Research Article

Psychology of Women Quarterly
2018, Vol. 42(1) 72-87

DOI: 10.1177/0361684317741889
journals.sagepub.com/home/pwq
SAGE

# Hostile and Benevolent Sexism and College Women's STEM Outcomes

**Sophie L. Kuchynka[1], Kristen Salomon[1], Jennifer K. Bosson[1], Mona El-Hout[1], Elizabeth Kiebel[1], Claudia Cooperman[2], and Ryan Toomey[3]**

**Abstract**

We examined associations of perceived ambivalent sexism with women's outcomes in university Science, Technology, Engineering, and Mathematics (STEM) courses as a function of their STEM identity. Women ($N = 592$) who varied in STEM identification reported on their personal experiences with benevolent and hostile sexism and indicated their STEM major intentions, STEM self-efficacy, and STEM grade point average (GPA). Women perceived more benevolent sexism (i.e., protective paternalism and complementary gender differentiation) than hostile sexism in STEM courses, and their STEM identity moderated the associations between sexism and STEM outcomes. Among weakly-identified (but not strongly-identified) women, protective paternalism predicted lower STEM major intentions, STEM self-efficacy, and STEM GPA; hostile sexism predicted lower STEM GPA. Male STEM students ($N = 163$) reported more protective paternalism attitudes than hostile sexism attitudes, suggesting that women's perceptions were not without warrant. We discuss implications of these results for understanding women's underrepresentation in STEM and advise STEM educators to avoid well-intended, but paternalistic, messages that convey negative stereotypes about women's STEM competence.

**Keywords**

women in STEM, gender stereotypes, hostile sexism, benevolent sexism

Women in the United States are underrepresented in Science, Technology, Engineering, and Mathematics (STEM) fields (Ceci & Williams, 2007; Hill, Corbett, & St. Rose, 2010). Gender stereotypes about women's lack of STEM aptitude play a key role in this disparity: Gender stereotypes can reduce women's STEM career interests (Cundiff, Vescio, Loken, & Lo, 2013), lower feelings of STEM self-efficacy (Bonnot & Croizet, 2007), and undermine STEM performance (Spencer, Steele, & Quinn, 1999). Past work illuminates the negative gender stereotypes that women face in STEM environments but does not distinguish between the hostile and benevolent ways in which such stereotypes are conveyed to women. In the current study, we used ambivalent sexism theory (Glick & Fiske, 1996) as a lens for examining women's experiences with sexism in STEM contexts. Specifically, we examined the independent associations of perceived hostile and benevolent sexism with women's intentions to major in a STEM field, STEM self-efficacy, and STEM grade point average (GPA). We also examined whether women's identification with STEM (their *STEM identity*) moderates the links between sexism and STEM outcomes.

## *Ambivalent Sexism Theory*

Glick and Fiske (1996, 2001) proposed that gender relations are characterized by a combination of hostile and benevolent attitudes and behaviors toward women. *Hostile sexism* consists of overtly negative, angry attitudes and behavior toward women who occupy masculine domains or push for social change; it often takes the form of disparaging attitudes about women's incompetence relative to men's or suspicion and resentment directed at women who are perceived as insubordinate. *Benevolent sexism* consists of affectively positive but condescending attitudes and reactions to women who embrace traditional gender roles; it can take the form of

[1] Department of Psychology, The University of South Florida, Tampa, FL, USA
[2] School of Public Affairs, The University of South Florida, Tampa, FL, USA
[3] Department of Chemical and Biomedical Engineering, The University of South Florida, Tampa, FL, USA

**Corresponding Author:**
Sophie L. Kuchynka, Department of Psychology, The University of South Florida, PCD 4118G, 4202 East Fowler Avenue, Tampa, FL 33620, USA.
Email: skuchynka@mail.usf.edu

*protective paternalism* (protectiveness toward women who are seen as weaker than men), *complementary gender differentiation* (flattering beliefs about women's morality and sweetness), or *heterosexual intimacy* (the belief that men are incomplete without a woman's love).

Women's responses to hostilely sexist treatment include anger, disgust, stress, and heightened cardiovascular reactivity (Barreto & Ellemers, 2005; Bosson, Pinel, & Vandello, 2009; Salomon, Burgess, & Bosson, 2015; Schneider, Tomaka, & Palacios, 2001). In the workplace, women exposed to hostilely sexist treatment experience lower job satisfaction and increased absenteeism, depression, and physical illness symptoms (Fitzgerald, 1993). In contrast, exposure to benevolent sexism elicits a different set of responses including self-objectification, body shame (Calogero & Jost, 2011), and prolonged cardiovascular recovery (Salomon et al., 2015). Benevolent sexism also impairs women's cognitive performance more than hostile sexism does (Dardenne, Dumont, & Bollier, 2007; Vescio, Gervais, Snyder, & Hoover, 2005) because it reduces working memory capacity by increasing intrusive thoughts about incompetence and by eliciting attempts to suppress these thoughts (Dardenne et al., 2007; Dardenne et al., 2013).

## Stereotypes and Ambivalent Sexism in STEM Environments

Women's underrepresentation in STEM fields reflects deeply embedded stereotypes that link men and masculinity more strongly than women and femininity with STEM competence (Nosek, Banaji, & Greenwald, 2002). To examine the role of these stereotypes in women's STEM experiences, researchers often measure people's endorsement of decontextualized statements about women's STEM abilities (e.g., "Men are just better at science than women"; Lane, Goh, & Driver-Linn, 2012). Research illuminates the pervasiveness of the gender stereotypes that women face in STEM environments (Gunderson, Ramirez, Levine, & Beilock, 2012; Hartman & Hartman, 2008) and demonstrates that gender-based stereotypes can undermine women's STEM interest (Cundiff et al., 2013), reduce their felt belonging in STEM environments (Cheryan, Plaut, Davies, & Steele, 2009), and undermine their STEM motivation (Fogliati & Bussey, 2013). Frequent stereotype exposure can also lead women to internalize the stereotypes, which can further degrade their interest in pursuing STEM careers (Schmader, Johns, & Barquissau, 2004).

One question that remains unaddressed, however, is how such gender stereotypes are delivered to women in STEM environments. Do women perceive that these stereotypes are delivered in a hostile manner, a benevolent manner, or both? Which delivery of gender STEM stereotype is perceived more frequently, and do they predict different outcomes for women? Hostile deliveries of STEM stereotypes may involve insulting jokes about women's inferiority in STEM fields or a tendency to ignore women's contributions to STEM tasks, whereas benevolent deliveries may involve treating women as if they cannot handle difficult STEM challenges (Bernstein & Russo, 2008; King et al., 2012) or have communal personalities that are ill-suited for competitive STEM environments. In other words, stereotypes about women's relative lack of STEM competence may arrive via either hostile or benevolent sexism.

There is reason to assume that gender STEM stereotypes assume both hostile and benevolent forms, although with different frequency. Women perceived as non-traditional—which likely characterizes many women who inhabit STEM environments—are more frequent targets of hostile sexism than are traditional women (Sibley & Wilson, 2004). And yet, because overt expressions of gender bias are normatively eschewed and legally sanctioned (Dovidio & Gaertner, 1998; Pettigrew & Meertens, 2006), even people who hold negative gender STEM stereotypes may be disinclined to communicate them overtly. This may be especially true in a university environment in which students and faculty alike receive regular messages regarding gender sensitivity. Moreover, across cultures, men tend to endorse benevolent sexism more strongly than hostile sexism (Glick & Fiske, 2001), suggesting that their behaviors may reflect more benevolence than hostility toward women. Thus, we expected women to report more benevolent than hostile sexism experiences in their university STEM courses.

We also expected benevolent sexism to predict negative STEM outcomes for women. As noted, benevolent sexism impairs women's cognitive performance, mediated by mental intrusions about competence and intrusive thought suppression (Dardenne et al., 2007; Dardenne et al., 2013). Similarly, subtle signs of bias—such as benevolent sexism—may produce cognitive disruption because their ambiguity elicits uncertainty and hypervigilance in targets (Murphy, Richeson, Shelton, Rheinschmidt, & Bergsieker, 2012). Moreover, benevolent sexism exposure elicits prolonged, low-level stress responses (Salomon et al., 2015), indicating that its stressful effects last a relatively long time. Finally, benevolent sexism can have a pacifying effect on women, reducing their motivation for social change (Becker & Wright, 2011) and increasing their tendencies to critically surveil and objectify themselves (Calogero & Jost, 2011). Thus, we expected perceived benevolent sexism to predict decreases in women's STEM major intentions, self-efficacy, and performance.

But which manifestation of benevolent sexism should predict negative STEM outcomes? As noted earlier, benevolent sexism comprises three subjectively positive stances toward women: protective paternalism, complementary gender differentiation, and heterosexual intimacy (Glick & Fiske, 1996). In the context of STEM classes, protective paternalism might involve treating women as if they are weaker or need more assistance than men, and complementary gender differentiation might involve treating women as if their virtuous and sweet personalities render them unsuited for competitive STEM fields. Because protective paternalism implies that women cannot handle STEM tasks without men's assistance, we expected it to impair women's cognitive abilities and undermine their confidence and motivation. In contrast, we reasoned that complementary gender differentiation does not directly communicate negative stereotypes about women's lack of STEM competence because it pertains to women's moral character rather than their intelligence. And yet, given cross-cultural stereotypes of women as higher in warmth and lower in competence than men (Fiske, Cuddy, Glick, & Xu, 2002), reminders of stereotypes about their warmth may also communicate to women—albeit indirectly—the assumption that they lack competence. If so, then complementary gender differentiation may predict negative STEM outcomes for women. We therefore treated complementary gender differentiation as an exploratory predictor of women's STEM outcomes in analyses. Finally, given that heterosexual intimacy involves the belief that men are incomplete without the romantic love of a woman (e.g., "Every man ought to have a woman he adores"; Glick & Fiske, 1996), we did not consider it here because of our exclusive focus on treatment that occurs within academic environments and communicates stereotypes about women's STEM competence.

Our expectations regarding the associations of hostile sexism with women's STEM outcomes were unclear. Although hostile sexism exposure predicts anger, negative affect, and stress in women, it also elicits approach-oriented motivations including desires to confront the source of sexism and engage in social change efforts (Becker & Wright, 2011; Chaudoir & Quinn, 2010; Schneider et al., 2001). Moreover, women's angry reactions to hostile sexism tend to fade quickly (Salomon et al., 2015), suggesting a capacity to cope promptly with hostile episodes and direct cognitive resources elsewhere (rather than perseverating). This may be because hostile sexism explicitly conveys animosity and is thus readily perceived as a form of prejudice (Dardenne et al., 2007). Recognizing sexism as prejudice is valuable for ego protection because attributing negative outcomes to prejudice can buffer targets' self-esteem (Major, Quinton, Kaiser, & McCoy, 2003). Thus, in the short-term time line of this study, it was not clear to us whether perceiving hostilely sexist treatment would predict negative STEM outcomes for women. We therefore treated hostile sexism as an exploratory predictor in analyses.

## Perceived Sexism and STEM Identity

We expected the links between perceived benevolent sexism and STEM outcomes to be moderated by the strength of women's *STEM identity*, defined as the importance of STEM domains to women's self-concept (Leach et al., 2008; Luhtanen & Crocker, 1992). Specifically, based on motivational intensity theory (Brehm & Self, 1989; Gendolla & Richter, 2010), we expected the associations of benevolent sexism and STEM outcomes to be especially strong among women low in STEM identity. Conversely, women with a strong STEM identity should experience relatively weak links between benevolent sexism and STEM outcomes.

Motivational intensity theory suggests that as investment in a given domain increases, motivation to perform well in the domain also increases, and this leads people to exert more effort during difficult tasks (Brehm & Self, 1989; Gendolla & Richter, 2010). Thus, individuals with a stronger domain identity should persist at higher levels of difficulty, whereas those with a weaker domain identity should lose motivation and withdraw effort sooner. Effort is maintained when the benefits of success (e.g., self-efficacy, achievement) outweigh the costs (e.g., exertion) and, conversely, effort is withheld when the costs outweigh the benefits. For women high in STEM identity, who are strongly invested in STEM domains (e.g., Steele, 1997; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), the benefits of success in the face of protective paternalism (i.e., reminders that the task is likely "too difficult" for them) should outweigh the costs of effort. Conversely, for women low in STEM identity, the effort needed to perform tasks that are presented as "too difficult" may outweigh the potential benefits of task success. As one illustration of this phenomenon, African American adolescents who identified more strongly with academics were buffered against the negative effects of discrimination on academic performance, while those who identified more weakly showed impaired performance after experiencing discrimination (Thompson & Gregory, 2011).

At first glance, this prediction may seem inconsistent with findings from the stereotype threat literature. Theoretically, stereotype threat should most strongly affect those who care the most about doing well in the threatened domain—in other words, those with high domain identification (Steele, 1997). For example, some studies find that women who identify more strongly with math tend to suffer the largest performance decrements following manipulated reminders of gender math stereotypes (Keller, 2007; Steinberg, Okun, & Aiken, 2012). However, the literature is inconsistent regarding the role of domain identification in stereotype threat effects. While one meta-analysis found that the harmful effects of stereotype threat on performance were stronger among domain-identified targets than non-identified targets (Walton & Cohen, 2003), another meta-analysis found that stereotype threat effects were smaller for highly math-identified women than moderately low math-identified women (Nguyen & Ryan, 2008). As Nguyen and Ryan noted, only a small number of stereotype threat experiments include domain identification as a moderator, and the findings from these studies are mixed.

Further, stereotype threat manipulations tend to involve the presence or absence of contextual cues that activate stereotypes and have immediate effects on testing outcomes. In contrast, our interest is in different deliveries (i.e., hostile vs. benevolent) of the same stereotype (i.e., "women lack STEM competence") that occur in everyday, naturalistic contexts. The stereotype threat literature is largely silent on the question of whether stereotype encounters that involve differential behavioral treatment—that is, overt insult versus patronizing protectiveness—are differently impactful among targets who vary in domain identification. Finally, stereotype threat is only one of multiple theorized pathways by which gender stereotypes may reduce women's performance in negatively stereotyped domains. While stereotype threat has acute effects on performance, specifically in testing contexts when women are concerned about confirming the stereotype, stereotypes may also more generally decrease women's interest and engagement in certain domains by lowering their self-confidence and the value that they place on the domain (Baron, Schmader, Cvencek, & Meltzoff, 2014). Little research has examined how domain identification moderates the effects of stereotypes on motivation and interest, two of the main goals of the current study. The current focus on the moderating role of STEM identity should add to our understanding of when and why domain identification matters among targets of negative stereotypes.

### Is Sexism "in the Air"?

Thus far, we have focused our discussion primarily on women's perceptions of hostile and benevolent sexism in STEM environments. But how do we know whether perceptions of sexism reflect real sexism experiences? For instance, women differ in the extent to which they expect to be gender stereotyped by others, or their *stigma consciousness* (Pinel, 1999), and those higher in stigma consciousness may report more sexism due to their perceptions rather than actual experiences. To bolster our confidence that the sexism reported by women in this study was actually "in the air," as opposed to merely "in their heads," we took two steps. First, we measured and controlled for individual differences in women's stigma consciousness to remove variance in our measure of perceived sexism that reflects women's chronic expectations of being stereotyped. Second, we measured the hostile and benevolent sexist attitudes of male students who were recruited from the same STEM courses as the female students. While a measure of privately held attitudes cannot reveal men's behavioral enactment of hostile and benevolent sexism toward women in their classes, it can at least offer insight into the extent to which the STEM environments that surround our female participants are occupied by male students who endorse hostile and benevolent sexist beliefs. Consistent with cross-cultural patterns (Glick & Fiske, 2001), we expected male STEM students to report higher levels of benevolent than hostile sexism.

### The Present Study

The harmful effects of gender stereotypes on women's STEM outcomes are well-established, but we do not yet know whether such stereotypes are typically conveyed to women in STEM environments via primarily hostile or benevolent treatment. Further, while most research on sexism's effects on women in STEM examines hostile sexism (Faulkner, 2009; Settles, Cortina, Buchanan, & Miner, 2013), there is reason to assume that benevolent sexism predicts negative STEM outcomes for women. Because no existing measure of women's personal experiences of sexism distinguishes between hostile and benevolent treatment, we created measures of these constructs for use in the current study. Our goal was not to fully develop and validate a scale but instead to create measures that (1) allowed us to assess women's self-reported sexism experiences with adequate internal reliability and convergent validity and (2) clearly differentiated between hostile and benevolent treatment. This latter goal allowed us to examine the unique roles of hostile and benevolent treatment in predicting STEM outcomes (STEM major intentions, STEM self-efficacy, and STEM GPA), moderated by women's STEM identity. We also examined whether women's hostile and benevolent sexism reports coincided with men's reports of their hostile and benevolent sexism attitudes.

Participants were women and men undergraduates currently enrolled in large, lower-level engineering and health science courses, as well as psychology students who had taken at least one undergraduate STEM course. We predicted that female STEM students would report more benevolent than hostile sexism experiences in their STEM courses (Hypothesis 1) and that male STEM students would endorse benevolent sexism attitudes more strongly than hostile sexism attitudes (Hypothesis 2). Next, we expected that protective paternalism would predict lower STEM major intentions, STEM self-efficacy, and STEM GPA among women in STEM courses (Hypothesis 3). Finally, we predicted that the associations between protective paternalism exposure and women's STEM outcomes would be moderated by the strength of women's STEM identity. That is, we expected protective paternalism to more strongly predict negative STEM outcomes for women who were weakly STEM-identified than women who were strongly STEM-identified (Hypothesis 4). We treated as exploratory analyses in which complementary gender differentiation and hostile sexism predicted women's STEM outcomes.

## Method

### Constructing Measures of Hostile and Benevolent Experiences

We wrote items to measure specific acts of hostile and benevolent sexism, based on ambivalent sexism theory (Glick & Fiske, 1996), and in consultation with instructors of university STEM courses (the last two authors) to ensure that they captured realistic, relevant experiences. To measure specific behaviors without imposing any interpretation on those behaviors, we did not use the word "sexism" in any items. Fourteen items assessed experiences with hostile sexism (e.g., "How often have you been harassed or threatened"), 8 items assessed experiences with protective paternalism (e.g., "How often have you been treated as if you need a man's help"), and 7 items assessed experiences with complementary gender differentiation (e.g., "How often have you been treated as if you are a 'nice girl'"). Respondents indicated how frequently, in their daily life, they experienced each type of treatment on 6-point scales. Scale labels were modified from the Schedule of Sexist Events (SSE; Klonoff & Landrine, 1995), which uses a scale in which 1 = *the event has never happened*, 2 = *the event happened once in a while (less than 10% of the time)*, 3 = *the event happened sometimes (10–25% of the time)*, 4 = *the event happened a lot (26–49% of the time)*, 5 = *the event happened most of the time (50–70% of the time)*, and 6 = *the event happened almost all of the time (more than 70% of the time)*. Because we wanted objective frequency estimates of recent and current sexism episodes, we converted the percentages from the SSE Scale to days per month and used the following scale: 1 = *the event has never happened to you*, 2 = *the event happens rarely (2–3 times per month)*, 3 = *the event happens sometimes (4–7 times per month)*, 4 = *the event happens often (8–14 times per month)*, 5 = *the event happens frequently (15–21 times per month)*, and 6 = *the event happens very frequently (more than 21 times per month)*.

*Sample 1.* Undergraduate women enrolled in psychology courses ($N = 144$, $Mdn_{age} = 20$) responded to the sets of hostile and benevolent sexism items as blocks (in counterbalanced order); after each block, they answered "To what extent do you view the behaviors described above as sexist?" on a scale of 1 (*not at all*) to 9 (*very much*). Women also completed the SSE (Klonoff & Landrine, 1995), which assesses how frequently women experienced 19 sexist episodes (e.g., "Called a sexist name") in both the past year and entire life. Importantly, women rated the sets of hostile and benevolent items as equally sexist, $t(144) < 1$, $p > .54$, and they rated both sets above the scale midpoint on sexism, $M$s = 5.90 and 5.97, $SD$s = 2.29 and 2.17, $t$s > 4.75, $p$s < .001.

*Sample 2.* Women enrolled in psychology courses ($N = 216$, $Mdn_{age} = 20$) responded to the hostile and benevolent sexism items on the same scales reported above; in this survey, hostile and benevolent items were randomly ordered on a participant-by-participant basis. After rating the sexism items, women completed Pinel's (1999) 10-item Stigma Consciousness Questionnaire (SCQ), which measures the extent to which they expect to be stereotyped by others (e.g., "Stereotypes about women have not affected me personally"). All items are rated on scales of 1 (*strongly disagree*) to 5 (*strongly agree*; see Procedure section for additional details about this scale).

*Combined Samples 1 and 2.* Results from a $\chi^2$ test of sphericity ($p < .001$) and a Kaiser–Meyer–Oklin test for sampling adequacy (.92) on the combined samples ($N = 360$) revealed that the data were appropriate for factor analysis and the sample size was large enough to evaluate the factor structure. We therefore submitted the 29 items to a principal axis factor analysis, using Promax rotation due to the presumed positive correlations among the factors. To determine the number of factors to retain, we examined the Scree plot, the percentage of total variance explained by each factor, and the interpretability of each factor in light of our theoretical framing. We also required a minimum loading of 3 items on each factor. Items were retained on a factor if they loaded at least .45 on that factor, did not load above .39 on any another factor, and did not have cross-loadings of less than .10 difference from their highest loading factor.

Using these criteria, the factor analysis results suggested that three factors should be retained. For instance, the Scree test (Cattell, 1966) indicated a three-factor solution, and the three-factor solution was superior on interpretability because it produced clear-cut hostile sexism, protective paternalism, and complementary gender differentiation factors. Inspection of the factor structure revealed that the first factor (accounting for 42.5% of the variance) contained items measuring hostile sexism, the second factor (accounting for 9.1% of the variance) contained items measuring protective paternalism, and the third factor (accounting for 4.7% of the variance) contained items measuring complementary gender differentiation. We therefore created a hostile sexism subscale (7 items; all factor loadings > .48; $\alpha = .84$), a protective paternalism subscale (5 items; all factor loadings > .46; $\alpha = .89$), and a complementary gender differentiation subscale (4 items; all factor loadings > .45; $\alpha = .80$).

*Sample 3.* Undergraduate women completed the sexism items as part of a large-scale survey required of all students in the psychology department's participant pool ($N = 935$). For the three-factor solution, the data yielded an excellent model fit, comparative fit index = .97, root mean square error of approximation = .05, standardized root mean residual = .04; (Hu & Bentler, 1999). The $\chi^2$ test was significant, indicating bad fit. Yet, because this significance test is oversensitive to sample size, we would expect it to be significant with this sample: $\chi^2(101) = 363.37$, $p < .01$.

Table 1 contains the final 16 sexism items, their factor loadings, and their communalities. Supporting convergent validity of the sexism measures, correlations with the SSE (collected in Sample 1) were significant or marginally significant for hostile sexism (SSE-year: $r = .69$, SSE-life: $r = .71$; $p$s < .001), protective paternalism (SSE-year: $r = .50$, SSE-life: $r = .57$; $p$s < .001), and complementary gender differentiation (SSE-year: $r = .15$, $p < .08$, SSE-life: $r = .20$, $p < .02$). Moreover, correlations with the SCQ (collected in Sample 2) were significant for hostile sexism ($r = .47$, $p < .001$), protective paternalism ($r = .48$, $p < .001$), and complementary gender differentiation ($r = .27$, $p < .001$). This indicates that the three sexism measures correlated meaningfully both

**Table 1.** Final Items From the Experiences With Ambivalent Sexism Scale.

| | Factor Loadings | | | Communalities | |
|---|---|---|---|---|---|
| Item ("How often have you . . . ") | HS | PP | CGD | Initial | Extraction |
| Been subjected to offensive comments | **.84** | .06 | .10 | .70 | .70 |
| Been treated in an insulting manner | **.74** | −.001 | .03 | .68 | .64 |
| Been harassed or threatened | **.80** | −.10 | .01 | .57 | .54 |
| Been the target of obscene sexual comments or gestures | **.73** | .11 | .11 | .60 | .56 |
| Been the target of insulting jokes | **.68** | .12 | .03 | .57 | .54 |
| Been accused of using your gender to your advantage | **.58** | −.01 | −.03 | .43 | .39 |
| Been treated in an angry manner | **.55** | −.14 | −.02 | .54 | .42 |
| Been treated as if you need more assistance than men do | −.07 | **.86** | .05 | .70 | .71 |
| Been treated as if you are weaker than men | −.05 | **.79** | .03 | .63 | .63 |
| Been treated as if you need a man's help | −.02 | **.77** | .06 | .67 | .66 |
| Been treated as if you are more vulnerable than men | .06 | **.55** | .18 | .61 | .60 |
| Been treated as if you cannot take care of yourself | .10 | **.45** | −.02 | .54 | .50 |
| Been treated as if you are a "nice girl" | −.008 | .01 | **.83** | .53 | .62 |
| Been treated as if you are innocent | −.06 | −.05 | **.77** | .49 | .55 |
| Been treated as if others expected you to be sweet and friendly | .03 | .11 | **.63** | .48 | 48 |
| Been treated as if you are more morally pure than men | .12 | .01 | **.60** | .40 | .41 |

*Note.* All scale items were preceded with the phrase "How often have you" and ended with the phrase "in your STEM courses and course-related activities." Bolded values indicate loadings on primary target factors. Items were retained if they loaded ≥.45 on the primary target factor, <.39 on all other factors, and the difference between loadings on the primary target factor and other factors was at least .10. STEM = science, technology, engineering, and mathematics; HS = hostile sexism; PP = protective paternalism; CGD = complementary gender differentiation.

with women's experiences of past sexism and their expectations of encountering gender-based stereotypes.

## Participants

Given the novelty of the constructs under investigation, the use of a new scale, and our desire to reach a wide sample of university STEM environments, we decided a priori to seek as large a sample as possible. Thus, while a power analysis conducted using G*Power indicated that a sample of $n = 109$ was sufficient to detect an interaction for a moderate effect size in the women's data, we continued collecting data for the entire semester. Participants were 755 undergraduates (592 women and 163 men) at a large, Southeastern public university. Women were recruited from engineering or health science courses ($n = 407$) or from psychology courses ($n = 185$). For the purposes of this study, we did not define psychology or other social sciences as STEM disciplines; therefore, to be eligible for participation, women recruited from psychology courses had to indicate that they had taken at least one university-level STEM course (which we defined for them). Men were recruited from the same engineering and health science courses as women participants, to ensure comparability of men's and women's reports of ambivalent sexism (see Table 2 for participant demographics). Participants recruited from psychology received course credit for their time, those recruited from one engineering course received a chance to win a US$100 gift card, and those recruited from other engineering courses or health science courses received extra course credit. No data were excluded from analyses, although *df*s differ across analyses because some participants provided incomplete data (see below on missing data).

To ensure comparability of the samples of women recruited from STEM courses versus from psychology, we conducted independent samples *t*-tests on women's STEM identity and their reports of the sexism that they experienced in their STEM courses. Women recruited from STEM versus psychology courses did not differ in their reports of hostile sexism, protective paternalism, or complementary gender differentiation (*p*s > .32). However, women recruited from STEM courses reported a stronger STEM identity than those recruited from psychology courses, $t(593) = 3.64$, $p < .01$.

## Procedure

After giving their informed consent, participants were introduced to an online survey. The first question asked them to report their gender and then routed women and men to different surveys. Women first read a definition of "STEM fields" accompanied by a detailed list of STEM disciplines and then rated the strength of their STEM identity. We placed the STEM identity scale first in the survey, given its role as the individual difference moderator in our theoretical model. Next, women reported their experiences with hostile and benevolent sexism, which served as our primary predictor variables. Finally, women completed measures of the three outcome variables: STEM major intentions, STEM self-efficacy, and STEM GPA. Women were instructed to keep their college STEM courses in mind as they completed these last measures. In contrast, men simply completed the Ambivalent Sexism Inventory (ASI; Glick & Fiske, 1996).

**Table 2.** Participant Demographics and Means and Standard Deviations for Study Variables by Recruitment Sample.

| | Women in STEM Courses | Women in Psychology Pool | Men in STEM Courses |
|---|---|---|---|
| Variables | $n = 410$ | $n = 185$ | $n = 163$ |
| Race | | | |
| White | 244 (60.0%) | 108 (58.4%) | 99 (60.7%) |
| Black | 70 (17.2%) | 22 (11.9%) | 28 (17.2%) |
| Asian | 28 (6.9%) | 22 (11.9%) | 9 (5.5%) |
| Native American | 3 (0.7%) | 0 | 0 |
| Middle Eastern | 12 (2.9%) | 7 (3.8%) | 5 (3.1%) |
| Biracial | 26 (6.4%) | 15 (8.1%) | 10 (6.1%) |
| Other | 19 (4.7%) | 6 (3.2%) | 7 (4.3%) |
| Age | 20.85 (2.72) | 20.04 (2.61) | 21.60 (3.72) |
| Class rank | | | |
| First year | 32 (7.9%) | 55 (29.7%) | 9 (5.5%) |
| Second year | 110 (27.0%) | 44 (23.8%) | 33 (20.2%) |
| Third year | 129 (31.7%) | 42 (22.7%) | 2 (1.2%) |
| Fourth year | 97 (23.8%) | 37 (20.0%) | 54 (33.1%) |
| Fifth year + | 38 (9.3%) | 7 (3.8) | 19 (11.7%) |
| STEM identity | 3.52 (0.85) | 3.21 (1.15) | 3.52 (0.77) |
| HS experiences | 1.49 (0.77) | 1.52 (0.82) | |
| CGD experiences | 2.39 (1.11) | 2.33 (1.01) | |
| PP experiences | 1.86 (1.02) | 1.95 (1.12) | |
| Number of STEM courses | 12.67 (3.15) | 10.52 (3.42) | 13.52 (3.86) |

*Note.* Age is presented as mean years of age with standard deviations in parentheses. STEM = science, technology, engineering, and mathematics; HS = hostile sexism; PP = protective paternalism; CGD = complementary gender differentiation.

All data and materials are available from the first author via email.

## Measures

*STEM identity.* Women's STEM identity was measured with a 6-item science identity scale modified from Chemers, Zurbriggen, Syed, Goza, and Bearman (2011); we replaced "scientist" with "STEM student" in each item (e.g., "In general, being a STEM student is an important part of my self-image" and "I am a STEM student"). Participants responded on scales from 1 (*strongly disagree*) to 5 (*strongly agree*). A principal components analysis on the 6 items yielded a one-factor solution that explained 73% of the total variance. We calculated mean scores ($\alpha = .92$), with higher scores reflecting a stronger STEM identity.

*Experiences with ambivalent sexism.* Women completed our measures of hostile and benevolent sexism with instructions and items modified to focus on their STEM courses (see Table 1 for the full scale). Items were rated on 6-point scales from 1 (*never*) to 6 (*very frequently/more than 21 times per month*). Due to an error, 1 item of the Hostile Sexism subscale ("Been treated in an insulting manner") was not included in study materials. This resulted in using a 6-item version of the Hostile Sexism subscale. We averaged items to create hostile sexism ($\alpha = .91$), protective paternalism ($\alpha = .94$), and complementary gender differentiation ($\alpha = .84$) composites.

*Intentions to major.* Three items assessed women's intentions to major in a STEM field: "How likely are you to major in a STEM field?" "How committed are you to majoring in a STEM field?" and "To what extent are you interested in majoring in a NON-STEM field?" (reverse scored). Items were rated on scales of 1 (*not at all*) to 5 (*very*), and a principal components analysis yielded a one-factor solution that explained 70.5% of the total variance. We calculated mean scores ($\alpha = .78$), with higher scores reflecting stronger intentions to major.

*STEM self-efficacy.* Six items from Schmader, Johns, and Barquissau (2004) that measured people's math self-efficacy were modified to assess STEM self-efficacy. Sample items are "I am confident that I could successfully complete the educational requirements for a STEM major" and "I feel like I have to work harder than other people in my STEM class(es) in order to do well" (reverse scored). These items were rated on scales of 1 (*strongly disagree*) to 5 (*strongly agree*). Participants also ranked their STEM ability compared to other students at their university on a scale that ranged from 0 to 100th percentile. We standardized these 7 items and calculated mean scores ($\alpha = .74$), with higher scores reflecting higher STEM self-efficacy.

*Stigma consciousness.* Women completed Pinel's (1999) 10-item SCQ, which measures the extent to which women expect to be stereotyped by others (e.g., "Stereotypes about women have not affected me personally"). All items are rated

**Table 3.** Bivariate Correlations, Means, and Standard Deviations for All Variables.

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. HS | | | | | | | | | | | |
| 2. PP | .80*** | | | | | | | | | | |
| 3. CGD | .55*** | .71*** | | | | | | | | | |
| 4. STEM identity | .14** | .13** | .19*** | | | | | | | | |
| 5. Major intentions | −.08 | −.09* | .01 | .49*** | | | | | | | |
| 6. STEM GPA | .01 | −.01 | .04 | .09* | −.003 | | | | | | |
| 7. Self-efficacy | −.09* | −.09* | .05 | .29*** | .38*** | .26*** | | | | | |
| 8. SCQ | .23*** | .32*** | .26*** | −.05 | −.03 | .08* | .05 | | | | |
| 9. STEM # | .10* | .08* | .15*** | .23*** | .28*** | .06 | .15*** | .02 | | | |
| 10. Age | −.08 | −.04 | −.08 | −.04 | −.01 | −.10* | −.02 | −.06 | .37*** | | |
| 11. Class year | −.03 | −.002 | .05 | .01 | .06 | −.01 | .09* | .01 | .62*** | .59*** | |
| 12. Race | .06 | .03 | .002 | .05 | −.02 | −.08 | −.06 | .03 | −.002 | −.01 | −.01 |
| Mean | 1.50 | 1.89 | 2.37 | 3.46 | 4.07 | 3.37 | 0.007 | 3.23 | 12.28 | 20.81 | 2.90 |
| Standard deviation | 0.78 | 1.05 | 1.10 | 0.93 | 0.95 | 0.51 | 0.62 | 0.64 | 3.55 | 3.01 | 1.16 |

*Note.* Race is coded 0 = *White*, 1 = *non-White*. HS = hostile sexism; PP = protective paternalism; CGD = complementary gender differentiation; SCQ = Stigma Consciousness Questionnaire; STEM = science, technology, engineering, and mathematics; STEM # = number of STEM courses taken.
*$p < .05$. **$p < .01$. ***$p < .001$.

on scales of 1 (*strongly disagree*) to 5 (*strongly agree*). Pinel used exploratory factor analysis to confirm that the 10 items reflect a unidimensional construct, and she established convergent and discriminant validity by demonstrating a positive correlation of the SCQ with self-consciousness and a negative correlation with a validated measure of sexism. The scale's test–retest reliability was reported as .76. We computed mean scores for the current sample ($\alpha = .79$), with higher scores indicating more stigma consciousness.

*STEM GPA.* Women reported their average final grades in each type of STEM class they took at the university level ("What is your average final grade in the science [technology, engineering, math, medicine/health sciences] classes you have taken at the university level?"), and we averaged across all grades ($\alpha = .73$). Thus, GPA estimates were based upon the total number of STEM classes taken by each female student, which ranged from 4 to 26 ($M = 11.48$).

*Sexism attitudes.* Men completed the 22-item ASI, which assesses benevolent and hostile sexist attitudes (Glick & Fiske, 1996). A sample hostile sexism item is "Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for equality," and a sample benevolent sexism item is "Women should be cherished and protected by men." All items are rated on scales of 1 (*strongly disagree*) to 6 (*strongly agree*). Researchers using this scale typically create one 11-item hostile sexism composite and one 11-item benevolent sexism composite that collapse across the three benevolent subfactors of protective paternalism (4 items), complementary gender differentiation (3 items), and heterosexual intimacy (4 items). These dimensions were confirmed by Glick and Fiske (1996) via exploratory and confirmatory factor analyses. Convergent validity was established by showing positive correlations between hostile sexism and four other validated sexism scales, while benevolent sexism was unrelated to previous measures of sexism because this was the first measure of sexism to assess affectively positive attitudes toward women. Benevolent sexism, hostile sexism, and Recognition of Discrimination (a translation of Katz and Hass's, 1988, pro-Black scale into a pro-gender scale) emerged as separate factors in an exploratory factor analysis demonstrating the ASI's discriminant validity.

Following Glick and Fiske (1996), we created hostile sexism ($\alpha = .84$) and benevolent sexism ($\alpha = .74$) scores. However, because we were interested in examining women's experiences of protective paternalism and complementary gender differentiation separately, we also created protective paternalism and complementary gender differentiation subscales with the men's data ($\alpha$s = .60 and .74, respectively). Although the $\alpha$ for protective paternalism was low, all of the items in this composite were positively correlated ($r$s ranging from 0.14 to 0.58), and the mean inter-item correlation was 0.28.

## Results

### Preliminary Analyses

Table 3 shows correlations and descriptive statistics for all study variables. Because the number of STEM courses taken correlated with our dependent measures, we controlled for it in analyses on STEM outcomes. We also covaried STEM GPA in analyses on STEM major intentions and STEM self-efficacy, to control for the possibility that low-performing women report more sexism because they do, in fact, elicit more hostile treatment or receive more paternalistic help. As noted, we also controlled for stigma consciousness in all analyses on women's data. Age and class year also correlated with one dependent measure each, but the present analyses do not include these as covariates. However, all findings reported here remained the same when we entered these two variables as covariates.

Missing data analyses were conducted on all of the primary variables. In the women's data, less than 0.7% of all items for all cases were missing, and 66.7% of variables were not missing data for any case. Overall, 95.4% of women participants had no missing data, and no item had more than 5% of missing values. Men had no missing data.

## Female STEM Students' Perceptions of Hostile and Benevolent Sexism

To test Hypothesis 1, that women would perceive more frequent experiences of benevolent sexism than hostile sexism, we ran a one-way repeated measures analysis of variance (ANOVA) comparing women's perceptions of the three sexism types (hostile sexism vs. protective paternalism vs. complementary gender differentiation). This analysis revealed a significant main effect, $F(2, 1188) = 343.31$, $p < .001$, $\eta_p^2 = 0.37$. Consistent with Hypothesis 1, women reported more complementary gender differentiation ($M = 2.37$, $SE = .04$) and protective paternalism ($M = 1.89$, $SE = .04$) than hostile sexism ($M = 1.50$, $SE = .03$). The average frequency of perceived complementary gender differentiation (2.36) corresponded to three to four episodes per month, the average frequency of perceived protective paternalism (1.89) corresponded to two to three episodes per month, and the average frequency of perceived hostile sexism (1.56) corresponded to zero to two episodes per month. Post-hoc least significant difference tests revealed that complementary gender differentiation was perceived more frequently than protective paternalism, $p < .001$, 95% CI [.41, .55], and protective paternalism was perceived more frequently than hostile sexism, $p < .001$, 95% CI [.34, .44]. This analysis remained significant when controlling for women's stigma consciousness, $F(2, 1184) = 8.92$, $p < .001$, $\eta_p^2 = 0.02$, as did the pattern of means and each of the post-hoc tests, $p$s < .001.

## Male STEM Students' Sexism Attitudes

Hypothesis 2 stated that the male STEM classmates of the female participants would endorse benevolent sexism attitudes more strongly than hostile sexism attitudes. A paired samples $t$-test revealed that the men endorsed benevolent sexism ($M = 3.67$, $SD = 0.73$) more strongly than hostile sexism ($M = 3.46$, $SD = 0.80$), $t(162) = 3.02$, $p < .01$. To examine more closely whether men's attitudes reflected the specific pattern of sexism that women reported experiencing, we ran a one-way repeated measures ANOVA comparing men's endorsement of the three sexism types (hostile sexism vs. protective paternalism vs. complementary gender differentiation). This analysis revealed a significant main effect of sexism type, $F(2, 159) = 35.45$, $p < .001$, $\eta_p^2 = 0.31$, and post-hoc tests revealed that men reported stronger endorsement of protective paternalism attitudes ($M = 4.03$, $SE = .07$) than hostile sexism attitudes ($M = 3.46$, $SE = .06$), $p < .01$, 95% CI [.40, .74]. However, men's endorsement of complementary gender differentiation attitudes ($M = 3.32$, $SE = .08$) did not differ significantly from their hostile sexism attitudes, $p = .09$. Thus, consistent with Hypotheses 1 and 2, women reported experiencing more protective paternalism than hostile sexism in their STEM courses, and men in these same courses endorsed protective paternalism more strongly than hostile sexism attitudes.

## Predicting STEM Outcomes

We used hierarchical multiple regression analyses and Hayes's (2013) PROCESS macro to test Hypotheses 3 and 4. In each model, we predicted one of the outcome variables from sexism type (protective paternalism, complementary gender differentiation, or hostile sexism), STEM identity, and their interaction. Simple slope analyses were conducted at $\pm 1$ standard deviation from the mean on each variable. Covariates were the number of STEM courses taken, stigma consciousness, and the two sexism types that were not treated as predictors (to examine unique effects of each sexism type). In analyses on major intentions and self-efficacy, we controlled for STEM GPA to rule out the possibility that low-achieving women drove the findings. Results for Hypotheses 3 and 4 without covariates remained largely unchanged.

*Sexism direct effects.* Hypothesis 3 stated that protective paternalism would predict lower STEM major intentions, STEM self-efficacy, and STEM GPA among women; in contrast, we treated hostile sexism and complementary gender differentiation as exploratory predictors of STEM outcomes. Model 2 in Tables 4 and 5 shows tests of direct effects of protective paternalism, complementary gender differentiation, and hostile sexism on the outcome variables. As shown in Tables 4 and 5, protective paternalism negatively predicted STEM self-efficacy ($p = .03$, 95% CI [−.18, −.007], and hostile sexism negatively predicted intentions to major in STEM ($p = .02$, 95% CI [−.31, −.02], but no other direct effects emerged for protective paternalism or hostile sexism on any outcome variable. Stronger STEM identity predicted major intentions, self-efficacy, and GPA ($p$s < .05). Moreover, because no interactive effects of complementary gender differentiation emerged on any outcome variables (all $t$s < 1.58, $p$s > .11), we did not include a separate table displaying the output from analyses in which complementary gender differentiation was treated as a primary predictor. However, complementary gender differentiation was a significant positive predictor of STEM self-efficacy in Model 2, $t = 2.39$, $p = .02$, 95% CI [.01, .13]. Therefore, Hypothesis 3 was partially supported.

*Moderation by STEM identity.* Hypothesis 4 stated that protective paternalism would more strongly predict negative

**Table 4.** Summary of Regression Analyses for Protective Paternalism Predicting STEM Outcomes.

| Outcomes | Variable | Model 1 | | | | Model 2 | | | | Model 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *B* | *SE*(*B*) | β | *t* | *B* | *SE*(*B*) | β | *t* | *B* | *SE*(*B*) | β | *t* |
| Intentions to major in STEM | STEM # | 0.09 | 0.01 | .29 | 7.08** | 0.06 | 0.01 | .19 | 5.32** | 0.06 | 0.01 | .20 | 5.48** |
| | GPA | −0.04 | 0.08 | −.02 | −0.46 | −0.12 | 0.07 | −.06 | −1.77 | −0.14 | 0.07 | −.07 | −2.03* |
| | SCQ | −0.05 | 0.07 | −.03 | −0.82 | 0.05 | 0.06 | .03 | 0.79 | 0.05 | 0.06 | .03 | 0.78 |
| | CGD | 0.05 | 0.04 | .05 | 1.06 | 0.01 | 0.05 | .01 | 0.26 | 0.01 | 0.05 | .01 | 0.32 |
| | HS | −0.18 | 0.06 | −.15 | −3.05** | −0.16 | 0.08 | −.13 | −2.27* | −0.18 | 0.08 | −.14 | −2.50* |
| | STEM ID | | | | | 0.49 | 0.04 | .49 | 13.14** | 0.33 | 0.04 | .32 | 4.71** |
| | PP | | | | | −0.08 | 0.07 | −.09 | −1.26 | −.09 | 0.07 | −.10 | −1.40 |
| | ID × PP | | | | | | | | | 0.09 | 0.03 | .19 | 2.71** |
| | $R^2$ change | | .10 | | | | .22 | | | | .01 | | |
| | *F* for $R^2$ change | | 12.03** | | | | 87.43** | | | | 7.34** | | |
| STEM self-efficacy | STEM # | 0.03 | 0.01 | .14 | 3.59** | 0.02 | 0.01 | .09 | 2.23* | 0.02 | 0.01 | .09 | 2.33* |
| | GPA | 0.29 | 0.05 | .24 | 6.07** | 0.26 | 0.05 | .22 | 5.54** | 0.25 | 0.05 | .21 | 5.31** |
| | SCQ | 0.04 | 0.04 | .04 | 0.96 | 0.09 | 0.04 | .09 | 2.12* | 0.09 | 0.04 | .09 | 2.12* |
| | CGD | 0.06 | 0.03 | .13 | 2.29* | 0.08 | 0.03 | .12 | 2.39* | 0.08 | 0.03 | .13 | 2.44* |
| | HS | −0.13 | 0.04 | −.17 | −3.54** | −0.10 | 0.06 | −.10 | −1.54 | −0.08 | 0.06 | −.11 | −1.70 |
| | STEM ID | | | | | 0.18 | 0.03 | .27 | 6.81** | 0.09 | 0.03 | .14 | 1.91 |
| | PP | | | | | −0.09 | 0.05 | −.16 | −2.13* | −0.10 | 0.04 | −.17 | −2.24* |
| | ID × PP | | | | | | | | | 0.05 | 0.02 | .15 | 2.02* |
| | $R^2$ change | | .11 | | | | .08 | | | | .006 | | |
| | *F* for $R^2$ change | | 13.65** | | | | 25.75** | | | | 4.09* | | |
| STEM GPA | STEM # | 0.01 | 0.01 | .06 | 1.42 | 0.01 | 0.01 | .04 | 0.90 | 0.01 | 0.01 | .04 | 1.01 |
| | SCQ | 0.07 | 0.04 | .09 | 1.91 | 0.09 | 0.04 | .11 | 2.38* | 0.08 | 0.04 | .10 | 2.35* |
| | CGD | 0.02 | 0.03 | .04 | 0.68 | 0.04 | 0.03 | .07 | 1.23 | 0.04 | 0.03 | .07 | 1.28 |
| | HS | −0.02 | 0.03 | −.04 | −0.69 | 0.02 | 0.05 | .04 | 0.58 | 0.02 | 0.05 | .03 | 0.37 |
| | STEM ID | | | | | 0.05 | 0.02 | .09 | 2.03* | 0.04 | 0.04 | .08 | −0.98* |
| | PP | | | | | −0.07 | 0.04 | −.14 | −1.71 | −0.07 | 0.04 | −.15 | −1.82 |
| | ID × PP | | | | | | | | | 0.05 | 0.02 | .19 | 2.41* |
| | $R^2$ change | | .01 | | | | .01 | | | | .01 | | |
| | *F* for $R^2$ change | | 1.69 | | | | 3.58* | | | | 5.81* | | |

*Note.* STEM = science, technology, engineering, and mathematics; STEM # = number of STEM courses taken; SCQ = Stigma Consciousness Questionnaire; GPA = grade point average; CDG = complementary gender differentiation; PP = protective paternalism; STEM ID = STEM identity; HS = hostile sexism. *$p$ < .05. **$p$ < .01.

STEM outcomes for women who were weakly STEM-identified than women who were strongly STEM-identified. Supporting Hypothesis 4, STEM identity moderated the associations of protective paternalism with major intentions, $p$ = .007, 95% CI [.02, .15]; see Figure 1a, such that protective paternalism negatively predicted major intentions among weakly-identified women, $B = -.17$, $SE = .78$, $t = -2.39$, $p = .01$, 95% CI [−.32, −.03,], but not among strongly-identified women ($p > .95$). Similarly, STEM identity moderated the associations of protective paternalism with STEM self-efficacy, $p = .04$, 95% CI [.001, .09]; see Figure 1b. Weakly-identified women who experienced more protective paternalism reported lower STEM self-efficacy, $B = -.14$, $SE = .05$, $t = -2.84$, $p = .01$, 95% CI [−.24, −.04], but protective paternalism was unrelated to self-efficacy among strongly-identified women ($p > .25$). Last, STEM identity moderated the associations of protective paternalism with STEM GPA, $p = .01$, 95% CI [.01, .09]; see Figure 1c. Protective paternalism experiences predicted lower GPA among weakly-identified women ($B = -.12$, $SE = .05$, $t = -2.64$, $p < .01$, 95% CI [−.21, −.03], but not among strongly-identified women ($p > .73$).

We used the Johnson–Neyman technique (Spiller, Fitzsimons, Lynch, & McClelland, 2013) to locate the values of STEM identity at which the associations of protective paternalism with each outcome became statistically significant. Protective paternalism significantly predicted lower STEM major intentions among women with STEM identity scores below 2.96 on a 5-point scale ($b_{JN} = -0.13$, $SE = .07$, $p = .05$), it predicted lower STEM self-efficacy among women with STEM identity scores below 3.69 ($b_{JN} = -0.09$, $SE = .04$, $p = .05$), and it predicted lower STEM GPA among women with STEM identity scores below 3.30 ($b_{JN} = -0.08$, $SE = .04$, $p = .05$). These scores fall at the 21st, 60th, and 38th percentiles of STEM identity strength, respectively. Thus, Hypothesis 4 received strong support.

Next, at Step 3, STEM identity did not moderate the associations of hostile sexism with major intentions ($p = .17$) or self-efficacy ($p = .12$), but a significant Hostile Sexism × STEM Identity interaction emerged on STEM GPA

**Table 5.** Summary of Regression Analyses for Hostile Sexism Predicting STEM Outcomes.

| | | Model 1 | | | | Model 2 | | | | Model 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Outcomes | Variable | *B* | *SE*(*B*) | β | *t* | *B* | *SE*(*B*) | β | *t* | *B* | *SE*(*B*) | β | *t* |
| Intentions to major in STEM | STEM # | .08 | 0.01 | .28 | 6.91** | .06 | 0.01 | .19 | 5.32** | 0.06 | 0.01 | .19 | 5.35** |
| | GPA | −.05 | 0.08 | −.02 | −0.59 | −.12 | 0.07 | −.06 | −1.77 | −0.13 | 0.07 | −.07 | −1.90 |
| | SCQ | −.03 | 0.07 | −.02 | −0.51 | .05 | 0.06 | .03 | 0.78 | 0.05 | 0.06 | .03 | 0.77 |
| | CGD | .08 | 0.06 | .09 | 1.64 | .01 | 0.05 | .01 | 0.03 | 0.01 | 0.05 | .01 | 0.28 |
| | PP | −.16 | 0.05 | −.17 | −3.06** | −.09 | 0.07 | −.09 | −1.26 | −0.08 | 0.07 | −.08 | −1.20 |
| | STEM ID | | | | | .49 | 0.04 | .49 | 13.14** | 0.39 | 0.04 | .39 | 4.94** |
| | HS | | | | | −.16 | .07 | −.13 | −2.27* | −0.19 | 0.07 | −.15 | −2.54* |
| | ID × HS | | | | | | | | | 0.07 | 0.05 | .11 | 1.38 |
| | $R^2$ change | | .10 | | | | .22 | | | | .002 | | |
| | *F* for $R^2$ change | | 12.04** | | | | 87.41** | | | | 1.91 | | |
| STEM self-efficacy | STEM # | .03 | 0.01 | .14 | 3.40** | .02 | 0.01 | .09 | 2.23** | 0.02 | 0.01 | .09 | 2.25* |
| | GPA | .29 | 0.05 | .24 | 5.89** | .26 | 0.05 | .22 | 5.53** | 0.25 | 0.05 | .21 | 5.34** |
| | SCQ | .06 | 0.04 | .06 | 1.37 | .09 | 0.04 | .09 | 2.12* | 0.09 | 0.05 | .09 | 2.10* |
| | CGD | .10 | 0.03 | .17 | 3.13** | .08 | 0.03 | .13 | 2.39* | 0.08 | 0.03 | .13 | 2.31* |
| | PP | −.13 | 0.03 | −.23 | −4.02** | −.09 | 0.05 | −.16 | −2.13* | −0.09 | 0.05 | −.15 | −2.08+ |
| | STEM ID | | | | | .18 | 0.03 | .27 | 6.81** | 0.18 | 0.03 | .17 | 2.00* |
| | HS | | | | | −.08 | 0.05 | −.10 | −1.54 | −0.10 | 0.06 | −.12 | −1.82 |
| | ID × HS | | | | | | | | | 0.05 | 0.03 | .12 | 1.35 |
| | $R^2$ change | | .12 | | | | .07 | | | | .003 | | |
| | *F* for $R^2$ change | | 14.47** | | | | 23.82** | | | | 1.82 | | |
| STEM GPA | STEM # | .00 | 0.00 | .06 | 1.37 | .00 | 0.00 | .04 | 0.90 | 0.006 | 0.007 | .04 | 0.92 |
| | SCQ | .08 | 0.04 | .10 | 2.15* | .09 | 0.04 | .11 | 2.38* | 0.08 | 0.04 | .10 | 2.32* |
| | CGD | .04 | 0.03 | .08 | 1.47 | .04 | 0.03 | .07 | 1.23 | 0.04 | 0.03 | .08 | 1.26 |
| | PP | −.05 | 0.03 | −.10 | −1.75 | −.06 | 0.04 | −.14 | −1.71 | −0.06 | 0.04 | −.13 | −1.59 |
| | STEM ID | | | | | .05 | 0.02 | .09 | 2.01* | −0.07 | 0.05 | .13 | 1.36 |
| | HS | | | | | .00 | 0.05 | .04 | 0.56 | −0.003 | 0.05 | −.004 | −0.06 |
| | ID × HS | | | | | | | | | 0.08 | 0.03 | .25 | 2.61** |
| | $R^2$ change | | .02 | | | | .008 | | | | .01 | | |
| | *F* for $R^2$ change | | 2.35* | | | | 2.27 | | | | 6.83* | | |

*Note.* STEM = science, technology, engineering, and mathematics; STEM # = number of STEM courses taken; SCQ = Stigma Consciousness Questionnaire; GPA = grade point average; CDG = complementary gender differentiation; PP = protective paternalism; STEM ID = STEM identity; HS = hostile sexism.
*$p$ < .05. **$p$ < .01.

($p$ = .009). Hostile sexism experiences were not significantly related to GPA among either weakly identified women ($p$ = .19) or strongly identified women ($p$ = .12). However, the Johnson–Neyman technique revealed that hostile sexism exposure significantly predicted lower STEM GPA among women with STEM identity scores below 1.28 ($b_{JN}$ = −0.17, $SE$ = .09, $p$ = .05), which falls at the third percentile of STEM identity strength. Thus, hostile sexism experiences negatively predicted STEM GPA among women who were exceedingly low in STEM identity strength.

## Discussion

This study is the first to our knowledge that examines perceived hostile and benevolent sexism experiences of women in university STEM courses, as well as the independent associations of these sexism types with women's STEM outcomes. Consistent with our expectations, women reported experiencing more protective paternalism and complementary gender differentiation (benevolent sexism) than hostile sexism in their STEM courses and course-related activities, even when controlling for individual differences in their expectations of being gender stereotyped. Moreover, male students from the same sample of STEM courses endorsed protective paternalism more strongly than hostile sexism, suggesting that women's perceptions of their gender-based treatment correspond to men's self-reported sexism attitudes. Of note, women who reported more frequent experiences with protective paternalism (but not complementary gender differentiation) also reported lower STEM major intentions, STEM self-efficacy, and STEM GPA, but these effects only emerged for weakly identified women. More frequent hostile sexism exposure also predicted lower intentions to major in STEM and lower GPA in STEM courses, but the latter effect only emerged among very weakly STEM-identified women. Overall, women who identified more strongly with STEM fields showed no associations between perceived sexism and STEM outcomes. This pattern is consistent with previous research that demonstrated that individuals with a weaker domain identity lose self-efficacy and withdraw effort earlier

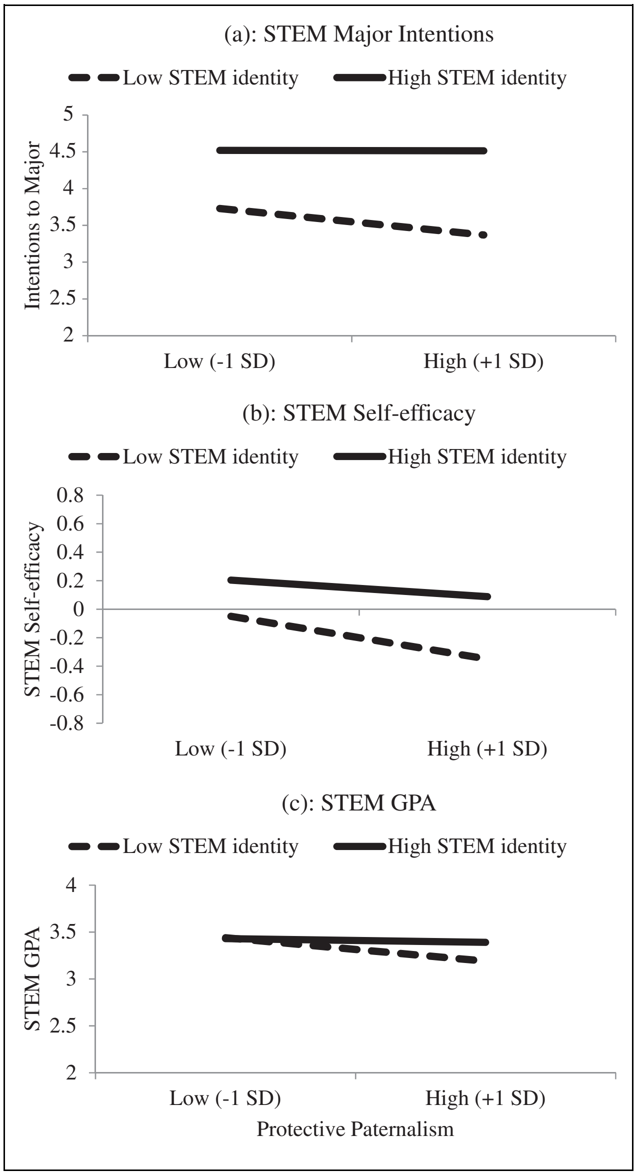


**Figure 1.** Interactions of protective paternalism and science, technology, engineering, and mathematics (STEM) identity on STEM major intentions, STEM self-efficacy, and STEM GPA.

in the face of increasing difficulty compared with those with a stronger domain identity (Brehm & Self, 1989; Gendolla & Richter, 2010).

Our findings suggest that benevolent sexism—in the form of protective paternalism—interacts with STEM identity to predict pernicious outcomes for women in university STEM courses, while complementary gender differentiation and hostile sexism are less predictive of harmful outcomes; in fact, complementary gender differentiation is a positive predictor of women's STEM self-efficacy. The relatively weak links between hostile sexism and negative outcomes may at first seem counterintuitive. Hostile sexism creates a stressful work environment that makes women feel unwelcome, and its harmful effects in general cannot be overlooked. Prolonged exposure to such treatment can undermine women's engagement, cause job burnout, and impair their mental and physical health (Fitzgerald, 1993). However, we did not examine prolonged exposure to hostile treatment, but merely self-reported exposure in university STEM courses among relatively young women. Moreover, our participants reported experiencing hostile sexism in their STEM courses relatively rarely (zero to two episodes per month), suggesting low baseline exposure to this type of treatment. In the short-term, hostile sexism may motivate women to prove insulting stereotypes wrong and push for social change (Becker & Wright, 2011). Furthermore, if the campus climate has programs that promote diversity awareness and support, it may reduce the psychological impact of hostile sexism by offering college women numerous resources for coping with it. Perhaps over a longer time, or in a less supportive environment, these low-frequency occurrences add up and the effort needed to repeatedly counteract hostile sexism eventually may predict disengagement. Further, hostile sexism may influence women's underrepresentation in STEM through more structural mechanisms (e.g., instructors and mentors who hold hostilely sexist attitudes) compared to interpersonal experiences with peers. Longitudinal studies and studies in other STEM workplace environments are needed to more fully examine the associations of hostile sexism with women's STEM outcomes.

Protective paternalism—which women reported experiencing 2–3 times per month in their STEM courses—was consistently related to less STEM engagement and worse STEM outcomes among women who identified weakly with STEM. Specifically, weakly STEM-identified women disengaged from STEM fields, doubted their STEM abilities, and performed more poorly in STEM courses to the extent that they received more protective, paternalistic treatment. This may occur because benevolent sexism activates thoughts of incompetence, which usurps cognitive resources (Dardenne et al., 2007). Thus, repeated exposure to protective paternalism may chip away at the STEM confidence and commitment of women for whom STEM pursuits are not centrally defining. Conversely, for women high in STEM identity, protective paternalism was unassociated with STEM outcomes. Perhaps these women's greater investment in STEM pursuits allowed them to maintain their motivation despite treatment that implied that they were weaker than men and needed extra assistance.

Of note, complementary gender differentiation—which was the sexism type most frequently reported by women—did not interact with STEM identity to predict any negative outcomes. Women who reported experiencing more complementary gender differentiation in their STEM classes also reported higher levels of STEM self-efficacy (see Tables 4 and 5, Model 2), but complementary gender differentiation was not related to any other STEM outcomes. Thus, being treated as if they were "warm" and "morally virtuous" did not appear to convey to women negative stereotypes about their lack of competence in STEM fields and, if anything, it predicted increases in women's STEM confidence. This unexpected finding should be interpreted with caution because complementary gender differentiation was not consistently related to women's STEM outcomes.

We treated performance (STEM GPA) as an outcome variable in the current study, but it is possible that women who perform more poorly in STEM courses elicit more protective paternalism and hostile sexism from their classmates, instructors, and teaching assistants. There are two reasons that we do not view this as a viable alternative explanation for the findings. First, we only found associations between perceived sexism and GPA among weakly identified women, and it is not clear why only these women would elicit more sexism for performing poorly. Second, we controlled for GPA in analyses predicting STEM intentions and self-efficacy and still found the predicted patterns.

## Limitations

The conclusions drawn from our findings are limited by the cross-sectional and self-report nature of the data. A more rigorous test of our hypotheses requires longitudinal assessment with sexism exposure in current STEM courses predicting future STEM outcomes. Such a design would also allow for direct tests of proposed mechanisms, including women's felt belonging in STEM environments, rumination about their STEM competence, and motivation to prove gender stereotypes wrong. Tracking women's experience with sexism over time would allow for a more direct test of the frequency of hostile and benevolent sexism in STEM environments. Also, as noted above, longitudinal assessment may reveal more negative effects of hostile sexism that accrue over time. Finally, perhaps protective paternalism does undermine performance and outcomes for highly STEM-identified women, but this effect is only detectable over a longer time. However, given that this is the first attempt to measure women's self-reported exposure to hostile and benevolent sexism separately and examine their associations with outcomes in daily life, we believe that our findings have important implications for understanding and increasing women's engagement in STEM disciplines.

Our conclusions regarding the lack of findings for complementary gender differentiation predicting negative STEM outcomes may be limited by the composition of our measure of complementary gender differentiation. Only 1 of the 4 items in our Complementary Gender Differentiation subscale makes a direct comparison between women and men (see Table 1). In contrast, items on the ASI (Glick & Fiske, 1996) that assess complementary gender differentiation explicitly compare women to men (e.g., "Women, compared to men, tend to have a superior moral sensibility"). Wording that explicitly contrasts women and men may be necessary in order to measure women's experiences with sexist, gender-based treatment as opposed to other, more benign forms of treatment. Although women in our pilot study interpreted the entire set of benevolent sexism items as reflecting sexism, we did not measure their perceptions of the protective paternalism and complementary gender differentiation items separately. Therefore, it is possible that the complementary gender differentiation items measure women's experiences with being treated in a genuinely flattering—rather than condescendingly sexist—manner. This could explain why scores on this scale did not predict negative STEM outcomes for women in the current study. Research currently underway in our lab therefore focuses on fine-tuning these items and further establishing the scales' validity and reliability.

Future research should also examine how the gender composition of STEM courses influences women's experiences with hostile and benevolent sexism. On the one hand, unequal gender ratios that favor male students may increase the salience of gender stereotypes about women's STEM competence, thus increasing the frequency with which women STEM students encounter hostile and benevolent sexism in their courses. On the other hand, sexist treatment that occurs in the context of heavily male-dominated environments may be easier to attribute to discrimination, which may reduce its impact on women's STEM outcomes. Future research may profit from examining how STEM instructors' endorsement of sexist attitudes translates into women's perceptions of sexist treatment. Last, conclusions regarding men's endorsement of protective paternalism should be interpreted with caution due to the low internal reliability of the scale for this sample. And the results for men's endorsement of hostile and benevolent sexism should also be interpreted with caution due to social desirability concerns.

## Practice Implications

These findings have practical implications for advisors and instructors of STEM courses. While instructors who wish to encourage women in STEM may view it as helpful to offer women extra assistance, the findings presented here suggest that such assistance is correlated with negative outcomes for women. We are reminded of an invitation sent out by the Office of Diversity at a large public university that advertised a "Research Boot Camp for Women" to help "women students develop the skills necessary for success in academia." Although undoubtedly well-intended, this invitation inadvertently communicated negative stereotypes about women's relative lack of competence. Moreover, the women who are most likely to lose confidence in the face of such paternalism may be women who are not yet firmly committed to STEM fields. This creates somewhat of a conundrum for STEM instructors: How does one encourage and support undecided female STEM students who are struggling, without simultaneously conveying negative gender stereotypes and sexist attitudes? One way may be to avoid messages that are directed only at women—like the Research Boot Camp described above—and instead convey clearly that help is offered to all struggling students based only on performance. Thus, STEM course instructors should be cognizant of how seemingly pro-social efforts to help advance women in STEM can have unintended consequences.

## Conclusions

While people generally disapprove of hostile sexism, those wishing to increase women's STEM involvement may view protective paternalism as appropriate or even helpful because it offers special accommodations and assistance to women in STEM environments. Our findings indicate that this sort of treatment may instead backfire by undermining women's self-efficacy, sense of engagement, and performance. In short, protective paternalism may lead women to disengage from STEM disciplines, especially if they are not yet fully committed to the pursuit of STEM. Ironically, a type of sexism that women are relatively more likely to encounter in STEM environments—protective paternalism—may erode their interests in STEM and further their underrepresentation in these male-dominated fields. Thus, we caution STEM educators and students to avoid treating women and other underrepresented individuals with paternalism. Instead, assistance should be offered to students on the basis of their performance, not on their gender.

### Acknowledgment

We would like to thank Scott Campbell, Mary Goodwin, Kate Johnson, and Will Brazelle for allowing us to collect data in their classes.

### Declaration of Conflicting Interests

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

### Funding

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

### References

Baron, A. S., Schmader, T., Cvencek, D., & Meltzoff, A. (2014). The gendered self-concept: How implicit gender stereotypes and attitudes shape self-definition. In P. J. Leman & H. R. Tenenbaum (Eds.), *Gender and development* (pp. 109–132). New York, NY: Psychology Press. doi:10.4324/9781315870878

Barreto, M., & Ellemers, N. (2005). The burden of benevolent sexism: How it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology*, *35*, 633–642. doi:10.1002/ejsp.270

Becker, J. C., & Wright, S. C. (2011). Yet another dark side of chivalry: Benevolent sexism undermines and hostile sexism motivates collective action for social change. *Journal of Personality and Social Psychology*, *101*, 62–77. doi:10.1037/a0022615

Bernstein, B. L., & Russo, N. F. (2008). Explaining too few women in STEM careers: A psychosocial perspective. In M. Paludi (Ed.), *The psychology of women at work: Challenges and solutions for our female workforce: Vol. 2. Obstacles and the identity juggle* (pp. 1–33). Westport, CT: Praeger. doi:10.5860/choice.46-3941

Bonnot, V., & Croizet, J. C. (2007). Stereotype internalization and women's math performance: The role of interference in working memory. *Journal of Experimental Social Psychology*, *43*, 857–866. doi:10.1016/j.jesp.2006.10.006

Bosson, J. K., Pinel, E. C., & Vandello, J. A. (2009). The emotional impact of ambivalent sexism: Forecasts versus real experiences. *Sex Roles*, *62*, 520–531. doi:10.1007/s11199-009-9664-y

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Reviews*, *40*, 109–131. doi:10.1146/annurev.psych.40.1.109

Calogero, R. M., & Jost, J. T. (2011). Self-subjugation among women: Exposure to sexist ideology, self-objectification, and the protective function of the need to avoid closure. *Journal of Personality and Social Psychology*, *100*, 211–228. doi:10.1037/a0021864

Cattell, R. B. (1966). The Scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245–276.

Ceci, S. J., & Williams, W. M. (Eds.). (2007). *Why aren't more women in science? Top researchers debate the evidence*. Washington, DC: APA Books. doi:10.1037/11546-000

Chaudoir, S. R., & Quinn, D. M. (2010). Revealing concealable stigmatized identities: The impact of disclosure motivations and positive first-disclosure experiences on fear of disclosure and well-being. *Journal of Social Issues*, *66*, 570–584. doi:10.1111/j.1540-4560.2010.01663.x

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Social Issues*, *67*, 469–491. doi:10.1111/j.1540-4560.2011.01710.x

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

Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education*, *16*, 541–554. doi:10.1007/s11218-013-9232-8

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: Consequences for women's performance. *Journal of Personality and Social Psychology*, *93*, 764–779. doi:10.1037/0022-3514.93.5.764

Dardenne, B., Dumont, M., Sarlet, M., Phillips, C., Balteau, E., Degueldre, C., . . . Collette, F. (2013). Benevolent sexism alters executive brain responses. *Neuroreport for Rapid Communication of Neuroscience Research*, *24*, 572–577. doi:10.1097/WNR.0b013e3283625b5b

Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3–32). Newbury Park, CA: Sage.

Faulkner, W. (2009). Doing gender in engineering workplace cultures. I. Observations from the field. *Engineering Studies*, *1*, 3–18. doi:10.1080/19378620902721322

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902. doi:10.1037//0022-3514.82.6.878

Fitzgerald, L. F. (1993). Sexual harassment: Violence against women in the workplace. *American Psychologist*, *48*, 1070–1076. doi:10.1037/0003-066X.48.10.1070

Fogliati, V. J., & Bussey, K. (2013). Stereotype threat reduces motivation to improve effects of stereotype threat and feedback on women's intentions to improve mathematical ability. *Psychology of Women Quarterly*, *37*, 310–324. doi:10.1177/0361684313480045

Gendolla, G. H., & Richter, M. (2010). Effort mobilization when the self is involved: Some lessons from the cardiovascular system. *Review of General Psychology*, *14*, 212. doi:10.1037/a0019742

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491–512. doi:10.1037/0022-3514.70.3.491

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, *56*, 109–118. doi:10.1037/0003-066x.56.2.109

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153–166. doi:10.1007/s11199-011-9996-2

Hartman, H., & Hartman, M. (2008). How undergraduate engineering students perceive women's (and men's) problems in science, math, and engineering. *Sex Roles*, *58*, 251–265. doi:10.1007/s11199-007-9327-9

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press. doi:10.1111/jedm.12050

Hill, C., Corbett, C., & Rose, A, St. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: American Association of University Women. doi:10.5195/jyd.2014.44

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893–905. doi:10.1037/t14267-000

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' math performance. *British Journal of Educational Psychology*, *77*, 323–338. doi:10.1348/000709906x113662

King, E. B., Botsford, W., Hebl, M. R., Kazama, S., Dawson, J. F., & Perkins, A. (2012). Benevolent sexism at work: Gender differences in the distribution of challenging developmental experiences. *Journal of Management*, *38*, 1835–1866. doi:10.1177/0149206310365902

Klonoff, E. A., & Landrine, H. (1995). The schedule of sexist events. *Psychology of Women Quarterly*, *19*, 439–470. doi:10.1111/j.1471-6402.1995.tb00086.x

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles*, *66*, 220–234. doi:10.1007/s11199-011-0036-z

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., ... Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*, 144–165. doi:10.1037/0022-3514.95.1.144

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*, 302–318. doi:10.1177/0146167292183006

Major, B., McCoy, S., Kaiser, C., & Quinton, W. (2003). Prejudice and self-esteem: A transactional model. *European Review of Social Psychology*, *14*, 77–104. doi:10.1080/10463280340000027

Murphy, M. C., Richeson, J. A., Shelton, J. N., Rheinschmidt, M. L., & Bergsieker, H. B. (2012). Cognitive costs of contemporary prejudice. *Group Processes & Intergroup Relations*, *16*, 560–571. doi:10.1177/1368430212468170

Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314–1334. doi:10.1037/a0012702

Nosek, B. A., Banaji, M., & Greenwald, A. (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

Pettigrew, T. F., & Meertens, R. W. (2006). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, *25*, 7–75. doi:10.1002/ejsp.2420250106

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*, 114–128. doi:10.1037/0022- 3514.76.1.114

Salomon, K., Burgess, K. D., & Bosson, J. K. (2015). Flash fire and slow burn: Women's cardiovascular reactivity and recovery following hostile and benevolent sexism. *Journal of Experimental Psychology: General*, *144*, 469–479. doi:10.1037/xge0000061

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50*, 835–850. doi:10.1023/b: sers.0000029101.74557.a0

Schneider, K. T., Tomaka, J., & Palacios, R. (2001). Women's cognitive, affective, and physiological reactions to a male coworker's sexist behavior. *Journal of Applied Social Psychology*, *31*, 1995–2018. doi:10.1111/j.1559-1816.2001.tb00161.x

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

Sibley, C. G., & Wilson, M. S. (2004). Differentiating hostile and benevolent sexist attitudes toward positive and negative female subtypes. *Sex Roles*, *51*, 687–696. doi:10.1007/s11199-004-0718-x

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

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., , Jr., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research*, *50*, 277–288. doi:10.1509/jmr.12.0420

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629. doi:10.1037//0003-066x.52.6.613

Steinberg, J. R., Okun, M. A., & Aiken, L. S. (2012). Calculus GPA and math identification as moderators of stereotype threat in highly persistent women. *Basic and Applied Social Psychology*, *34*, 534–543. doi:10.1080/01973533.2012.727319

Thompson, A. R., & Gregory, A. (2011). Examining the influence of perceived discrimination during African American adolescents' early years of high school. *Education and Urban Society*, *43*, 3–25. doi:10.1177/0013124510379827

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, England: Basil Blackwell. doi:10.2307/2073157

Vescio, T. K., Gervais, S. J., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology*, *88*, 658–672. doi:10.1037/0022-3514.88.4.658

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467. doi:10.1016/S0022-1031(03)00019-2