Psychology of Women Quarterly
1-17

DOI: 10.1177/0361684317720186
journals.sagepub.com/home/pwq
SAGE

Fitting in to Move Forward: Belonging, Gender, and Persistence in the Physical Sciences, Technology, Engineering, and Mathematics (pSTEM)

Karyn L. Lewis[1,2], Jane G. Stout[3], Noah D. Finkelstein[4], Steven J. Pollock[4], Akira Miyake[1], Geoff L. Cohen[5], and Tiffany A. Ito[1]

Abstract

Social science researchers have increasingly focused on understanding the precursors to gender disparities favoring men in the physical sciences, technology, engineering, and mathematics (pSTEM). In the current work, we hypothesized that the core social need to belong explains persistence in pSTEM for women more so than for men. We conducted three field studies with data from close to 3,000 participants bridging a wide span of higher education levels and differing pSTEM fields. In each study, we found gender disparities on sense of belonging in pSTEM favoring men. Moreover, sense of belonging explained persistence intentions for both women and men in one study and explained persistence intentions and actual persistence in pSTEM coursework for women, more so than for men, in the other two studies, even after controlling for two conventional predictors of academic achievement (self-efficacy and exam performance). These results highlight the role of belonging in gender differences in pSTEM persistence and indicate STEM educators should strive to create inclusive learning environments for all students. *Additional online materials for this article are available on* PWQ's *website at* http://journals.sagepub.com/doi/suppl/10.1177/0361684317720186

Keywords

belonging, self-efficacy, gender, persistence, pSTEM

> It's the slow drumbeat of being underappreciated, feeling uncomfortable, and encountering roadblocks along the path to success. These subtle distinctions help make women feel out of place . . . feeling out of place over and over again eventually soaks in . . . after enough of this kind of thing, women feel beaten down and underappreciated, or worse, they feel incapable.
>
> Urry (2005)

The participation of women in physical science, technology, engineering, and mathematics (pSTEM) still severely lags behind that of men. For instance, in 2012, women constituted a small fraction of undergraduate students who earned degrees in engineering (20%), physics (19%), and computer science (18%; National Science Foundation [NSF], 2015a). Not surprisingly, this pattern is similar in the workforce, where women make up 15% of engineers, 12% of physicists and astronomers, and 24% of computer and information scientists (NSF, 2015b). This representation gap also is accompanied by an achievement gap. For example, women on average earn lower course grades in introductory calculus-based physics courses and score lower on standardized nationally normed surveys of conceptual mastery of physics (Kost-Smith, Pollock, & Finkelstein, 2009; Lorenzo, Crouch, & Mazur, 2006; Pollock, Finkelstein, & Kost-Smith, 2007). These disparities, however, do not extend to all aspects of science. In the social sciences and the life sciences, such as biology, the number of women is at parity with, if not greater than, the number of men (National Science Board, 2012), which highlights the importance of understanding why the pSTEM disciplines, in particular, suffer from gender disparities in accomplishment and retention.

[1] Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA
[2] Center for Research, Evaluation, and Analysis, Education Northwest, Portland, OR, USA
[3] Center for Evaluating the Research Pipeline, Computing Research Association, Washington, DC, USA
[4] Department of Physics, University of Colorado Boulder, Boulder, CO, USA
[5] Department of Psychology, Stanford University, Stanford, CA, USA

Corresponding Author:
Karyn L. Lewis, Center for Research, Evaluation, and Analysis, Education Northwest, 101 SW Main, Ste. 500, Portland, OR 97204, USA.
Email: karyn.lewis@educationnorthwest.org

Belonging and pSTEM Gender Gaps

Given the pervasiveness of gender gaps in pSTEM achievement and persistence, it is likely that the causes are complex and multiply determined. Rather than trying to enumerate all possible causes, in the current study we identify a core psychological process through which multiple factors, such as stereotype threat and lack of women's representation in the field, could operate—namely, one's *sense of belonging* in pSTEM. A sense of belonging is defined as the subjective feeling of fitting in and being included as a valued and legitimate member in a particular setting, such as a pSTEM learning environment (Baumeister & Leary, 1995; Good, Rattan, & Dweck, 2012; Goodenow, 1993; Walton & Cohen, 2007). The importance of a sense of belonging is illustrated by the above comments of lauded astrophysicist Meg Urry. Urry's experience with the consequences of feeling "out of place" echoes decades of social psychological research outlining the fundamental importance of belonging for well-being (Baumeister & Leary, 1995). Unfortunately, recent empirical evidence converges with Urry's comments in showing that in academic contexts, members of underrepresented and negatively stereotyped groups experience greater threats to their belonging, on average, than members of majority groups (Walton & Cohen, 2007). This difference is problematic because students' sense of belonging is a robust predictor of academic motivation, engagement, and achievement in STEM fields, as well as in achievement domains more generally (Hausmann, Schofield, & Woods, 2007; Pittman & Richmond, 2007; Walton, Cohen, Cwir, & Spencer, 2012; Zumbrunn, McKim, Buhs, & Hawley, 2014). We examined a gender difference in sense of belonging in pSTEM and assessed whether that gender difference might explain the gender differences in pSTEM persistence intentions and actual persistence. At the same time, we assessed whether the gender difference in self-efficacy in pSTEM (the belief that one can succeed or accomplish a task; Bandura, 1986) could explain gender differences in pSTEM intentions and actual persistence (discussed below). Thus, in the current work, we aimed to make a unique contribution to the growing body of research explaining gender representation in pSTEM by simultaneously pitting theoretically rooted sources of the phenomenon against each other (belonging and self-efficacy) and pinpointing particularly potent sources of persistence in pSTEM. This examination is critical in order to better refine interventions designed to ameliorate gender gaps.

Situational Influences on Belonging

Research points to a number of situational factors typical of pSTEM settings that lead women to feel lower belonging than men. First, and perhaps most obviously, the dearth of female peers and role-models in pSTEM settings can be a blatant cue about who does and does not belong. Dasgupta (2011) theorizes that in achievement contexts, contact with in-group peers and experts functions as a kind of psychological vaccine that protects people from the negative effects of identity threat and serves to boost sense of belonging. Consistent with this, Murphy and colleagues have shown experimentally that women exposed to typical asymmetric pSTEM gender ratios (i.e., women were far outnumbered by men) report a lower sense of belonging, relative to women exposed to balanced gender ratios (Murphy, Steele, & Gross, 2007).

Another threat to women's sense of belonging in pSTEM settings comes from the mismatch between women's perceptions of themselves and stereotypes about what members of pSTEM fields are like (or should be like). Stereotypes can take the form of the personal characteristics that pSTEM inhabitants are expected to have (e.g., socially inept, loner, self-focused; Schott & Selwyn, 2000), which are often incompatible with typical female gender roles (Diekman, Clark, Johnston, Brown, & Steinberg, 2011; London, Rosenthal, Levy, & Lobel, 2011; Settles, Jellison, & Pratt-Hyatt, 2009). Stereotypes can also take the form of assumptions about the intellectual abilities that pSTEM inhabitants are expected to possess (e.g., innate talent and natural brilliance; Leslie, Cimpian, Meyers, & Freeland, 2015), which stand in stark opposition to negative stereotypes about women's quantitative abilities (Spencer, Steele, & Quinn, 1999). Converging evidence suggests that exposure to these stereotypes challenges women's perceived compatibility with the field (Cheryan, Drury, & Vichayapai, 2012; Cheryan, Plaut, Davies, & Steele, 2009; Good et al., 2012; Smith, Lewis, Hawthorne, & Hodges, 2013). The end result is that women tend to conclude that participation in pSTEM requires them to be different from who they are or who they want to be, which can decrease their sense of belonging in that field.

Gender Differences in Effect of Sense of Belonging on Persistence

Experimental research on the situational determinants of students' sense of belonging in pSTEM suggests women should experience a lower sense of belonging than men. Thus, a gender discrepancy in pSTEM sense of belonging warrants close empirical scrutiny. In the current work, we predicted overall gender differences in pSTEM sense of belonging but also expected women's subjective sense of belonging in pSTEM might affect persistence outcomes to a greater degree than would be the case for men. This prediction aligns with what Hughes, Im, and Allee (2015) refer to as the vulnerability hypothesis; they argue that people who are most at risk of failure in an academic setting will be most affected by their subjective experiences in that setting. Given women's relatively lower representation and performance in pSTEM, the vulnerability hypothesis predicts that sense of belonging will have a bigger effect on pSTEM motivation and achievement for women than for men. Walton and Cohen (2007) proposed, in a similar hypothesis, that members of socially stigmatized groups experience uncertainty as to whether they belong (i.e.,

"belonging uncertainty"), which can lead to greater sensitivity to, and biased interpretation of, ambiguous cues. For instance, receiving a poor quiz grade becomes an indictment of one's abilities overall and an unfriendly lab partner is taken as evidence that one's group is not welcome in general.

Consistent with a vulnerability or belonging uncertainty perspective, Walton and Cohen (2007) found that Black undergraduates who underwent a manipulation that made them think they would have few friends in computer science were more pessimistic about their potential for success in that field, compared to Black students whose sense of belonging was not threatened; the same manipulation had no effect on White students' perceptions of their potential to succeed (i.e., White students were equally optimistic about their potential, regardless of whether they thought they would have friends in the domain or not). Furthermore, Black students' sense of belonging was related to their daily academic adversity (e.g., feeling lonely or stressing over an exam), but White students' sense of belonging was more stable and independent of such events.

Further evidence supporting a vulnerability or belonging uncertainty perspective comes from Murphy and Zirkel (2015), who compared how sense of belonging is linked to outcomes for Black, versus White, students. In a study of middle school students, they found sense of belonging in school predicted the educational aspirations of Black students, but not of White students. In a separate study of college students, Murphy and Zirkel also found a sense of belonging predicted academic achievement (grade point average; GPA) for Black students, but not for White students.

The work of Walton and Cohen (2007) and Murphy and Zirkel (2015) suggests that sense of belonging is tightly linked with educational outcomes for people who are negatively stereotyped in the domain. Although not yet tested among women in pSTEM, these results suggest that women's vulnerable social status in pSTEM may result in a stronger relation between the degree to which they feel they belong in pSTEM and their intentions to persist. In contrast, men should feel a higher sense of belonging. But even when men feel a relatively low sense of belonging, they should still experience protective benefits from being a member of the dominant majority group. That is, as the dominant group in pSTEM, men are part of a group that is subject to positive stereotypes about their ability, so they should be less vulnerable to the effects of low belonging, compared to women (see Walton & Cohen, 2003). We expected this to translate into a weaker relation between sense of belonging and persistence in pSTEM for men in general. This is not to say that men cannot be affected by a low sense of belonging. Rather, the effects of a sense of belonging may be less extreme for men than for women because men's place in pSTEM domain is more culturally sanctioned than is the case for women.

To our knowledge, the current work is the first to extend the work of Walton and Cohen (2007) and Murphy and Zirkel (2015), to test a vulnerability hypothesis with a sample of women in a broad array of pSTEM fields. We sought to understand whether belonging is more predictive of outcomes for women than for men in these domains. There is initial empirical evidence that a sense of belonging in mathematics is linked to intent to persist and performance for undergraduate women. Good, Rattan, and Dweck (2012) found that for women, but not for men, a sense of belonging in calculus classes eroded when women perceived others in their calculus class (e.g., other students, instructors) believed (1) women are less talented than men in math and (2) that math is a fixed ability (as opposed to a malleable skill that can be improved over time). Our current work builds on Good and colleagues' work to (1) examine additional domains of pSTEM where women are even more underrepresented than in math (e.g., computer science, physics) and (2) understand whether a sense of belonging may be differentially linked to actual persistence for women and men. The latter is important to investigate, given extensive research indicating that while intentions are an excellent predictor of actual behavior, behavioral intentions and actual behavior do not always show a 1:1 relation (Ajzen, 1985, 1987, 1991, 2011). Given the ultimate goal of increasing women's actual (and not merely their intended) persistence in pSTEM, the current work extends prior studies on the effects of women's sense of belonging on their intentions to persist in pSTEM, to assess whether women's sense of belonging also relates to behavioral persistence in pSTEM (i.e., course taking) in Study 2.

Sense of Belonging Versus Self-Efficacy

Another variable frequently cited as the root of gender gaps in pSTEM persistence is the so-called confidence gap (Orenstein, 1994), or women's relatively lower self-efficacy than men. Self-efficacy refers to beliefs about one's ability to plan for and execute steps necessary for future success (Bandura, 1986). This construct has received extensive attention in educational psychology as a potential predictor of academic outcomes and gender disparities. For example, Lent, Brown, and Larkin (1986) have shown that self-efficacy promotes academic performance and motivation. Women often report lower pSTEM self-efficacy than men, even when women and men have performed at parity, for example, when they earn statistically equivalent pSTEM course GPAs (Kost-Smith et al., 2009; Stout, Dasgupta, Hunsinger, & McManus, 2011). These findings suggest markedly different experiences in pSTEM among women and men: At best, women hold a more realistic (less optimistic) view of their ability than men; at worst, women have an inaccurately negative view of their ability. This rather grim picture of women's self-efficacy in pSTEM aligns with persistent and pervasive cultural stereotypes that women are "naturally" less gifted in pSTEM fields than men (Miller, Eagly, & Linn, 2015). Other research has found when there is a gender difference in pSTEM performance (favoring men), it can be explained by the fact that women hold lower pSTEM self-efficacy than men (Pajares &

Miller, 1994). This research indicates that self-efficacy is associated with students' achievement and motivation; women's and men's experience of self-efficacy in pSTEM differs in magnitude and quality, raising the possibility that gender differences in self-efficacy underlie gender differences in persistence in pSTEM.

While there is evidence suggesting independent links of belonging and self-efficacy with gender disparities in pSTEM (e.g., Good et al., 2012; Pajaras & Miller, 1994), there has been no concurrent assessment of the two variables, so their relative contributions are unknown. By simultaneously evaluating the effects of two variables known to be related to pSTEM persistence, we sought to provide a strong test of the degree to which each of these variables could better explain gender gaps. Thus, the current work makes a unique contribution to the growing body of research explaining precursors to gender representation and achievement in pSTEM by simultaneously pitting belonging and self-efficacy (two known precursors) against each other. This examination is critical in order to better refine interventions designed to ameliorate gender gaps.

Overview and Hypotheses

The current work extends previous research on gender disparities in pSTEM in three important ways by examining (1) multiple pSTEM domains where women are especially underrepresented (e.g., computer science, physics); (2) individuals at different stages of their career progression (early- to late-career undergraduates and graduate students); and (3) two related but distinct important outcomes (intentions and actual persistence). We predicted that (1) women would report lower levels of belonging than men and (2) the relation between a sense of belonging and academic persistence would be stronger for women than men. These effects were expected to occur even after controlling for a well-documented predictor of academic persistence: self-efficacy. First, we tested our hypotheses at a single time point in large nationally representative samples of undergraduate students (Study 1a) and graduate students (Study 1b) pursuing computing-related degrees (e.g., computer science, computer engineering). This design allowed us to capture a snapshot of how the relation between a sense of belonging (vs. self-efficacy) and intentions to persist in pSTEM might differ for women and men. Next, we expanded our design to assess students from multiple pSTEM disciplines at the start of their pSTEM major and tracked these students beyond the semester to measure actual persistence (Study 2). This longitudinal design provides better evidence for the causal role of a pSTEM sense of belonging and self-efficacy on actual persistence for women.

Studies 1a and 1b

Studies 1a and 1b focused on the issues of women's low representation in pSTEM by looking specifically at gender disparities in the field of computing. Women's representation in the field of computing is particularly poor (e.g., 18% of bachelor's degree earners in computing in 2013 were women; NSF, 2015a) and the gender gap in representation in computing professions has even grown in the last two decades (Landivar, 2013). We assessed our hypotheses with two large-scale survey data sets, each collected at a single time point: one from undergraduate computing majors (Study 1a) and a second from graduate computing students (Study 1b).

Study 1a: Method and Participants

Data were collected from 1,946 undergraduate students majoring in a computing field who participated in our study in exchange for being entered into a raffle to win a US$100 gift card. Data from 107 of these students were excluded because they reported that they were not computing majors. Data from an additional 80 undergraduate students were excluded because these students did not report their gender or our primary dependent variable of thoughts about leaving and thus could not be included in the invariance testing and structural equation modeling. Finally, because these data were collected entirely online, we paid special attention to the quality of participants' responses and found that 24 students' survey responses were suspect. These individuals gave inconsistent responses (i.e., their ratings were uniform across all scale items even when some items were keyed in the opposite direction). These considerations resulted in a final sample of 1,735 students (453 women; 1,282 men).

Students were recruited from 68 academic computing departments across the United States who were participating of a larger, semiannual data collection initiative. Items used in the present analyses were part of this regular data collection initiative, which also included additional items not of relevance to the present analyses. Measures for the larger study have been collected since 2013 and will continue indefinitely. All survey items were voluntary, so students could skip items they did not wish to answer. All students who completed the survey (regardless of the number of questions answered) were entered into the raffle, with the exception of (1) students who indicated they did not wish to be entered into the raffle and (2) students enrolled at institutions whose institutional review boards (IRBs) did not allow raffles to be used as incentives. The data collection site for Studies 1a and 1b obtains approval from each IRB for each institution involved with the study. More information on the larger data collection effort can be found at http://cra.org/cerp/data-buddies/.

Prior to this study, participating departments had been recruited via stratified random sampling to include roughly equivalent representation of (1) highly-ranked PhD programs (approximately top 20%, see U.S. News & World Report, 2014); (2) other PhD programs; (3) terminal master's and bachelor's granting departments; and (4) 4-year program institutions or liberal arts colleges. Within the current sample, 17% of student respondents were enrolled at an institution

with a top-ranked PhD program; 41% from an institution that granted PhDs but was not ranked in the top 20%; 17% from an institution whose highest awarded computing degree was a terminal master's, but who also awarded bachelor's degrees; and 22% from institutions that only offered 4-year programs/liberal arts colleges. The participating departments distributed the survey to eligible students. Because we never had access to lists of potentially eligible students, there was no way to gauge student response rate. The data collection rule for stopping was that we would keep the survey open for 2 months and only analyze data collected during that window of time. Data were collected for 2 months during the Fall 2013 academic semester in order to allow departments ample time to distribute the survey and send out survey reminders and to allow students ample time to complete the survey. The representation of women within our sample (~26% women) was roughly in line with the gender representation of students who earn their bachelor's degree in computing (~19 women collapsing across 2004–2014; NSF, 2015a). The racial and ethnic makeup of the sample was 4.6% African American, 14.4% Asian American, 54.3% Caucasian, 10.5% Latina/Latino, and 16.2% Other. The sample consisted of 16.2% first-year students, 19.1% second year students, 30.5% third year students, and 34.2% fourth year or greater students.

Procedure and Measures

Students who were majoring in a computing field within each department in our sample were invited to complete an online survey via an email invitation sent by their department chair or an administrative staff person in their department. Embedded within the survey were questions pertaining to students' sense of belonging in, and self-efficacy with, the field of computing and whether they had ever considered changing to a non-computing major (we did not ask about students' interest in a specific non-computing majors).

Sense of belonging. Three items assessing belonging were adapted from Walton and Cohen's (2007) Sense of Academic Fit measure: "I feel like I 'belong' in computing," "I feel like an outsider in the computing community," and "I feel welcomed in the computing community." Each item used a scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. An additional belonging item, "I do not have much in common with the other students in my computing classes," was excluded from analyses because its low factor loading (−.57) indicated it was a poor indicator of belonging. The Cronbach's α of the three retained items was .79.

Self-efficacy. Students' self-efficacy in computing was assessed via the following prompt: "I am confident that I can:" and items "earn a 'B' or better grade in all my computing courses," "complete my undergraduate degree in computing," and "get admitted to a graduate computing program," using a scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Two additional self-efficacy items were excluded. One item, "I am confident I can find employment in my field," was excluded because its relatively low factor loading (.48) indicated it was a poor indicator of self-efficacy. The second item, "I am confident that I can complete a graduate degree in computing," was excluded because analyses indicated this item was nearly redundant with the other self-efficacy item regarding graduate school. Because this item added no unique information, it was dropped, as is consistent with standard best practice in confirmatory factor analysis (Brown, 2006). The Cronbach's α of the three retained items was .77. The substantive results did not change if the excluded items were retained in the analyses. (This was true for all studies; see the Online Supplemental Materials at journals.sagepub.com/doi/suppl/10.1177/0361684317720186 for more details.)

Persistence intentions. The following item assessed students' thoughts about leaving their major: "Since declaring or planning to declare your computing major, have you seriously considered changing to a non-computing major?" and responded with either "yes" or "no".

Study 1a: Results and Discussion

In order to test our hypotheses, we first needed to be sure that sense of belonging was measured the same way for women and men, so analyses began with testing for measurement invariance of a confirmatory factor model (for details see Supplemental Materials available at journals.sagepub.com/doi/suppl/10.1177/0361684317720186). After establishing measurement invariance of the latent constructs, we ran preliminary analyses, then examined our focal hypotheses by testing for structural invariance in the latent variable means and paths of interest.

All analyses were conducted in Mplus Version 7 (Muthén & Muthén, 1998–2012). In line with the established guidelines (Kline, 2005) and general benchmarks (Hu & Bentler, 1999; van de Schoot, Lugtig, & Hox, 2012), model fit was evaluated using the root mean square error of approximation (RMSEA; values less than .05 are consistent with good fit and values less than .08 are satisfactory) and the comparative fit index (CFI; values greater than .95 indicate good fit and values greater than .90 are satisfactory). Model χ^2 tests are reported because they are useful for model comparisons, but given our large sample size, they are largely uninformative of model fit (Kline, 2005). The measurement models were tested using a maximum likelihood estimator; the structural models required a weighted least squares estimator because the outcome was a categorical variable.

Descriptive Statistics

Table 1 displays the zero-order intercorrelations, means, and standard deviations of the raw variables separately for women and men. We computed raw scores for belonging and

Table 1. Correlations and Descriptive Statistics for Study 1a Raw Variables Separately for Women and Men.

	Sense of Belonging	Self-Efficacy	Considered Leaving
Belonging	—	.43*	−.41*
Self-efficacy	.36*	—	−.41*
Considered leaving	−.24*	−.23*	—
Women's *M* (*SD*)	61.20 (23.02)	78.97 (21.13)	0.23 (.42)
Men's *M* (*SD*)	73.96 (19.51)	85.36 (16.33)	0.10 (.30)
	$t(1,733) = 11.39$, $p < .001$	$t(1,733) = 6.60$, $p < .001$	$t(1,733) = 6.68$, $p < .001$

Note. The correlations for women are above the diagonal of the matrix and the correlations for men are below the diagonal. Belonging and self-efficacy are represented as percentage of maximum possible transformation scores with 0 representing the minimum possible score and 100 representing the maximum possible score. The mean for "considered leaving" represents the proportion of students who reported they had considered leaving their major. The *t* tests assess the effect of gender for each variable.
*$p < .05$.

self-efficacy using a percentage of maximum possible transformation (POMP scores) for intuitive interpretation (Cohen, Cohen, Aiken, & West, 1999). POMP scoring linearly transforms the raw metric of the scores such that 0 represents the minimum possible score and 100 represents the maximum possible score. Tests of mean differences using the latent variable models (presented below) are preferable, given these analyses can account for measurement error but, for completeness, we also present tests of mean differences using the raw scores in the bottom row of Table 1. Analysis of the patterns of missing data revealed that less than 3% of all items for all cases were missing, and none of the items were not missing data for any case. Considering individual cases, 92.5% of participants had no missing data. Finally, no item had 8% or more of missing values.

Model Testing

We used structural invariance testing to determine whether belonging differentially predicted intent to persist for women and men. To accomplish this, we built upon the best fitting model identified from measurement invariance testing (see Supplemental Materials available at journals.sagepub.com/doi/suppl/10.1177/0361684317720186) and included intent to persist (1 = *Yes, I have seriously considered changing to a non-computing major*, 0 = *No, I have not seriously considered changing to a non-computing major*) as an outcome variable. This model was a good fit to the data, $\chi^2(32) = 55.20$, $p \leq .01$, RMSEA = .031, 90% CI [.018, .044], CFI = .987. Figure 1 displays the standardized path estimates from the unconstrained model.

Before examining whether there were significant gender differences in the path estimates, we examined our first hypothesis that women would experience a lower sense of belonging in computer science than men. We constrained the means of the latent sense of belonging and self-efficacy variables to 0 for men, but freely estimated them for women. The test of whether women's means differ from zero then became a test of group differences; because the latent variables are standardized, the estimate for women's means can be interpreted akin to a Cohen's *d* (i.e., it represents the difference between women and men's sense of belonging in standardized units). As we predicted, women reported a significantly lower sense of belonging ($M_{\text{belonging}} = -0.63$, $p < .001$) and lower self-efficacy ($M_{\text{self-efficacy}} = -0.35$, $p < .001$) than men. In other words, women's self-reported sense of belonging on average was over half a standard deviation lower than men's, and women's self-efficacy on average was about a third of a standard deviation lower than men's. The variances in sense of belonging, $\sigma^2 = .67$ for women and $\sigma^2 = .45$ for men; $\chi^2_\Delta(1) = 9.24$, $p = .002$, differed for women and men, as did the variance in latent

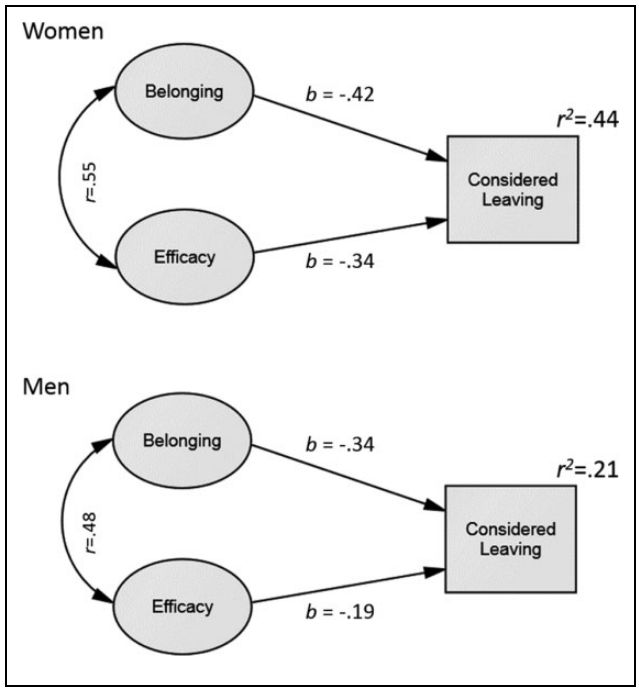

Figure 1. The Study 1a unconstrained structural model of the effect of belonging and self-efficacy on persistence intentions separately for women and men. The coefficients are standardized parameter estimates; all parameters are significant at the .05 level. For simplicity, the individual indicators of the latent variables and the measurement residuals are not depicted.

self-efficacy, $\sigma^2 = .68$ for women and $\sigma^2 = .41$ for men; $\chi^2_\Delta(1) = 15.51, p < .01$, as indicated by a significant worsening in model fit when these variances were constrained to be equal. Based on these results, the latent variable means and the variances were allowed to vary for women and men in subsequent models.

Next, we estimated our hypothesized structural model and tested whether there were gender differences in the path estimates for the effect of sense of belonging and self-efficacy on intent to persist. We hypothesized that sense of belonging would be a stronger predictor of intent to persist for women, compared to men. In order to test this prediction, we compared model fit in a constrained model, where this parameter was forced to be equal for women and men, to the fit for a less restrictive model, where the parameter was freely estimated for women and men. Improvement in model fit when constraints were lifted indicates that the parameter in question differs for women and men. When the path from belonging to intent to persist was constrained to be equal for women and men, model fit did not significantly worsen, $\chi^2_\Delta(1) = 1.00, p = .32$; belonging was only a slightly stronger predictor of intentions to persist for women than men and this difference was not significant (see coefficients for paths from sense of belonging to considered leaving in Figure 1). Note that standardized path estimates can be interpreted as standardized effect sizes (Kelley & Preacher, 2012); thus, for women, every one standardized unit decrease in sense of belonging was associated with a .42 standardized unit increase in self-reported thoughts about leaving, whereas this effect was somewhat smaller for men; for men, a one standardized unit decrease in belonging was associated with a .34 standardized unit increase in self-reported thoughts about leaving.

Self-efficacy was a significant predictor of intent to persist for women and men, so, for completeness, we also tested whether gender moderated the path from self-efficacy to intent to persist. When the path from self-efficacy to intent to persist was forced to be equal for women and men, model fit did not worsen, $\chi^2_\Delta(1) = 1.90, p = .17$, which indicates no gender difference in the predictive utility of self-efficacy on intent to persist (see coefficients for paths from Self-Efficacy to Considered Leaving in Figure 1).

In sum, our results show that, among undergraduate computing majors, women report a lower sense of belonging (and less self-efficacy) than men; and sense of belonging is linked to intentions to persist for women and men, even when controlling for self-efficacy (see Figure 1). However, while sense of belonging was a directionally stronger predictor of thoughts about leaving for women than for men, this difference was not statistically significant. Nevertheless, the lower mean levels of sense of belonging among women than men suggest a higher likelihood of considering leaving computing among women than men, implicating the importance of sense of belonging in women's persistence in pSTEM. To further test our primary hypothesis and demonstrate that sense of belonging matters all along the pSTEM path and not just at the outset, in Study 1b, we tested the same hypotheses in a more advanced sample of students.

Study 1b: Method and Participants

Data were collected from 1,010 graduate students, enrolled in a computing graduate program, who participated in our study in exchange for being entered in a raffle to win a US$100 gift card. Data collection ended 2 months after the online survey was opened. Data from 131 graduate students were excluded because these students were enrolled in a predominately online graduate program. We opted to exclude these individuals because we were concerned that sense of belonging would be difficult to quantify for students who do not have face-to-face interactions with others in their program. Data from 10 graduate students were also excluded because these students did not report their gender and thus could not be included in the invariance testing. Finally, data cleaning indicated that 14 students' survey responses were suspect. These individuals gave inconsistent responses (i.e., their ratings were uniform across all scale items even when some items were keyed in the opposite direction). These considerations resulted in a final sample of 855 students (315 women, 540 men). As with Study 1a, we were unable to calculate a response rate because we never had access to lists of potentially eligible students.

Sampling occurred from the same departments using the same recruiting strategy as Study 1a: Participants in Study 1b were entered into a raffle to win a US$100 gift card. Also, as with Study 1a, students took part in Study 1b as part of a larger study. Data collected for this study were collected alongside data collected for the larger study; no additional measures were created for the current study. Measures for the larger study have been collected since 2013 and will continue indefinitely. All survey items were voluntary, so students could skip items they did not wish to answer. All students who completed the survey (regardless of the number of questions answered) were entered into the raffle, with the exception of students who indicated they did not wish to be entered into the raffle and students enrolled at institutions whose IRBs did not allow raffles to be used as incentives.

Within the current sample, 26% of student respondents were enrolled at an institution with a top-ranked PhD program, 60% from an institution that granted PhDs but was not top-ranked, 7% from an institution whose highest awarded computing degree was a terminal master's, and 7% had unknown institution classification. In terms of type of graduate program, 1.8% participants were enrolled in a joint bachelor's and master's program, 50.2% in a terminal master's program, and 48% in a PhD program. The representation of women within our sample was inflated ($\sim$37% women) compared to the gender representation of students who earn their graduate degrees in computing (MS: $\sim$28% women; PhD: $\sim$21% women; NSF, 2013). The racial and

Table 2. Correlations and Descriptive Statistics for Study 1b Raw Variables Separately for Women and Men.

	Belonging	Self-Efficacy	Considered Leaving
Belonging	—	.37*	−.37*
Self-efficacy	.40*	—	−.27*
Considered leaving	−.08	−.22*	—
Women's *M* (*SD*)	66.77 (22.38)	78.73 (19.65)	0.33 (.47)
Men's *M* (*SD*)	75.45 (20.06)	82.93 (17.19)	0.25 (.43)
	$t(852) = 5.84$, $p < .001$	$t(853) = 3.27$, $p = .001$	$t(853) = 2.53$, $p = .01$

Note. The correlations for women are above the diagonal of the matrix and the correlations for men are below the diagonal. Belonging and self-efficacy are represented as percentage of maximum possible transformation scores with 0 representing the minimum possible score and 100 representing the maximum possible score. The mean for "considered leaving" represents the proportion of students who reported they had considered leaving their major. The *t* tests assess the effect of gender for each variable.
*$p < .05$.

ethnic makeup of the sample was 2.9% African American, 53.5% Asian American, 32.7% Caucasian, 2.4% Latina/Latino, and 8.5% Other.

Procedure and Measures

Students who were enrolled in a graduate program in a computing field within each department in our sample were invited to complete an online survey via an email invitation sent by their department chair or an administrative staff person in their department. Embedded within the survey were questions pertaining to students' sense of belonging, self-efficacy, and whether they had ever considered leaving their graduate program. The survey was administered during the Fall 2013 academic semester.

The same 3 items from Study 1a were used in Study 1b to measure sense of belonging. The Cronbach's α of the 3 items for the graduate sample was .74.

Students' self-efficacy in computing was assessed via the following prompt and 3 items: "I am confident that I can": "find a job in my field," "become an expert in my field," and "complete my department's milestones towards earning my degree," using a scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. An additional self-efficacy item, "I am confident that I can publish in the top journals in my field," was excluded because its low factor loading (.57) indicated it was a poor indicator of self-efficacy. The Cronbach's α of the three retained items was .74.

The following item assessed students' thoughts about leaving their graduate program: "During your academic career, have you ever seriously considered leaving your graduate program?" using the following three response options: "I have seriously considered leaving, and I still think of it sometimes"; "I have seriously considered leaving, but I don't think of it anymore"; or "I have never seriously considered leaving." For consistency with Study 1a and because we were most interested in comparing students who had ever considered leaving their program to those who had not, the first two responses were coded as "yes" and the last response option was coded as "no." Results do not differ if this variable is instead treated as continuous instead of dichotomized.

Study 1b: Results and Discussion

Descriptive Statistics

Table 2 displays the zero-order intercorrelations, means, and standard deviations of the raw variables (POMP scored) separately for women and men. Analysis of the patterns of missing data revealed no instances where all items for all cases were missing, and 10% of the items were not missing data for any case. Considering individual cases, 98.5% of participants had no missing data. Finally, no item had more than 2% missing values. Again, we first tested whether women and men differed in their sense of belonging and self-efficacy, then we assessed whether sense of belonging and self-efficacy differentially predicted intent to persist for women and men. To accomplish this, we built upon the best fitting model identified from the measurement invariance testing (see Supplemental Materials available at journals.sagepub.com/doi/suppl/10.1177/0361684317720186) and included intent to persist (yes vs. no coding described above) as an outcome variable. This model was adequate to good fit to the data, $\chi^2(29) = 86.08$, $p < .01$, RMSEA = .068, 90% CI [.052, .085], CFI = .93. Figure 2 displays the standardized path estimates from the unconstrained model.

Before examining whether there were gender differences in the path estimates, we examined whether women and men differed in mean levels of sense of belonging. This was accomplished by constraining the means of the two latent variables to 0 for men but freely estimating them for women. The test of whether women's means differ from zero then becomes a test of group differences. As we predicted, women reported a significantly lower sense of belonging ($M_{\text{belonging}} = -0.53$, $p < .001$) and lower self-efficacy ($M_{\text{self-efficacy}} = -0.35$, $p = .004$) than men. Neither the variances in sense of belonging, $\sigma^2 = .54$ for women and $\sigma^2 = .47$ for men; $\chi^2_{\Delta}(1) = 1.30$, $p = .25$, nor latent self-efficacy, $\sigma^2 = .38$ for women and $\sigma^2 = .30$

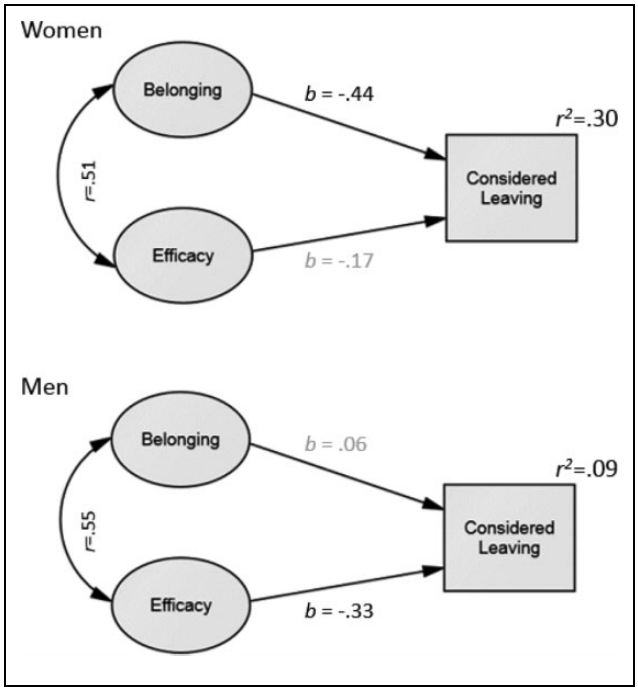

Figure 2. The Study 1b unconstrained structural model of the effect of belonging and self-efficacy on persistence intentions separately for women and men. The coefficients are standardized parameter estimates, and black-font parameters are significant at the .05 level. For simplicity, the individual indicators of the latent variables and the measurement residuals are not depicted.

for men; $\chi^2_\Delta(1) = 2.85$, $p = .09$, differed for women and men, as indicated by a nonsignificant worsening in model fit when these variances were constrained to be equal. Based on these results, the latent variable means were allowed to vary, but the variances were constrained to be equal for women and men in subsequent models.

Next, we tested whether sense of belonging differentially predicted intent to persist for women and men. We hypothesized that sense of belonging would be a stronger predictor of intent to persist for women compared to men. Consistent with this, the structural model results showed that sense of belonging significantly predicted intent to persist for women but not for men (see coefficients for paths from sense of belonging to considered leaving in Figure 2). Specifically, whereas for women, every one standardized unit decrease in sense of belonging was associated with a .44 standardized unit increase in thoughts about leaving, this effect was significantly smaller and not significantly different from zero for men. For men, a one standardized unit decrease in sense of belonging was associated with a .06 standardized unit decrease in thoughts about leaving. To further test our hypothesis, we also compared model fit in a constrained model where this parameter was forced to be equal for women and men to the fit for a less restrictive model. As we expected, when the path from belonging to intent to persist was constrained, model fit was significantly worse, $\chi^2_\Delta(1) = 9.43$, $p < .01$; sense of belonging was a stronger predictor of intentions to persist for women than men.

Self-efficacy was a significant predictor for men ($b = -.33$, $p < .001$) and marginally so for women ($b = -.17$, $p = .07$), so for completeness, we also tested whether gender moderated the path from self-efficacy to intent to persist. When the path from self-efficacy to intent to persist was forced to be equal for women and men, model fit did not worsen, $\chi^2_\Delta(1) = 1.68$, $p = .20$, which indicates no gender difference in the predictive utility of self-efficacy on intent to persist.

Together, these results show that among graduate computing students, women reported a lower sense of belonging than men. Moreover, sense of belonging was more strongly linked to intentions to persist for women than it is for men, even when controlling for self-efficacy. In fact, this model explained nearly a third of the variance in intent to persist for women but less than a tenth for men.

Study 2

Using large nationally representative samples of computing majors, Studies 1a and 1b provide initial evidence for our theorized explanation for why women are more likely than men to withdraw from pSTEM fields: Women experienced a relatively lower sense of belonging in computing. Moreover, within the graduate sample, there was a closer relation between belonging and thoughts about leaving the field, even after accounting for self-efficacy. The relation was directionally similar in the undergraduate sample, but in that sample, the strength of the association did not differ significantly for women and men. Although Study 1a did not find a stronger relation between sense of belonging and thoughts about leaving among women, compared with men, sense of belonging was still a significant predictor of thoughts about leaving among women. Coupled with the significant relation between sense of belonging and thoughts about leaving among only women in Study 1b and the significantly lower mean level of belonging among women than men in both samples, the results highlight the importance of sense of belonging for women's persistence in pSTEM.

Studies 1a and 1b each assessed data at a single time point, which limits our ability to make a causal argument for the impact of sense of belonging on women's tendency to leave pSTEM. The main goal of Study 2 was to provide a stronger test of our hypothesis that belonging is linked with pSTEM persistence for women, by assessing how sense of belonging and efficacy in an initial gateway pSTEM course was related to actual persistence in the form of completing the second course in a sequence of courses required for one's pSTEM major. To enhance generalization, we also tested our hypotheses in a different pSTEM domain, that of physics.

Study 2 also addressed an alternative explanation to our findings from Studies 1a and 1b: Women's low persistence in pSTEM may be driven by their performing poorly in pSTEM, rather than a low sense of belonging. Simply put, doing

poorly in an earlier class can prevent a student from moving forward in their degree program. We thus expected performance to affect persistence for both men and women. Of note, women often underperform in high-stakes pSTEM testing situations—particularly when they are confronted with signals that they are less apt in pSTEM than men (Danaher & Crandall, 2008; Inzlicht & Ben-Zeev, 2000; Spencer et al., 1999). However, we hypothesized that some students (and female students in particular) would have a low sense of belonging, regardless of their objective course performance, and that sense of belonging is what threatens their persistence. Thus, accounting for gender differences in performance allows for a particularly strong test of the effects of sense of belonging and self-efficacy on women and men's persistence in pSTEM.

Study 2: Method and Participants

To expand our assessment into domains of pSTEM other than computing, we recruited students from an introductory level, calculus-based physics course required for pSTEM majors at the university where Study 2 was conducted. Compensation for participation was a small amount of course credit. We attempted to survey every student; thus, sample size and stopping rule were determined by course enrollment. A total of 736 students were enrolled in the course. We excluded students from analyses who were not pSTEM majors ($n = 297$). This inclusion criterion was set in order to make the test of our hypotheses more conservative, since we reasoned that non-pSTEM majors would report less belonging and be less likely than pSTEM majors to enroll in the second course in the sequence, as this would not be a requirement for non-majors. We also excluded students who did not complete the course ($n = 23$), since this meant they were not eligible to enroll in the next course in the sequence, which was the outcome of interest. These considerations resulted in a final sample of 416 students (122 women, 294 men). The distribution of specific pSTEM majors in our sample was as follows: 10.3% physical sciences, 1.4% computer science, 83.7% engineering, and 4.6% mathematics. The racial and ethnic makeup of the sample was 0.5% African American, 0.5% American Indian, 5.4% Asian American, 72.5% Caucasian, 11.2% Latina/Latino, and 9.8% other. The sample was primarily first-year students (60%) and second-year students (34%), with a few third-year students (4%) and students in their fourth year or beyond (2%).

Procedure and Measures

Students were invited to complete an online survey during the penultimate week of the 16-week semester during the spring of 2012. Embedded within this survey were questions very similar to the items used in Study 1 to measure belonging but modified to relate to the physics course. We obtained the following from official university records: (1) students' exam scores and (2) whether students completed the second physics class in the sequence. Official records were available for all the students who met our inclusion criteria; thus there were no missing data for our main dependent variable, but not all students completed the individual difference measures (the response rate of final sample was 81.97%). However, because our analyses (invariance testing in a structural equation modeling framework) used a maximum likelihood estimator, complete data are not required and this method of estimation can implicitly cope with missingness and provide less biased parameter estimates (Enders, 2010).

Sense of belonging and self-efficacy. Students' sense of belonging in physics was assessed via the following 3 items: "I feel like I belong in physics," "People in physics accept me," and "I feel like an outsider in physics," each using a scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. The Cronbach's α of the 3 items was .70.

Students' self-efficacy was assessed via the following prompt: "Rate how confident you are that you can do each of the following in your next science course:" followed by the items "learn the course concepts," "perform well on exams," and "complete homework assignments by myself," each using a scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. An additional self-efficacy item, "I am confident that I can complete the course with a B or better," was excluded because its low factor loading (.52) suggested that it was a relatively poor indicator of self-efficacy. The Cronbach's α of the three retained items was .83.

Exam performance. Students completed four multiple choice exams during the semester as part of the course curriculum. Percent correct on each exam was averaged together and the composite was then standardized. The Cronbach's α of the four exam scores was .83. We opted to treat exam performance as an observed rather than a latent variable because, although individual exam scores are potentially unreliable measures of a students' latent ability and class knowledge, exam scores in their imperfectly measured form can, and do, have real consequences for college students' persistence; this relation was of interest to us.

Completion of second semester physics course. We obtained official student records in Spring 2014 in order to determine whether each student completed the second course in the sequence at any time in the 2 years since their completion of the first physics course in the sequence, operationalized as received a final grade (0 = *did not complete*, 1 = *completed*). We selected completing the second course as our measure of persistence because both courses are required for pSTEM majors. Completion of that second course would be necessary to be eligible to take many upper-level courses in the major (which is why we chose to follow-up at 2 years out as opposed to a longer time frame). Thus, failing to complete the sequence in a timely manner would reflect a major

Table 3. Correlations and Descriptive Statistics for Study 2 Raw Variables Separately for Women and Men.

	Belonging	Self-Efficacy	Exam Scores	Persistence
Belonging	—	.53*	.48*	.37*
Self-efficacy	.54*	—	.45*	.24*
Exam scores	.30*	.46*	—	.31*
Persistence	.14*	.12	.36*	—
Women's *M* (*SD*)	52.54 (26.61)	50.33 (26.09)	63.69 (14.07)	0.84 (.37)
Men's *M* (*SD*)	63.61 (21.11)	65.22 (23.55)	71.47 (14.35)	0.81 (.40)
	$t(339) = 4.11$, $p < .001$	$t(337) = 5.48$, $p < .001$	$t(414) = 5.06$, $p < .001$	$t(414) = 0.72$, $p = .48$

Note. The correlations for women are above the diagonal of the matrix and the correlations for men are below the diagonal. Degrees of freedom differ for the *t*-tests because of missing data for the individual difference measures. Belonging and self-efficacy are represented as percentage of maximum possible transformation scores with 0 representing the minimum possible score and 100 representing the maximum possible score. Exam scores are reported as mean % correct. The mean for "persistence" represents the proportion of students who went on to complete the next course. The *t* tests assess the effect of gender for each variable.
*$p < .05$.

impediment to degree completion and makes this an important assessment of persistence.

Major change. We also examined whether students changed to a non-pSTEM major as another potential dependent measure. There was a significant zero-order point-biserial correlation between belonging and changing major for women ($r = -.21$, $p = .04$), but not men ($r = -.10$, $p = .16$); however, this correlation became nonsignificant after controlling for exam performance and self-efficacy ($r = -.05$, $p = .64$ and $r = -.03$, $p = .67$ for women and men, respectively). Very low variability in this outcome (only 8.3% of women and 6.3% of men switched to a non-pSTEM major) may have made it difficult to detect our anticipated effects of our independent measures. This possibility was not considered further.

Study 2: Results and Discussion

Table 3 displays the zero-order intercorrelations, means, and standard deviations of the raw variables (POMP scored) separately for women and men. Analysis of the patterns of missing data revealed no instances where all items for all cases were missing, and 25% of the items were not missing data for any case. Considering individual cases, 79.3% of participants had no missing data. Finally, no item had more than 20% missing values.

We tested whether women and men differed in their sense of belonging and also whether belonging differentially predicted intent to persist for women and men. To accomplish this, we built upon the best fitting model identified from the measurement invariance testing (see Supplemental Materials available at journals.sagepub.com/doi/suppl/10.1177/0361684317720186) and included actual persistence (0 = *did not take next course in the sequence*, 1 = *took next course in the sequence*) as an outcome variable. Exam performance was added as a covariate of sense of belonging and self-efficacy and was also modeled as an exogenous predictor of persistence. The unconstrained structural model was a good fit to the data, $\chi^2(39) = 54.83$, $p = .05$, RMSEA = .044, 90% CI [.005, .070], CFI = .964; Figure 3 displays the standardized path estimates from this model.

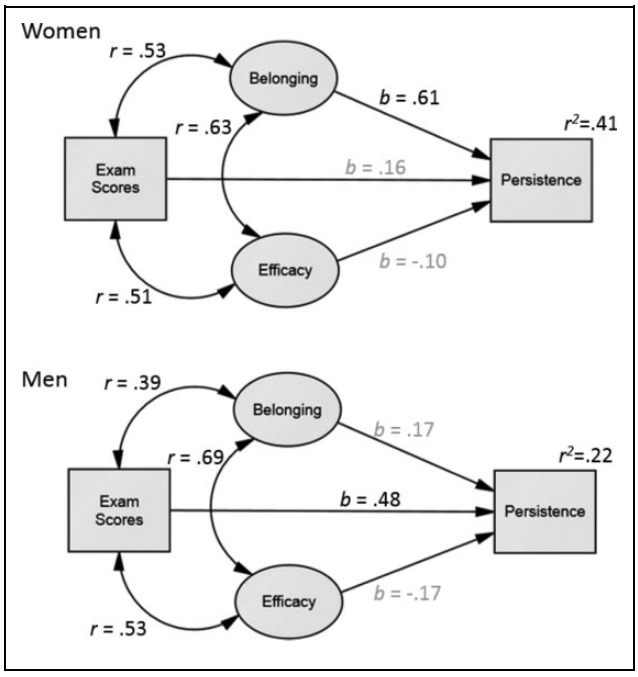

Figure 3. The Study 2 unconstrained structural model of the effect of belonging, self-efficacy, and exam performance on persistence separately for women and men. The coefficients are standardized parameter estimates, and black-font parameters are significant at the .05 level. For simplicity, the individual indicators of the latent variables and the measurement residuals are not depicted.

Before examining whether there were gender differences in the path estimates, we examined whether women and men differed in mean levels of sense of belonging. We constrained the means of the two latent variables to 0 for men but freely estimated them for women. The test of whether women's means differ from zero then became a test of group differences. As we predicted, women reported a significantly lower sense of belonging ($M_{\text{belonging}} = -0.36$, $p = .02$) than men.

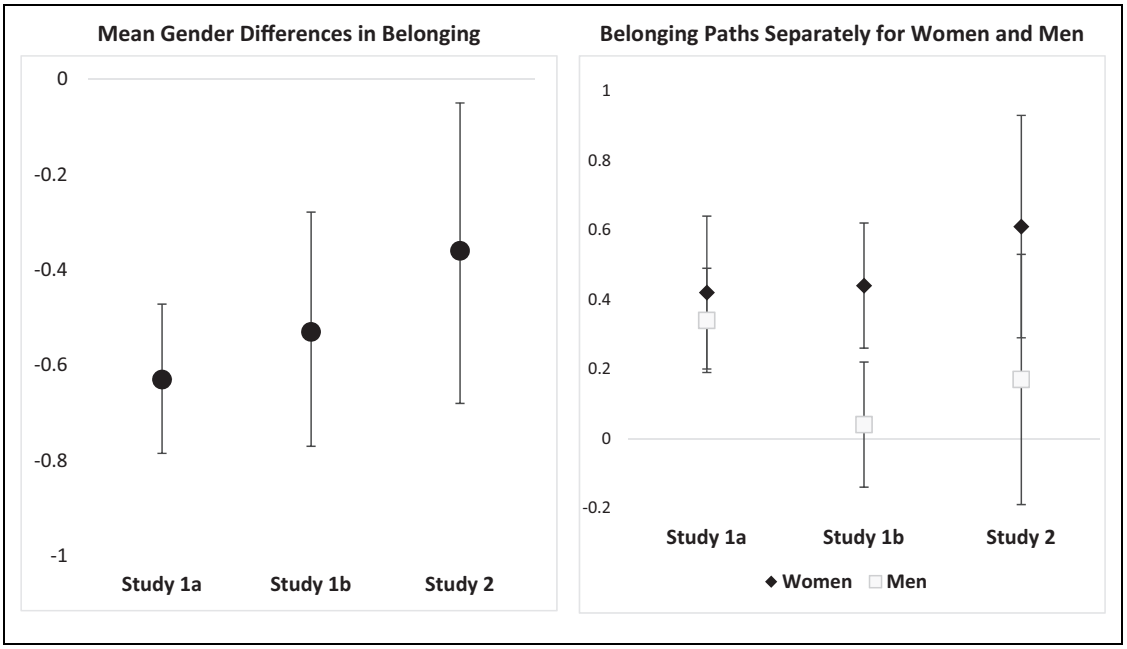

Figure 4. The left panel displays the mean standardized gender differences in belonging (with 95% confidence intervals) for each of the three studies. More negative values indicate relatively lower belonging for women than for men. The right panel displays the absolute value of the standardized path estimates for the effect of belonging on persistence (with 95% confidence intervals) separately for women and men for each of the three studies.

Women in this sample also reported lower self-efficacy ($M_{\text{self-efficacy}} = -0.67$, $p < .001$) than men. The variances in the latent sense of belonging variable differed for women and men, $\sigma^2 = 1.27$ for women and $\sigma^2 = .68$ for men; $\chi^2_\Delta(1) = 4.86$, $p = .02$, but the variances in the latent self-efficacy variable did not, $\sigma^2 = 1.17$ for women and $\sigma^2 = .99$ for men; $\chi^2_\Delta(1) = 0.34$, $p = .56$. Based on these results, the latent variable means were allowed to vary in subsequent models. The variance for self-efficacy was constrained, but the variance for belonging was freely estimated for women and men.

Next, we tested whether there were gender differences in the path estimates for the effect of belonging on intent to persist. We hypothesized that sense of belonging would be a stronger predictor of intent to persist for women, compared to men. Consistent with this prediction, the structural model results show that sense of belonging significantly predicted persistence for women but not for men (see coefficients for paths from sense of belonging to persistence in Figure 3). Specifically, whereas for women, every one standardized unit decrease in sense of belonging was associated with a .61 standardized unit decrease in the likelihood of persistence, this effect was significantly smaller and not significantly different from zero for men. For men, a one standardized unit decrease in belonging was associated with a .17 standardized unit decrease in likelihood of persistence. Furthermore, model fit worsened, $\chi^2_\Delta(1) = 3.05$, $p = .08$, when the path from sense of belonging to persistence was constrained to be equal for women and men, which further indicates that sense of belonging was a stronger predictor of persistence for women than men. The only significant predictor of persistence for men was exam performance. Model fit also worsened, $\chi^2_\Delta(1) = 4.52$, $p = .03$, when the path from exam scores to persistence was constrained, which indicates that exam performance was a stronger predictor of persistence for men than women (see coefficients for paths from exam scores to persistence in Figure 3).

Study 2 shows that among undergraduate pSTEM majors in a gateway physics course, women reported a lower sense of belonging than men, and sense of belonging was more strongly linked to persistence for women than it was for men, even when controlling for self-efficacy and exam performance. It is noteworthy that sense of belonging did not predict persistence for men, but exam performance did. By contrast, exam performance did not predict persistence for women. Unlike Study 1a and 1b, self-efficacy in this sample did not predict persistence for either men or women, which is likely because Study 2 controlled for actual achievement.

General Discussion

Past approaches to explaining gender disparities in naturalistic pSTEM settings often have focused on ability-relevant factors, such as women's relatively lower self-efficacy in pSTEM, relative to men (e.g., Kost-Smith et al., 2009; Vogt, Hocevar, & Hagedorn, 2007). Although these explanations are clearly valuable in understanding precursors to pursuing pSTEM, our research demonstrates that this focus is incomplete, at least when considering women who choose to

initially pursue pSTEM in post-secondary education. Across three field studies (for comparison purposes, see Figure 4) with data from close to 3,000 students in differing pSTEM domains and educational levels, using cross-sectional and longitudinal designs, we found that women question their ability in pSTEM more than men (as evidenced by lower self-efficacy) and women feel a lower sense of belonging than men in pSTEM (see left panel of Figure 4). Moreover, in each study, we found that women's sense of belonging in pSTEM reliably served as a predictor of outcomes critical to retaining women en route to pSTEM careers—namely, reported intentions to persist in pSTEM and actual persistence on the pSTEM path. We found that women's relatively lower sense of belonging was a more consistent predictor of persistence than either women's lower self-efficacy or academic achievement, again highlighting the critical role of social fit in pSTEM pursuits for women. Our results converge with the comments of accomplished astrophysicist Meg Urry that opened this article: Even women with demonstrable accomplishments (e.g., Urry 2005; women who have secured admission to graduate school in Study 1b) are susceptible to the negative effects of a low sense of belonging. Our results showing the importance of sense of belonging in accounting for academic outcomes coincides with the theoretical "core" nature of belonging for well-being; feeling appreciated and welcomed in the social milieu is a universal and unwavering human need (see Baumeister & Leary, 1995). Thus, our findings that having a sense of belonging outshines other conventional and robust explanations for women's intentions to persist and for their actual persistence in pSTEM further validated just how "core" belonging is.

Of note, in two of the three studies, belonging was more strongly linked with women's than with men's persistence in pSTEM, which we expect is due to a greater sense of belonging uncertainty among women in this domain where their status and ability are more in question (Murphy & Zirkel, 2015; Walton & Cohen, 2007). That is, women (but not men) are continually faced with cultural messages that their gender does not belong in pSTEM, creating uncertainty as to whether they as an individual belong in pSTEM. This makes the link between perceptions of belonging in the local environment and persistence particularly strong for women. For men, whose sense of belonging in pSTEM is consistently strong (as seen by gender differences in belonging favoring men throughout all of our studies), belonging appears to be less relevant to their decisions to persist in pSTEM. This is not to say that belonging is not important for men, but merely that on average in this domain, the current socio-cultural context makes belonging status more influential on persistence for women than men.

Given the critical need for belonging, it seems feasible that the wide range of cultural, social, and environmental predictors of women's low representation and achievement within pSTEM operate in part through their effects on women's sense of belonging. Although the current work neither measured nor manipulated predictors of sense of belonging, a growing body of literature articulates a number of common characteristics of pSTEM settings that hinder women's sense of fit within them, including the virtual absence of other women (Murphy, Steele, & Gross, 2007; Stout et al., 2011), role-models and peers who seem to embody stereotypes (Cheryan et al., 2012), women's perceptions that others endorse negative gender stereotypes about women's relative inaptitude in pSTEM (Good et al., 2012), and explicit forms of bias against women such as outright discrimination (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012). Our work extends prior research to showcase why, and the degree to which, women's relatively lower sense of belonging compares to men's matters. Even after controlling for other correlates of achievement outcomes, low sense of belonging for women predicted weaker self-reported intentions to persist in pSTEM and a lower tendency to actually persist on the academic track to obtaining a pSTEM degree in all three samples.

Belonging only predicted intent to persist for men in Study 1a (see right panel of Figure 4). The best predictor of outcomes for men was performance (Study 2). Thus, when it comes to navigating the path to a pSTEM career, our results suggest that, whereas men may be more likely to calibrate their decisions with objective indicators of future potential (i.e., performance), women may attend more to their subjective sense of belonging in the field. Why was belonging less likely to be linked with persistence for men than women? We propose this is because low belonging is less threatening for men who, unlike women, continually receive indications that they do belong (e.g., membership in the dominant majority group for whom positive stereotypes abound). In other words, when men question "Do I belong here?," the environment responds with a resounding "Yes!"; when women question their belonging, they encounter silence at best or complete disavowal at worst. We did not anticipate that the best predictor of persistence for men would be objective performance, but we speculate that students will persist depending on the most salient indicator for their future success and enjoyment in a field. Whereas a male student may have an easier time brushing off concerns about belonging as temporary and situation specific, a female student may be more likely to perceive those concerns as global and enduring for her experiences in the entire domain (though this is speculative and warrants further study). The most salient indicator for men in this study seems to be how they were objectively performing in the course.

Limitations and Future Research

One limitation of this work is the inherent disadvantage of most survey research: We cannot know how the results for non-responders may have differed from those of the individuals who did participate in our study. This is a particular concern in drawing conclusions from the results of Studies

1a and 1b where we only know the numbers of students who chose to participate but do not have access to the total of number of students surveyed; thus, we have no information about non-responders and cannot calculate a response rate. However, this concern is somewhat mitigated given the consistent replication of the predicted effects across all three studies (see Figure 4). Furthermore, our confidence is bolstered by the results of Study 2 where the response rate was high ($\sim 82\%$) and our analysis technique was able to implicitly cope with missingness due to non-response on the survey measures, and we calculated less biased parameter estimates that were consistent with our hypothesized effects and the results of Studies 1a and 1b.

Another limitation of this work is that we only compared men and women within pSTEM domains, so it is not clear whether the gender differences we observed in these data generalize to other domains. However, as outlined in the Introduction, we believe that women's experiences in pSTEM domains differ from those of women in fields that are less male-dominated because contextual factors specific to pSTEM fields set up a unique set of circumstances that depress women's sense of belonging and also amplify the effects of low belonging. Consistent with this, in two samples of non-pSTEM majors, Lewis and Hodges (2015) showed that women and men do not differ in their level of belonging with their major and that belonging predicted intentions to persist in the major equally for women and men.

Related to this point, Leslie, Cimpian, Meyers, and Freeland (2015) draw attention to the fact that female underrepresentation is also problematic in some fields outside of pSTEM (e.g., philosophy, music composition). They reported that the percentage of women in a field is highly correlated with common beliefs among that field's practitioners about what is required for success in that domain; namely, male-dominated fields tend to place more emphasis on brilliance and innate talent, whereas female-dominated fields place greater emphasis on effort and determination. Given the negative stereotypes that exist regarding women's intellect (Spencer et al., 1999), ability beliefs about a field may be another cue that leads women to conclude that they do not belong in male-dominated fields. Consistent with this, Lewis and Hodges (2015) theorize that belonging uncertainty may take the form of worrying about one's social fit and/or one's intellectual fit within a domain, and each component of belonging can uniquely account for academic outcomes. The measures used in the current study focused more on the social fit dimension of belonging. Future research should consider whether the perceived ability fit of pSTEM students might show gender discrepancies similar to those documented here.

In our analyses, we collapsed across year in major and program; study results did not change when we controlled for this variable. However, further investigation of grade level is an important consideration for future research. Our measure of intent to persist in Study 1a, in particular, may have failed to capture important variability in persistence intentions for students at differing levels on the path to degree completion; we only asked if students had ever considered leaving their major or degree program, but not when or how often they had considered leaving. Future research that further considers when along the path to degree completion students are most vulnerable to low belonging will be important in order to better target interventions and additional supports.

A final limitation is that only partial measurement invariance was established in Studies 1b and 2 (i.e., model fit significantly worsened when the factor loading for 1 item, the intercepts for 2 items in Study 1b and the intercept for 1 item in Study 2 were constrained to be equal, thus these parameters were allowed to vary for women and men). Although partial measurement invariance is typically considered satisfactory (Steenkamp & Baumgartner, 1998), because full measurement invariance was not established in these data, it is not possible to fully know to what degree the observed differences between women and men are caused by differences in the measurement of the constructs of interest. Also, we excluded some of the original items that were used to measure belonging and self-efficacy, based on the results of the confirmatory factor analysis. Although we did not find that our main results differed substantially when the poor-functioning items were retained, it is always possible that researchers using all measured items may find different results.

We are presently unable to determine why persistence was more strongly related to sense of belonging for women than for men in two of our samples, but not the third. We note that the influence of sense of belonging on persistence among women was consistent across all samples, highlighting the importance of focusing more attention on women's sense of belonging in pSTEM. What remains uncertain are the conditions under which sense of belonging also affects persistence among men. The two samples in which we found a stronger association of sense of belonging with persistence varied in education level (graduate students and undergraduates) and content area (computing exclusively vs. a broad representation of pSTEM fields). By contrast, the sample in which sense of belonging equally predicted thoughts about leaving for women and men involved undergraduates in computing. This suggests that neither education level nor pSTEM field accounts for the gender difference in our results. Further research might consider the conditions under which sense of belonging influences persistence for men, as well as women.

Practice Implications

Gender differences in pSTEM attrition are an enduring problem with real social costs; for example, women miss out on the economic benefits of high paying pSTEM careers and the pSTEM labor force suffers from a dearth of qualified job candidates (Augustine, 2005). We identified one reason that

women participate in pSTEM at lower rates than men, which has implications for direct intervention efforts: Our work suggests that a reliable route to increased representation of women in pSTEM is to narrow the gap between women and men's perceptions of belonging and create inclusive environments that affirm women's belonging just as much as men's. What can educators do to work toward this goal and foster women's sense of belonging in pSTEM contexts? Fortunately, research indicates that women's sense of belonging in pSTEM can be modulated by environmental modifications (for a review, see Lewis, Stout, Pollock, Finkelstein, & Ito, 2016). For instance, women's sense of belonging is boosted when exposed to environments that are clearly gender-neutral (rather than stereotypically masculine; Cheryan et al., 2009), that showcase female peers and role models (Dennehy & Dasgupta, 2017; Murphy et al., 2007; Stout et al., 2011), or that emphasize effort and hard work over brilliance and innate talent (Smith et al., 2013).

Conclusions

Bolstering women's sense of belonging may go a long way toward promoting increased gender diversity in pSTEM. In turn, greater gender diversity will bring a wide range of viewpoints to bear on scientific problems, facilitating innovation, creative problem-solving, and novel discoveries. To that end, we hope scientists, educators, and policy makers will commit to fostering inclusive pSTEM environments, so that all students have the opportunity to engage with pSTEM.

Authors' Note

The data reported in this article are available from the authors upon request, pending institutional review board approval for distribution.

Acknowledgments

We thank Jessica Cundiff and Ama Nyame-Mensah for data collection efforts, Blake Redabaugh for providing student academic information, and Lauren Kost-Smith for contributions to prior research leading to Study 2's development.

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 work was funded by National Science Foundation grants DRL0910373 and DRL0833364 to University of Colorado Boulder and CNS-1246649 and DUE-1431112 to the Computing Research Association.

References

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhi & J. Beckmann (Eds.), *Action—Control: From cognition to behavior* (pp. 11–39). Heidelberg, Germany: Springer.

Ajzen, I. (1987). Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 1–63). New York, NY: Academic Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Ajzen, I. (2011). Behavioral interventions: Design and evaluation guided by the theory of planned behavior. In M. M. Mark, S. I. Donaldson, & B. C. Campbell (Eds.), *Social psychology for program and policy evaluation* (pp. 74–100). New York, NY: Guilford.

Augustine, N. R. (2005). *Rising above the gathering storm: Energizing and employing America for a brighter economic future. Committee on Prospering in the Global Economy of the 21st Century*. Washington, DC: National Academies Press.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529. doi:10.1037/0033-2909.117.3.497

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Cheryan, S., Drury, B. J., & Vichayapai, M. (2012). Enduring influence of stereotypical computer science role models on women's academic aspirations. *Psychology of Women Quarterly*, *37*, 72–79. doi:10.1177/0361684312459328

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

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*, 315–346. doi:10.1207/S15327906MBR3403_2

Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, *38*, 1639–1655. doi:10.1111/j.1559-1816.2008.00362.x

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

Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences*. Advance online publication. doi:10.1073/pnas.1613117114

Diekman, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology*, *101*, 902–918. doi:10.1037/a0025199

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in

mathematics. *Journal of Personality and Social Psychology*, *102*, 700–717. doi:10.1037/a0026659

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools*, *30*, 79–90. doi:10.1002/1520-6807(199301)30:1<79:: AID-PITS2310300113>3.0.CO;2-X

Hausmann, L. R., Schofield, J. W., & Woods, R. L. (2007). Sense of belonging as a predictor of intentions to persist among African American and White first-year college students. *Research in Higher Education*, *48*, 803–839. doi:10.1007/s11162-007-9052-9

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Hughes, J. N., Im, M. H., & Allee, P. J. (2015). Effect of school belonging trajectories in grades 6–8 on achievement: Gender and ethnic differences. *Journal of School Psychology*, *53*, 493–507. doi:10.1016/j.jsp.2015.08.001

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

Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, *17*, 137–152. doi:10.1037/a0028086

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics—Physics Education Research*, *5*, 1–14. doi:10.1103/PhysRevSTPER.5.010101

Landivar, L. C. (2013). *Disparities in STEM employment by sex, race, and Hispanic origin* (American Community Survey Reports ACS-24). Washington, DC: U.S. Census Bureau.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology*, *33*, 265–269. doi:10.1037/0022-0167.33.3.265

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*, 262–265. doi:10.1126/science.1261375

Lewis, K. L., & Hodges, S. D. (2015). Expanding the concept of belonging in academic domains: Development and validation of the Ability Uncertainty Scale. *Learning and Individual Differences*, *37*, 197–202. doi:10.1016/j.lindif.2014.12.002

Lewis, K. L., Stout, J. G., Pollock, S. J., Finkelstein, N. D., & Ito, T. A. (2016). Fitting in or opting out: A review of key social-psychological factors influencing a sense of belonging for women in physics. *Physical Review Physics Education Research*, *12*, 1–10. doi:10.1103/PhysRevPhysEducRes.12.020110

London, B., Rosenthal, L., Levy, S. R., & Lobel, M. (2011). The influences of perceived identity compatibility and social support on women in non-traditional fields during the college transition. *Basic and Applied Social Psychology*, *33*, 304–321. doi:10.1080/01973533.2011.614166

Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics*, *74*, 118–122. doi:10.1119/1.2162549

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *107*, 631–644. doi:10.1037/edu0000005

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., & Zirkel, S. (2015). Race and belonging in school: How anticipated and experienced belonging affect choice, persistence, and performance. *Teachers College Record*, *117*, 1–40. Retrieved from http://www.tcrecord.org/Content.asp?ContentId=18154

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

National Science Board. (2012). *Science and engineering indicators 2012*. Arlington, VA: National Science Foundation.

National Science Foundation, National Center for Science and Engineering Statistics. (2015a). *Integrated Postsecondary Education Data System, Completions Survey. Table 5-2. Bachelor's Degrees Awarded, by Field and Sex: 2004–14* (Special Tabulations of U.S. Department of Education, National Center for Education Statistics). Retrieved from https://www.nsf.gov/statistics/2015/nsf15311/tables.cfm

National Science Foundation, National Center for Science and Engineering Statistics. (2015b). *Scientists and Engineers Statistical Data System (SESTAT), Table 9-5. Employed Scientists and Engineers, by Occupation, Highest Degree Level, and Sex: 2013*. Retrieved from https://www.nsf.gov/statistics/2015/nsf15311/tables.cfm

Orenstein, P. (1994). *Schoolgirls: Young women, self-esteem, and the confidence gap*. New York, NY: Anchor Books.

Pajares, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology*, *86*, 93–203. doi:10.1037/0022-0663.86.2.193

Pittman, L. D., & Richmond, A. (2007). Academic and psychological functioning in late adolescence: The importance of school belonging. *The Journal of Experimental Education*, *75*, 270–290. doi:10.3200/JEXE.75.4.270-292

Pollock, S. J., Finkelstein, N. D., & Kost-Smith, L. E. (2007). Reducing the gender gap in the physics classroom. *Physics Review: Special Topics in Physics Education Research*, *3*, 1–4. doi:10.1103/PhysRevSTPER.3.010107

Schott, G., & Selwyn, N. (2000). Examining the "male, antisocial" stereotype of high computer users. *Journal of Educational*

Computing Research, *23*, 291–304. doi:10.2190/V98R-5ETX-W9LY-WD3 J

Settles, I. H., Jellison, W. A., & Pratt-Hyatt, J. S. (2009). Identification with multiple social groups: The moderating role of identity change over time among women scientists. *Journal of Research in Personality*, *43*, 856–867. doi:10.1016/j.jrp.2009.04.005

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

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

Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*, 78–90. doi:10.1086/209528

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*, 255–270. doi:10.1037/a0021385

U.S. News & World Report. (2014). *Best grad schools: Computer science*. Retrieved from http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-science-schools/computer-science-rankings?int=983508

Urry, M. (2005, February 6). Diminished by discrimination we scarcely see. *The Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/articles/A360-2005Feb5.html

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*, 486–492. doi:10.1080/17405629.2012.686740

Vogt, C. M., Hocevar, D., & Hagedorn, L. S. (2007). A social cognitive construct validation: Determining women's and men's success in engineering programs. *The Journal of Higher Education*, *78*, 337–364. doi:10.1353/jhe.2007.0019

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

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

Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal of Personality and Social Psychology*, *102*, 513–532. doi:10.1037/a0025731

Zumbrunn, S., McKim, C., Buhs, E., & Hawley, L. R. (2014). Support, belonging, motivation, and engagement in the college classroom: A mixed method study. *Instructional Science*, *42*, 661–684. doi:10.1007/s11251-014-9310-0