute to SSAB. However, this lack of relation may have also emerged due to lower overall SSAB scores at Time 2 (not significantly different from 0) reducing variance in SSAB change scores, making it difficult to draw conclusions about the relation between changes in SSAB over time and perceived negative environments.

Discussion

The results from Study 1 supported our four hypotheses and provided evidence for the existence of SSAB in response to men's and women's STEM setbacks. At Time 1 of the study, all participants regardless of gender showed the predicted SSAB pattern of spontaneously providing more internal explanations (e.g., low ability) for women's than men's STEM setbacks, and more external explanations (e.g., unfavorable circumstances) for men's than women's STEM setbacks. In line with our first hypothesis, the SSAB pattern did not reliably emerge regarding successes in STEM. Consistent with research on naive causal theories (Ybarra, 2002), this result may reflect that respondents perceived STEM setbacks to be more meaningful and ability-diagnostic than STEM successes.

Moreover, SSAB emerged only at Time 1 and appeared to diminish at Time 2. This may suggest that SSAB varies over time; however, it is also likely that the reduction in SSAB at Time 2 emerged because participants responded to the exact same sentence beginnings at Time 1 and Time 2 and may have engaged in more deliberative and controlled responding to the items at Time 2 because they had seen them before. Although this possibility cannot be tested with the existing data, in support of predictions, we found that perceiving a negative environment for women in STEM predicted greater SSAB within both Time 1 and Time 2, but only among women. These results indicate both men and women show SSAB, but perceiving a negative environment for women seems to be contributing to this bias primarily among women.

Perceived negative environment for women at Time 1 did not predict SSAB at Time 2 over and above the effect of Time 1 SSAB, suggesting that the strong relation between Time 1 and Time 2 SSAB eclipsed any relation between perceived negative environment for women across time points. However, correlational results indicated that increases in SSAB over time tended to be associated with diminishing levels of belonging in one's science field and lower intentions to stay in science among women, but not among men. These results provided preliminary evidence that SSAB influences women's STEM outcomes in particular.

Perceived negative environment for women did not predict academic attributional style as assessed by the AASQ (Peterson & Barrett, 1987). Negative environments for women might have less influence on more deliberative assessments of attribution because the influence of gender-STEM stereotypes activated by the setting might be more controllable on such measures. Moreover, the AASQ measures responses to general academic outcomes, whereas our measure focused on STEM-specific events, suggesting that perceived negative environments for women may have a stronger relation to attributions for STEM-specific events than for general academic events. Furthermore, the academic events presented in the AASQ are imagined to have occurred to one's self, whereas the SSAB measures responses to outcomes described as happening to other men and women. It could be the case that in settings where the implications of one's performance are highlighted (such as giving an STEM performance in a small group), deliberative attribution styles may also be influenced by perceived negative environments for women. To address such possibilities, we retained the AASQ in Study 2, which examined STEM outcomes in small groups that varied in perceived negativity towards women.

Study 2

In Study 1, perceived negative environment for women was tested in terms of blatant instances of negative treatment of women (e.g., hearing derogatory comments about women). However, negative treatment of women also takes more subtle forms characterized by microaggressions (Sue, 2010) and modern sexism (Swim & Cohen, 1997). Moreover, these subtle forms of negative treatment have been associated with negative outcomes for women in STEM (e.g., Camacho & Lord, 2011; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012) and women in the workplace generally (Basford, Offermann, & Behrend, 2014; Sojo, Wood, & Genat, 2015). Therefore, in Study 2, we examined perceived negative environments for women in terms of more subtle negative behaviors enacted towards women in an STEM setting.

In Study 2, we sought to replicate and extend the findings from Study 1 by manipulating an STEM setting to be either a positive or a negative environment. Under the guise of testing an audiovisual (AV) communication system, we were able to create a group setting in which female STEM students viewed other female STEM students (prerecorded confederates) speaking on a science topic and receiving negative nonverbal feedback from male group members. By manipulating the nonverbal feedback of the men towards the women in the group, we created an unwelcoming environment for women, defined by witnessing the negative treatment of other women in an STEM setting.

In line with our first three hypotheses and the results of Study 1, we predicted that participants who saw other women receive negative nonverbal feedback would show greater SSAB and more negative STEM-relevant outcomes (i.e., belonging and future intentions in STEM) than women who saw other women being treated positively. We also tested the possibility that women's STEM performance may be inhibited by viewing the negative treatment of women. Finally, we tested whether SSAB mediates the relation between witnessing the negative treatment of women and STEM-relevant outcomes (Hypothesis 4).

Method

Given the results from Study 1 showing that perceived negative environment for women in science predicted SSAB only among women and that women's (not men's) STEM outcomes were associated with increases in SSAB over time, Study 2 included women only. Participants were told that they would be testing an AV communication system, which allows remotely located students to communicate with each other, a procedure used by Sekaquaptewa and Thompson (2002, 2003). Participants were placed in private rooms and believed they could see and hear other participants located in separate rooms using the AV system. Participants saw prerecorded footage of other participants that appeared as though it were being broadcast live from the other rooms. Participants presented STEM material to their group on camera. Negative treatment of women was manipulated by having male actors provide negative nonverbal feedback (e.g., eye rolling and scoffing), or more positive feedback (e.g., nodding and smiling), when women in the group were speaking on camera. Finally, participants completed two STEM outcome measures (belonging and postgraduate intentions in STEM), a brief oral exam on the STEM material, demographic measures, and a funneled debriefing.

Participants

The original sample consisted of 98 undergraduate women. Data from 11 participants (7 from the negative treatment condition and 4 from the control condition) were excluded from analyses due to suspicion about the use of videotaped confederates reported during the debriefing. The final sample was 87 women majoring in STEM-related fields. We recruited women majoring in STEM fields through fliers and e-mailed announcements; participants were compensated with US$15.00. The sample ranged in age from 18 to 37 ($M = 19.94$, $SD = 2.34$). The majority of the sample identified as White/European American (50.6%) or Asian/Asian American (26.4%). Given that we were targeting an underrepresented group, our sample size was based on the number of participants we could recruit across two academic semesters. A post-hoc one-way ANOVA power analysis was conducted in G*Power 3.1 (Faul et al., 2009) to ensure that we had adequate power to detect the hypothesized condition difference in SSAB. Power was computed using an effect size of 0.5 and an error probability of .05. The power analysis confirmed that we had adequate power to detect our effects (power > .95).

Procedure

Prior to the experiment, five nonprofessional student actors were filmed and their footage was embedded in a video presentation that participants would view during the experiment. Participants were led to believe that a live, AV system was being used.

When participants arrived, they were oriented to an AV system and told the system was designed to examine video communication among college students. Participants were told they would view the other group members on one monitor and themselves on another monitor and that when they spoke, they would not see the group members, but that the group members would still be able to see them. Participants were told this so that when speaking, they would look directly at the camera and would not be distracted by looking at another monitor showing footage of other group members. In addition, when the participant spoke the camera was turned off, which isolated the effects of a negative environment for women from the effects of participants directly receiving negative feedback themselves: They saw the behaviors of men directed towards the other women in the group, not towards themselves.

The experimenter, located in another room, controlled both of the monitors in the participant's cubicle and used a headset to communicate with the participant throughout the experiment. The computer program launched by the experimenter first explained the cover story and all phases of the experiment. Then, participants were given 5 minutes to read a short passage of science information and were told that they would have to summarize the passage to their group members later.

Next, participants removed a covering from the camera lens and watched as the experimenter asked everyone in the group, including the participant, if each person could see and hear everyone. The participant's group consisted of five others: either three men and two women or four men and one woman (not including the participant). Although we predicted that a group with fewer women would be perceived as more unwelcoming to women, no significant differences emerged by gender composition of the group, indicating that the effect of seeing one versus two women in the group was negligible and/or overshadowed by the negative treatment of women.

Participants were then told that it was time for everyone to summarize for the group the science passage they had read and were reminded that at the end of the experiment there would be an oral exam covering the passage summaries given by each group member. Participants first watched, as four of the group members summarized their own passages. During these summaries, participants were introduced to the negative treatment or control manipulation.

Participants in the negative treatment condition viewed male actors displaying negative nonverbal feedback (e.g., frowning, scoffing, crossing their arms, rolling their eyes, and looking disinterested) towards the female group members but positive nonverbal feedback (e.g., nodding, smiling, and looking interested) towards male group members. Participants in the control condition viewed male actors displaying the same positive nonverbal feedback to all group members, including women. Having only the male actors enact the manipulated behaviors towards female members of the group was intended to heighten perceptions that the behavior was gender-based in the negative treatment condition (as opposed to negative behaviors enacted towards everyone regardless of gender).

In addition, in the negative treatment condition the experimenter did not ask any of the female group members a follow-up question but did ask all the male group members a follow-up question. The experimenter also gave praise to the male group members after their summaries saying, "good job" or "well done," but did not do the same for the female group members. In the control condition, all group members were asked a follow-up question and given praise regardless of gender. Within the positive and negative treatment conditions, half the participants had a female and half had a male experimenter. Results did not significantly differ by experimenter gender. Next, the experimenter deactivated the AV system and participants were asked to rate the ease of using the AV system and to complete assessments of their academic experiences and attitudes. They then completed the SSAB measure, measures of belonging in the field, postgraduate intentions, and a manipulation check. The experimenter then reactivated the AV system for the oral exam. The participant was selected as the first group member to be called on to complete the oral exam on the passage that they had previously read.

The video recorded oral exam contained seven questions about the passage assigned to the participant, five multiple-choice questions and two open-ended questions. All seven questions were answered by the participant aloud, and they also wrote down their responses to the multiple choice items on an answer sheet. After answering the questions, the experimenter abruptly stopped, stating that he or she forgot to have the participant complete a questionnaire prior to the exam. The participant completed the questionnaire which included the measure of academic attributional style and demographic questions. Next, the experimenter told participants the experiment was over due to time constraints and no one else would be taking the oral exam. Participants were probed for suspicion, fully debriefed about the true nature of the study, and compensated. Overall, the debriefing interview indicated that participants largely believed the cover story and did not think their group members were actors.

Measures

SSAB was measured in the same way as Study 1. STEM successes assessed whether SSAB would again emerge primarily in response to STEM setbacks, or whether SSAB in response to positive STEM behaviors might emerge in this experimental context. Similar to Study 1, each sentence completion was coded by independent raters as to whether it provided an internal explanation, an external explanation, or no explanation for the behavior depicted in the sentence beginning. Because we were coding the SSAB measure for the first time in Study 1, we used three raters in order to develop a concise coding manual. Having established the coding procedure, we used only two raters (one male and one female) in Study 2. Across all sentences, raters agreed an average of 82.2% of the time and attribution scores were averaged across raters as in Study 1. The average number of attributions (internal or external) was 5.01 ($SD = 3.19$), and three participants (3.4%) did not provide any attributions across the SSAB items. Across the 16 stereotype-relevant items, 22% were completed with internal attributions, 9% were completed with external attributions, and 69% were completed with no attribution (*Mean* $\kappa = .68$, $ps < .001$).

As in Study 1, the AASQ was administered ($M = 4.46$, $SD = .57$). AASQ subscale scores were averaged to form one academic attributional style score, $\alpha = .74$. Reliabilities for the subscales were $\alpha = .43$ for the internality subscale, $\alpha = .80$ for the stability subscale, and $\alpha = .75$ for the global subscale. In order to address the low reliability of the internality subscale, an analysis of the reliability of the full scale (with all three subscales) was performed. Results revealed that three of the academic scenarios internality ratings did not contribute significantly to the scale; however, dropping these 3 items from our analysis did not alter our results, so for the sake of consistency all items were retained in the analyses presented.

All participants read a GRE-based essay on a science topic (solar activity) at the beginning of the experiment. Performance was measured using a seven-question oral exam developed by the authors that had five multiple-choice questions worth 1 point each. Participants gave answers aloud as well as written on paper. Two open-ended questions were also included ("What are some of the documented effects of solar activity on Earth and its inhabitants?" and "Name one positive outcome associated with sunspots"). Two independent coders rated each participant's videotaped oral exam performance on the open-ended questions, giving 1 point for each correct detail given in the answer, of a total of 7 points possible across the two questions (e.g., "dangerous radiation levels" worth 1 point). There was sufficient reliability among coders (Question 1 $\kappa = .41$, Question 2 $\kappa = .77$, $ps < .001$), so their scores were averaged to form an open-ended question performance score ($M = 3.25$, $SD = 1.20$).

Participants were first asked to identify their field of study (i.e., their major), and their feelings of belonging in their field of study, using 3 items adapted from Walton and Cohen (2007). We used fewer items in Study 2 to save time. Items (e.g., "I feel accepted by other students in my field of study") were rated on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), $M = 5.11$, $SD = 1.10$, $\alpha = .83$.

We assessed participants' intentions to stay in their field of study after graduation with 2 items rated on a 7-point scale, ranging from 1 (*not at all likely*) to 7 (*very likely*), $M = 5.43$, $SD = 1.50$. The 2 items used (fewer than in Study 1 to save time) were "How likely is it that you will pursue graduate study related to your field of study?" and "How likely is it that your eventual career after graduation will directly pertain to your field of study?" ($r = .56$, $p < .001$).

In order to assess whether the video manipulation did indeed lead women to perceive the group setting as negative, participants' perception of the group depicted in the video was assessed with 2 items rated on a 10-point scale, ranging from 1 (*strongly disagree*) to 10 (*strongly agree*). The 2 items

Table 5. Study 2 SSAB Actor Gender Effects by Condition.

		Female Actor	Male Actor	Simple Effects Contrast	
Condition	Attribution Type Provided for SSAB Item Type	*M* (*SD*)	*M* (*SD*)	*F*(1, 69)	*p*
Negative treatment	Internal attributions for STEM successes	0.83 (0.97)	0.90 (0.99)	2.77	.600
	External attributions for STEM successes	0.12 (0.34)	0.15 (0.36)	0.25	.619
	Internal attributions for STEM setbacks	1.12 (0.97)	0.79 (0.95)	5.87	.018
	External attributions for STEM setbacks	0.55 (0.69)	0.73 (0.81)	1.70	.196
Control	Internal attributions for STEM successes	0.82 (0.89)	0.65 (0.68)	1.24	.268
	External attributions for STEM successes	0.12 (0.29)	0.21 (0.35)	0.48	.220
	Internal attributions for STEM setbacks	0.96 (0.95)	1.09 (1.05)	1.53	.490
	External attributions for STEM setbacks	0.49 (0.61)	0.39 (0.55)	0.48	.492

Note. STEM = science, technology, engineering, and math; SSAB = STEM stereotypic attribution bias.

used were “The other group members probably expect me to answer my questions quite well” ($M = 4.79$, $SD = 1.90$) and “I feel like I fit in well with the rest of the group” ($M = 4.84$, $SD = 2.20$).

Results

Manipulation Check

Responses to the 2 manipulation check items were only weakly correlated ($r = .25$, $p = .02$), so each item was analyzed separately in a one-way ANOVA, comparing the negative treatment versus control conditions. Results revealed a marginal effect of condition, $F(1, 78) = 3.71$, $p = .058$, $\eta^2 = .045$, for the statement “The other group members probably expect me to answer my questions quite well” and a marginal main effect of condition, $F(1, 77) = 3.61$, $p = .061$, $\eta^2 = .045$, for the statement “I feel like I fit in with the rest of the group.” In both cases, women in the negative treatment condition ($M = 4.40$, $SD = 2.16$ and $M = 4.39$, $SD = 2.22$, respectively) felt less positive about their group than those in the control condition ($M = 5.21$, $SD = 1.47$ and $M = 5.32$, $SD = 2.09$, respectively).

SSAB Analyses

The number of attributions (averaged across both judges) for each SSAB sentence type (female actor/success, female actor/setback, male actor/success, and male actor/setback) was analyzed in a 2 (Actor Gender: Female/Male Actor) × 2 (Behavior Type: STEM Success/Setback) × 2 (Attribution Type: Internal/External) × 2 (Condition: Negative Treatment vs. Control) mixed model ANOVA, with the first three factors within subjects and the fourth factor between subjects. Results of this omnibus ANOVA revealed a main effect of behavior type, $F(1, 85) = 47.22$, $p = .001$, $\eta^2 = .357$; a main effect of attribution type, $F(1, 85) = 55.40$, $p = .001$, $\eta^2 = .395$, and the predicted Actor Gender × Behavior Type × Attribution Type × Condition interaction, $F(1, 85) = 9.25$, $p = .003$, $\eta^2 = .098$. No other significant effects emerged, all Fs < 3.51, all ps > .06. Means are presented in Table 5.

Simple interaction analyses indicated that the Actor Gender × Behavior Type × Attribution Type interaction was significant in the negative treatment condition, $F(1, 85) = 6.28$, $p = .014$, $\eta^2 = .067$, and marginally significant in the control condition, $F(1, 85) = 3.30$, $p = .073$. Inspection of the means indicated that attributions generally followed the predicted SSAB pattern in the negative treatment condition, but in the control condition, the means tended to be in a pattern opposing SSAB. Although the SSAB attribution means in the negative treatment condition follow the predicted pattern, simple effects contrast analyses showed that only one of the mean differences was statistically significant (see Table 5).

L (Lambda) Scores

As in Study 1, overall L scores (derived from both setbacks and successes) were computed from contrast weights, in order to produce a single index of SSAB. Overall L scores ($M = 0.10$, $SD = 1.76$) were analyzed with one-way ANOVA comparing the negative treatment versus control conditions. Results revealed a significant effect of condition, $F(1, 85) = 9.25$, $p < .01$, $\eta^2 = .098$. Women in the negative treatment condition showed greater overall SSAB ($M = 0.62$, $SD = 1.98$) than women in the control condition ($M = -0.48$, $SD = 1.25$), supporting the results of the omnibus ANOVA. Thus, the pattern of overall L scores followed predictions.

As in Study 1, SSAB subscale scores were examined separately for setbacks versus successes. Consistent with Study 1, results supported our first hypothesis and showed that the effect of negative treatment was most apparent for SSAB derived from STEM setbacks. Analysis of L scores specific to setbacks revealed a significant effect of negative treatment condition, $F(1, 85) = 6.90$, $p = .01$, $\eta^2 = .075$, such that women in the negative treatment condition ($M = 0.58$, $SD = 1.40$) showed more setback-specific SSAB than women in the control condition ($M = -0.22$, $SD = 1.42$). Parallel analysis of L scores specific to successes revealed no significant difference, $F(1, 85) = 1.56$, $p = .22$, between women in the negative treatment condition ($M = 0.04$, $SD = 1.25$) and

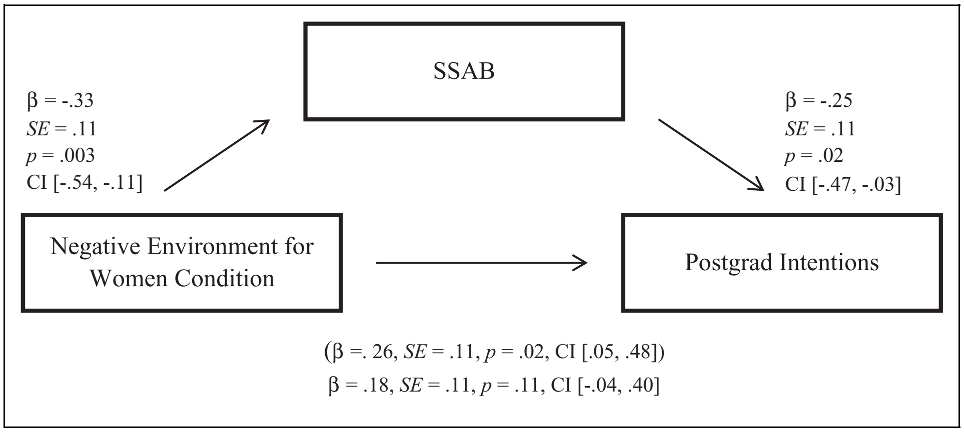

Figure 1. Perceptions of a negative environment for women condition (1 = *negative*, 2 = *control*) affected postgraduate intentions mediated through science, technology, engineering, and math stereotypic attribution bias (SSAB). Higher SSAB and postgraduate intention scores reflect greater SSAB and postgraduate intentions. Represented path coefficients are standardized *B*s. The indirect X–Y path is presented below the direct X–Y path.

women in the control condition ($M = -0.26$, $SD = 0.94$). Therefore, further analyses focused on SSAB specific to setbacks only.

Academic Attributional Style

AASQ scores were analyzed in one-way ANOVA, comparing the negative treatment ($M = 4.52$, $SD = 0.57$) versus control conditions ($M = 4.39$, $SD = 0.57$). Results revealed no significant effect of condition on overall academic attributional style scores, $F(1, 85) = 1.09$, $p = .30$, nor on the three academic attributional style subscale scores (internal, global, and stable; all Fs < 1.60, all ps > .20).

Oral Exam Performance

Malfunctions with the computerized video and/or audio recording of the participant's oral exam performance prevented usable recording for 28 participants. Participants wrote down their answers to the multiple choice questions, and these responses were available for use. However, no responses to the open-ended questions were available for the 28 participants whose video recording malfunctioned. Because this represents a loss of 32% of the sample, analyses of responses to the open-ended questions suffered from low statistical power. Nevertheless, to test part of our third hypothesis, both the multiple choice ($N = 87$) and oral exam performance scores ($N = 59$) were analyzed in two separate one-way ANOVAs. Results showed no significant effect of the negative treatment ($M = 3.50$, $SD = 1.91$) verses control condition ($M = 2.09$, $SD = 1.82$) on multiple choice items, $F(1, 84) = 0.84$, $p = .36$. Women in the negative treatment condition ($M = 3.50$, $SD = 2.05$) also did not perform significantly different from women in the control condition ($M = 2.94$, $SD = 1.89$) on open-ended items, $F(1, 57) = 1.15$, $p = .29$. Scores on the multiple choice and open-ended items were uncorrelated, $r(59) = .05$, $p = .72$.

Belonging in Field of Study

Belonging in field of study scores were analyzed in one-way ANOVA, comparing the negative treatment ($M = 5.32$, $SD = 1.21$) versus control conditions ($M = 5.61$, $SD = 1.03$). In contrast to the results of Study 1, results revealed no significant effect of condition, $F(1, 81) = 1.39$, $p = .24$.

Postgraduate Intentions in One's Field of Study

To test postgraduate intentions, scores were analyzed in a one-way ANOVA, comparing the negative treatment ($M = 5.06$, $SD = 1.74$) versus the control condition ($M = 5.84$, $SD = 1.03$). Results revealed a significant effect of condition, $F(1, 79) = 5.88$, $p < .02$, $\eta^2 = .069$, such that women who witnessed the negative treatment of women in their group had weaker postgraduate intentions in their field of study than those who did not witness negative treatment of women.

Testing Models of Mediation

A mediation model was analyzed to assess whether the direct effect of negative treatment condition on postgraduate intentions was mediated by SSAB (*L* scores derived from STEM setbacks). Following the procedures described by Hayes (2013, PROCESS Model 4), we assessed whether the effect of condition on this outcome ($B = .26$, $SE = 0.11$, $p = .02$) was weakened when SSAB was taken into account. Results showed that when SSAB was included in the model, the effect of condition on postgraduate intentions decreased ($B = .18$, $SE = 0.11$, $p = .11$; see Figure 1). Bootstrapping confidence intervals tested whether the indirect effect was significantly different from zero, using a 95% bias-corrected bootstrap confidence interval based on a sample of 10,000 iterations (Hayes, 2013). In regards to the mediation effect size, we followed the recommendation of Preacher and Kelly (2011), and report more than one indicator of effect size. One such indicator is the effect size index k^2, defined as the proportion of the maximum possible indirect effect that could have occurred to the indirect effect obtained. This index has the desired qualities of being scaled on a meaningful metric (0–1), amenable to the construct of confidence intervals, and largely independent of sample size (Preacher & Kelley, 2011). Results of the k^2 test in the bootstrapping analyses suggested that SSAB significantly mediated the relation between condition and postgraduate intentions, $k^2 = .08$, $SE = 0.05$, CI [0.003, 0.15], indicating a small effect size.

Other researchers have pointed out mathematical problems with k^2, and recommended an alternative P_M, defined as the ratio of the indirect effect to the total effect (Wen & Fan, 2015). However, P_M was found to be unstable in simulation sample sizes of less than 500 (MacKinnon, Warsi, & Dwyer, 1995). The P_M in the current mediation analysis was nonsignificant, $P_M = .31$, $SE = .73$, CI [−0.17, 1.58]. This may have been due to our relatively small sample size; or perhaps more likely, because the mediating role of SSAB is modest. In addition to the influence of SSAB, other unmeasured mechanisms may contribute to the effects of the negative treatment of women manipulation on this outcome.

Results of analyses testing an alternative model (in which postgraduate intentions are tested as a mediator of the effect of condition on SSAB) showed inferior support for the alternative model, $k^2 = .03$, $SE = 0.03$, CI [0.001, 0.12]; $P_M = .14$, $SE = 1.54$, CI [−0.08, 0.90], primarily because postgraduate intentions did not predict SSAB ($B = -.07$, $SE = 0.10$, $p = .27$). Overall, the results of the mediation model testing suggests a small but perhaps significant role of SSAB in explaining the effect of witnessing the negative treatment of women in an STEM setting on lowering women's postgraduate intentions to continue in STEM.

Discussion

Study 2 extended the correlational results of Study 1 by providing an experimental manipulation of experiencing a negative environment for women in STEM. In Study 2, female science majors were randomly assigned to experience a negative environment for women (operationalized as witnessing the negative treatment of women in a science setting). As in Study 1, SSAB emerged in response to STEM setbacks, but not successes, and was more influenced by experiencing a negative environment for women; consequently, we focused on SSAB in response to STEM setbacks only. Women who experienced a negative environment for other women showed greater SSAB than those assigned to a control condition. In addition, experiencing a negative environment for women was associated with lower intentions to continue in STEM after graduation, an effect that was predicted to be mediated by SSAB. However, mediation tests were only suggestive, as the mediation effect size was significant regarding one index (k^2) but not an alternative index (P_M). By either index, the role of SSAB as a mediator of the effect of negative environment for women on women's postgraduate intentions in STEM was small. This finding suggests that being in an STEM environment that is negative for women may lead women to feel less likely to continue in STEM via several psychological mechanisms. One potential explanation is found in the increase in stereotypically biased attributions for men's and women's STEM setbacks. The stereotypically biased attribution pattern may support gender-STEM stereotypes, which then have a negative influence on women's STEM perceptions and motivation.

Although experiencing a negative environment for women significantly influenced women's perceptions of their experimental STEM group setting as negative, it did not have a significant effect on feelings of belonging in the field generally. Feelings of belonging in one's field of study (i.e., STEM) is an important factor in predicting women's success in STEM fields (e.g., Dasgupta, 2011; Good, Rattan, & Dweck, 2012; Smith, Lewis, Hawthorne, & Hodges, 2013), and previous research has linked this type of belonging to the experience of stereotype threat (e.g., Dasgupta, 2011; Schmader, Johns, & Forbes, 2008). That feelings of belonging in the field were not affected in this study might mean that the relevance of stereotypes triggered by witnessing the negative treatment of women in the group does not transfer to the broader context of the field. On the other hand, it may be that many, if not most, women in the experiment were majoring in an STEM field that differed from the STEM domain on which they were tested in the study (solar activity). Low perceived similarity to the woman in the video could diminish the perceived implications for their own performance in the study, and their belonging in their own STEM fields of study (Lockwood & Kunda, 1997; Marx & Ko, 2012). Finally, as in Study 1, experiencing a negative environment for women in STEM was not associated with negative academic attributional style, as measured by the AASQ (Peterson & Barrett, 1987).

General Discussion

Results from the two studies suggest that being in STEM environments perceived as negative for women can bias attributions made for men's and women's STEM performances. We defined SSAB as the spontaneous attribution of women's STEM setbacks to internal causes (e.g., low ability) and men's STEM setbacks to external causes (e.g., bad luck). Perceived negative environments for women, both self-reported (Study 1) and experimentally induced (Study 2) were associated with greater SSAB, and greater SSAB was associated with lower intentions to continue in STEM after graduation (Studies 1 and 2).

Consistent with our first hypothesis, we found evidence for the predicted SSAB pattern. The finding that SSAB emerged primarily in response to negative events (STEM setbacks) as opposed to positive events (STEM successes) is consistent with work on naive causal theory (Ybarra, 2002). Naive causal theory suggests that we find negative behaviors particularly meaningful and diagnostic of what people are really like, because they are non-normative and violate our general expectations for positive behaviors from others (Fiske, 1980; Kanouse & Hanson, 1972). It may be the case that individuals perceive greater stereotype-incongruency in men's STEM setbacks, and greater stereotype-congruency in women's STEM setbacks, which leads to greater stereotypically biased attributions for these events.

The finding that outcomes were predicted by one form of SSAB (derived from setbacks) and not another form (derived from successes) is also aligned with previous research using a measure of SSAB assessing race stereotypic explanatory bias. Sekaquaptewa, Espinoza, Thompson, Vargas, and von Hippel (2003) demonstrated that some subscales of the measure (those based on external attributions) showed greater predictive utility than other subscales. Although these studies focused on stereotypic attribution patterns involving race stereotypes and not gender-STEM stereotypes, and an interracial social context, not an STEM performance context, in general, it appears that patterns of stereotypic attribution bias may be dynamic. The various factors involved (e.g., type of actor, type of event, and type of attribution) may be more or less influential depending on the context or targeted stereotype. Future research may reveal the factors that make SSAB subscale scores more or less critical in different performance domains.

Results revealed that both self-reported (Study 1) and manipulated (Study 2) perceived negative environments for women predicted SSAB among women but not among men. When women witness other women in the setting being devalued or treated negatively, this can activate the perception that STEM is a stereotypically male domain in which men's performance and contributions are valued over women's performance and contributions. These stereotypic perceptions may in turn influence spontaneous attributional processes, such that the initial reaction to men's setbacks is attributions to circumstances, whereas the initial reaction to women's STEM setbacks is attributions to ability.

Our focus on spontaneous, unprompted attributions connects the current research to other work on unconscious or implicit stereotype processes that influence women in STEM. Implicit gender-STEM stereotypes influence women's vulnerability to stereotype threat (Kiefer & Sekaquaptewa, 2007). Increases in implicit stereotyping across a semester predict lowered math course grades for women (Ramsey & Sekaquaptewa, 2011), and implicit stereotypes influenced outcomes when explicit endorsement of stereotypes did not (see also Nosek, Banaji, & Greenwald, 2002). In the current study, a measure of AASQ was not affected by exposure to an unwelcoming environment for women and did not predict STEM outcomes, perhaps because these more deliberative attributional responses reflect more controlled or self-monitored responding. As previously mentioned, however, several other differences between the AASQ and the SSAB measure may also account for these results. Taking this into consideration, we interpret the nonsignificant effects on the AASQ as providing some support for our assertion that negative environments for women exert a subtle negative influence on spontaneous attributional processing of stereotype-relevant information, but not on more direct controlled processing; however, more research that directly compares indirect, spontaneous measures such as SSAB to more direct, deliberative measures such as AASQ, is needed to support this claim.

The current work also established that SSAB is associated with negative outcomes for women in STEM. SSAB was associated with less perceived belonging and weaker postgraduate intentions to stay in STEM. Engaging in this stereotype-biased pattern of attribution of negative events may influence women's STEM outcomes because SSAB promotes gender-STEM stereotype maintenance over stereotype change. When gender-STEM stereotype-inconsistency (i.e., men's poor performance in STEM) is explained away to bad luck or other fleeting circumstances, the stereotype-inconsistent event seems less likely to occur again in the future, under different circumstances. Moreover, when gender-STEM stereotype-consistent events (i.e., women's setbacks in STEM) are attributed to stable abilities and dispositions, the stereotype-consistent event seems more likely to reoccur across different circumstances. Engaging in SSAB may lead to stereotype maintenance, even in the face of disconfirming information, and that may lead to diminished intentions to continue in STEM in the future. Partial support for this assertion was found in Study 2 that demonstrated that SSAB mediated the difference between the negative treatment condition and control condition on intentions to continue in STEM. However, this mediation model was significant by only one of the two effect size indicators, suggesting that the role of SSAB in explaining the effect of witnessing the negative treatment of women in an STEM setting may be small. SSAB may be one of several factors influenced by negative environments for women that lead to diminished STEM outcomes among women.

Overall, the current work contributes to scientists understanding how negative environments influence stigmatized individuals. Of particular interest, the current work used a newly created assessment of attribution bias, the SSAB measure, which may represent an important methodological advance in research on women in STEM. One important novel aspect of the SSAB measure is its focus on attributions for the STEM behaviors of other men and women, rather than one's own behaviors. Consistent with work demonstrating that people more easily perceive discrimination as happening to others, than to oneself (Taylor et al., 1990), stereotype-biased attributions for the behaviors of others may be of particular utility in predicting STEM outcomes. The current research provides evidence that when STEM environments are perceived as promoting gender-STEM stereotypes, the negative implications of group stereotype-biased attribution styles transfer to the self.

Practice Implications

The findings of this research can inform the work of those who seek to understand and address the effects of a negative climate on the retention of women in fields where they are currently underrepresented. Our work demonstrates that environments that are perceived as negative for women can significantly affect women's sense of belonging and intentions to persist in their field, outcomes that are important to the retention of women in those fields. Efforts to increase the representation and participation of women in STEM areas have identified the important role of climate and environmental factors on these outcomes (Sekaquaptewa, 2014; Settles, Cortina, Buchanan, & Miner, 2013). Our work contributes to this line of research by identifying an important process by which a negative environment harms women, namely, by promoting patterns of attribution that support gender stereotypes among women. Those seeking ways to retain more women in STEM fields should develop strategies to promote more favorable explanations for women's STEM setbacks, perhaps by providing role models that point out the external and situational influences on outcomes, and steer women away from internal attributions for their setbacks (Good, Aronson, & Inzlicht, 2003; Wilson & Linville, 1982). Such strategies may enhance the retention of women in STEM, even in the face of a negative climate, by reducing the emergence of SSAB.

Although the current work has focused on the influence of negative environments on SSAB among women, it is important to note that negative environments may exist for other marginalized groups in a given domain. For example, members of racial/ethnic groups such as African Americans, Hispanics and Latino/as, and Native Americans are also severely underrepresented in many fields and may also experience "chilly" climates (Smith, Cech, Mertz, Huntoon, & Moyer, 2014; Taxis, 2002). Therefore, SSAB may emerge when members of these groups experience negative environments. Interventions based on reducing internal attributions for setbacks among racial/ethnic minorities experiencing a negative climate in settings where they are underrepresented may serve to enhance retention of these groups in those settings. Finally, it is important to note that SSAB was less apparent in STEM settings that were more positive for women, indicating that efforts to create a positive climate for women in STEM may eliminate the need to create strategies to reduce SSAB.

Limitations

The current research integrated several areas of study relevant to understanding the influence of gender-STEM stereotypes on women, including achievement attribution, gender discrimination, and implicit biases. However, there were several limitations. First, although Study 2 provided a conceptual replication of the results of Study 1, it is important to note that in Study 1, perceived negative environment for women was operationalized as hearing derogatory comments about other women (a blatant cue), whereas in Study 2, it was operationalized as negative nonverbal behavior enacted towards women by her group members and the experimenter (a subtle cue). Our results suggest that SSAB is associated with both; however, more research is needed to clarify the specific types of cues that trigger SSAB.

Another limitation of the current work was the inability to fully test hypotheses concerning oral exam performance due to video recording malfunctions in Study 2. Responses to multiple choice questions were available, yet no significant differences between conditions emerged although significant differences would be predicted, according to stereotype threat theory (e.g., Inzlicht & Ben-Zeev, 2000; Spencer et al., 1999). It should be noted that the test used in Study 2 may have been unrealistically difficult, as the overall mean scores on multiple choice items (2.27 of the 5 possible points) and open-ended items (3.25 of the 7 possible points) were quite low. Moreover, previous work using a similar videotaping procedure demonstrated diminished performance on open-ended oral exam questions for women experiencing gender salience, a situation that promotes stereotype threat (e.g., Sekaquaptewa & Thompson, 2003). Therefore, it may be that oral exam performance is most affected by stereotype salience for open-ended, rather than multiple choice question types, or for computational math items, as opposed to comprehension and memory of science material. A re-examination of the effects of stereotype salience, induced by exposure to unwelcoming environments, on both closed and open-ended test items should be a goal of future research.

Last, it is important to note that minor improvements to the administration of our measure of SSAB could be made. First, we suggest that when the SSAB measure is administered online, software could be used that measures the amount of time elapsed before participants initiate their response to each sentence. This addition would allow for a more stringent test of our hypothesis that people engage SSAB spontaneously and without much deliberation. In addition, our longitudinal assessment of SSAB in Study 1 may have been affected by using an identical version of the SSAB measure at both time points. We recommend that future research assessing SSAB longitudinally use two different versions of the measure.

Conclusions

Our research demonstrated that STEM women's attributions for men's and women's STEM setbacks can be biased by gender-STEM stereotypes and that this biased attribution pattern is increased when cues in the setting promote the stereotype. Witnessing the negative treatment of other women in an STEM setting may be a common cue, perhaps even more so when the negative treatment involves subtle, nonverbal behaviors. When women perceive that women are undervalued, belittled, or ignored in STEM settings, this can affect their spontaneous attributional responses in a way that promotes and maintains gender-STEM stereotypes, and ultimately lowers women's intentions to remain in STEM fields or pursue STEM careers. To the extent that behavioral intentions predict behavioral choices (Fishbein & Ajzen, 2010), this research identifies a process by which subtle discrimination against women in STEM contributes to the "leaky pipeline."

The current work highlights the importance of developing strategies to reduce negative behaviors towards women in STEM settings. Of interest, negative behaviors enacted towards women can often be unintentionally negative, reflecting (unconscious) stereotypic biases, as in the case of microaggressions (Sue, 2010). Of course, more explicit gender discrimination exists as well (Heilman, Wallen, Fuchs, & Tamkins, 2004), and the current work suggests the need to develop strategies designed to improve "chilly climates" for women and increase their participation in STEM. Perhaps more importantly, given an often chilly climate for women, the current work suggests that one way to prevent the consequences of negative environments for women may be through interventions designed to change the attributions women make for other women's STEM setbacks. Like many other unconscious cognitive biases, making people aware of the bias may reduce the deleterious effects of the bias (e.g., Johns, Schmader, & Martens, 2005; Pronin, 2007). Our findings indicate that teaching STEM majors, especially female STEM majors, to make fewer internal attributions, and more external attributions, for female setbacks could help improve women's feelings of belonging in STEM and encourage women to continue in their field of study, even when experiencing a negative climate.

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) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by grants from the National Science Foundation GSE/RES grant number 0936434, awarded to the second author.

References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261–271. doi:10.1037/0022-0663.84.3.261

Basford, T. E., Offermann, L. R., & Behrend, T. S. (2014). Please accept my sincerest apologies: Examining follower reactions to leader apology. *Journal of Business Ethics*, *119*, 99–117. doi:10.1007/s10551-012-1613-y

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370. doi:10.1037/1089-2680.5.4.323

Beller, M., & Gafni, N. (1996). The 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology*, *88*, 365–377. doi:10.1037/0022-0663.88.2.365

Camacho, M. M., & Lord, S. M. (2011). Quebrando fronteras: Trends among Latino and Latina undergraduate engineers. *Journal of Hispanic Higher Education*, *10*, 134–146. doi:10.1177/1538192711402354

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching*, *32*, 243–257. doi:10.1002/tea.3660320305

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient acceptance: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045–1060. doi:10.1037/a0016239

Church, M. A., Elliot, A. J., & Gable, S. C. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology*, *93*, 43–54. doi:10.1037/0022-0663.93.1.43

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry*, *22*, 231–246. doi:10.1080/1047840X.2011.607313

Eccles, J. S., & Blumenfeld, P. (1985). Classroom experiences and student gender: Are there differences and do they matter? In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 79–114). New York, NY: Academic Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*, 889–906. doi:10.1037/0022-3514.38.6.889

Furr, R. M., & Rosenthal, R. (2003). Repeated-measures contrasts for "multiple-pattern" hypothesis. *Psychological Methods*, *8*, 275–293. doi:10.1037/1082-989X.8.3.275

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology*, *55*, 685–694. doi:10.1037/0022-3514.55.5.685

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*, 733–740. doi:10.1037/0022-3514.54.5.733

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*, 645–662. doi:10.1016/j.appdev.2003.09.002

Good, C. D., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of acceptance and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*, 700–717. doi:10.1037/a0026659

Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, *46*, 35–60. doi:10.1111/j.1540-4560.1990.tb01922.x

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, *46*, 44–56. doi:10.1037/0022-3514.46.1.44

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York, NY: The Guilford Press.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male tasks. *Journal of Applied Psychology*, *89*, 416–427. doi:10.1037/0021-9010.89.3.416

Henry, J. W., Martinko, M. J., & Pierce, M. A. (1993). Attributional style as a predictor of success in a first computer course. *Computers in Human Behavior*, *9*, 341–352. doi:10.1016/0747-5632(93)90027-P

Hong, Y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*, 588–599. doi:10.1037/0022-3514.77.3.588

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. doi:10.1111/1467-9280.00272

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, *16*, 175–179. doi:10.1111/j.0956-7976.2005.00799.x

Kanouse, D. E., & Hanson, L. R., Jr. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 47–62). Morristown, NJ: General Learning Press.

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457–501. doi:10.1146/annurev.ps.31.020180.002325

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender math stereotyping influences women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*, 825–832. doi:10.1016/j.jesp.2006.08.004

Kiefer, A. K., & Shih, M. (2006). Gender differences in persistence and attributions in stereotype relevant contexts. *Sex Roles*, *54*, 859–868. doi:10.1007/s11199-006-9051-x

Koch, S. C., Müller, S. M., & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education*, *51*, 1795–1803. doi:10.1016/j.compedu.2008.05.007

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. doi:10.1037/0022-3514.73.1.91

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, *30*, 41–62. doi:10.1207/s15327906mbr3001_3

Martinez, R., & Sewell, K. W. (2000). Explanatory style in college students: Gender differences and disability status. *College Student Journal*, *34*, 72–78. Retrieved from http://psycnet.apa.org/psycinfo/2000-03398-003

Marx, D. M., & Ko, S. J. (2012). Superstars "like" me: The effect of role model similarity on performance under threat. *European Journal of Social Psychology*, *42*, 807–812. doi:10.1002/ejsp.1907

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*, 16474–16479. doi:10.1073/pnas.1211286109

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. doi:10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., & Taylor, V. J. (2012). The role of situational cues in signaling and maintaining stereotype threat. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and applications* (pp. 17–33). New York, NY: Oxford University Press.

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. doi:10.1037/0022-3514.83.1.44

Peterson, C. (1992). Explanatory style. In C. P. Charles, J. W. Atkinson, D. C. McClelland, & J. Veroff (Eds.), *Motivation and personality: Handbook of thematic content analysis* (pp. 376–382). New York, NY: Cambridge University Press.

Peterson, C., & Barrett, L. C. (1987). Explanatory style and academic performance among university freshmen. *Journal of Personality and Social Psychology*, *53*, 603–607. doi:10.1037/0022-3514.83.1.44

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93–115. doi:10.1037/a0022658

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, *11*, 37–43. doi:10.1016/j.tics.2006.11.001

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. *Social Psychology of Education*, *14*, 377–387. doi:10.1007/s11218-010-9150-y

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review*, *12*, 85–110. doi:10.1023/A:1009037101170

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356. doi:10.1037/0033-295X.115.2.336

Sekaquaptewa, D. (2014, May/June). Calling for a change in the STEM climate: Research suggests the need to rethink science settings to recruit and retain women. *Observer*, 27. Retrieved from http://www.psychologicalscience.org/index.php/publications/observer/2014/may-june-14/calling-for-climate-change.html

Sekaquaptewa, D., & Espinoza, P. (2004). Biased processing of stereotype-incongruency is greater for low than high status groups. *Journal of Experimental Social Psychology*, *40*, 128–135. doi:10.1016/S0022-1031(03)00093-3

Sekaquaptewa, D., & Thompson, M. (2002). The differential effects of solo status on members of high and low status groups. *Personality and Social Psychology Bulletin*, *28*, 694–707. doi:10.1177/0146167202288013

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotypes, and performance expectancies: Their effects on women's public performance. *Journal of Experimental Social Psychology*, *39*, 68–74. doi:10.1016/S0022-1031(02)00508-5

Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Psychology*, *39*, 75–82. doi:10.1016/S0022-1031(02)00512-7

Settles, I. H., Cortina, L. M., Buchanan, N. T., & Miner, K. (2013). Derogation, discrimination, and (dis)satisfaction with jobs in science: A gendered analysis. *Psychology of Women Quarterly*, *37*, 179–191. doi:10.1177/0361684312468727

Shaw, A. K., & Stanton, D. E. (2012). Leaks in the pipeline: Separating demographic inertia from ongoing gender differences in academia. *Proceedings of the Royal Society B-Biological Sciences*, *279*, 3736–3741. doi:10.1098/rspb.2012.0822

Smith, J. L., Cech, E., Metz, A., Huntoon, M., & Moyer, C. (2014). Giving back or giving up: Native American student experiences in science and engineering. *Cultural Diversity & Ethnic Minority Psychology*, *20*, 413–429. doi:10.1037/a0036945

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard isn't natural: Women's acceptance with and motivation for male-dominated STEM fields as a function of expenditure concerns. *Personality and Social Psychology Bulletin*, *39*, 131–143. doi:10.1177/0146167212468332

Sojo, V. E., Wood, R. E., & Genat, A. E. (2015). Harmful workplace experiences and women's occupational well-being: A meta-analysis. *Psychology of Women Quarterly*, *40*, 10–40. doi:10.1177/0361684315599346

Spencer, S. J., Steele, C. M., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28. doi:10.1006/jesp.1998.1373

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, *34*, 379–440. doi:10.1016/S0065-2601(02)80009-0

Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. Hoboken, NJ: John Wiley.

Swim, J. K., & Cohen, L. L. (1997). Overt, covert, and subtle sexism: A comparison between the attitudes toward women and modern sexism scales. *Psychology of Women Quarterly*, *21*, 103–118. doi:10.1111/j.1471-6402.1997.tb00103.x

Swim, J. K., & Sanna, L. J. (1996). He's skilled, she's lucky: A meta-analysis of observers' attributions for women's and men's successes and failures. *Personality and Social Psychology Bulletin*, *22*, 507–519. doi:10.1177/0146167296225008

Taylor, D. M., Wright, S. C., Moghaddam, F. M., & Lalonde, R. N. (1990). The personal/group discrimination discrepancy: Perceiving my group, but not myself, to be a target for discrimination. *Personality and Social Psychology Bulletin*, *16*, 254–262. doi:10.1177/0146167290162006

Taxis, C. J. (2002). The underrepresentation of Hispanics/Latinos in nursing education: A deafening silence. *Research and Theory for Nursing Practice*, *16*, 249–262. doi:10.1891/rtnp.16.4.249.53029

Thoman, D. B., Arizaga, J. A., Smith, J. L., Story, T. S., & Soncuya, G. (2014). The grass is greener in non-science, technology, engineering, and math classes: Examining the role of competing belonging to undergraduate women's vulnerability to being pulled away from science. *Psychology of Women Quarterly*, *38*, 246–258. doi:10.1177/0361684313499899

Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., & Lee, J. K. (2013). Beyond performance: A motivational experience model of stereotype threat. *Educational Psychology Review*, *25*, 211–243. doi:10.1007/s10648-013-9219-1

Vargas, P., Sekaquaptewa, D., & von Hippel, W. (2004). It's not just what you think, it's also how you think: Prejudice as biased information processing. In J. D. Williams, W.-N. Lee, & C. P. Haugvedt (Eds.), *Diversity in advertising: Broadening the scope of research directions* (pp. 93–119). Mahwah, NJ: Lawrence Erlbaum.

Vargas, P., Sekaquaptewa, D., & von Hippel, W. (2007). Armed only with paper and pencil: "Low-tech" measures of implicit attitudes. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 103–124). New York, NY: Guilford Press.

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. doi:10.1037/0022-3514.92.1.82

Weary, G., Vaughn, L., Stewart, B. D., & Edwards, J. A. (2006). Adjusting for the correspondence bias: Effects of causal uncertainty, cognitive busyness, and causal strength of situational information. *Journal of Experimental Social Psychology*, *42*, 87–94. doi:10.1016/j.jesp.2005.01.003

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, *71*, 3–25. doi:10.1037/0022-0663.71.1.3

Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: Questioning kappa-squared as mediation effect size measure. *Psychological Methods*, *20*, 193–203. doi:10.1037/met0000029

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, *42*, 367–376. doi:10.1037/0022-3514.42.2.367

Wong, A. F. L., & Fraser, B. J. (1996). Environment-attitude association in the chemistry laboratory classroom. *Research in Science and Technological Education*, *64*, 29–40. doi:10.1080/0263514960140107

Ybarra, O. (2002). Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin*, *128*, 421–441. doi:10.1037//0033-2909.128.3.421