Received: 16 April 2020 | Revised: 1 November 2021 | Accepted: 6 December 2021

DOI: 10.1002/tea.21741

RESEARCH ARTICLE

JRST | WILEY

Sources of gender differences in competency beliefs and retention in an introductory premedical science course

Eben B. Witherspoon | **Christian D. Schunn**

Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Correspondence
Eben B. Witherspoon, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara St., Pittsburgh, PA 15217.
Email: eben.witherspoon@pitt.edu

Funding information
Division of Undergraduate Education, Grant/Award Number: 1524575

Abstract

Gender disparities in retention in pathways to science continue to vary widely by course. Undergraduates intending to study prehealth and premedicine often represent a majority of students enrolled in introductory science courses, contribute to a large number of eventual science degree earners, and are a population that typically includes a high number of women. However, gender differences in attrition, grades, and attitudes persist in the introductory science courses required by undergraduate preheath and premedical programs, particularly within the physical sciences (i.e., Chemistry and Physics). We use structural equation modeling to study 416 undergraduate students across multiple sections of an Algebra-based Physics course, a common course on the prehealth and premedical track where large gender differences in grades, retention and competency beliefs have been documented. Our analysis focuses on identifying potential academic and attitudinal sources for gender differences in students' beliefs about their Physics abilities at the end of the course, and retention to the second physics course, which is often influenced by these competency beliefs. Results suggest that while men's ability beliefs in Physics are relatively stable and largely derived from early performance indicators, this is a smaller source of ability beliefs for women. Instead, women's ability beliefs are mediated during the course through their sense of belonging in Physics, and the extent to which they believe that Physics ability is fixed or malleable. These findings can inform the design of interventions in Physics courses that specifically target the development of ability beliefs for women intending medical careers.

KEYWORDS

competency beliefs, gender, premedical education, retention, science education

1 | INTRODUCTION

As a result of decades of intervention, research, and activism, great strides have been made toward increasing equity in undergraduate STEM participation for women (see Brotman & Moore, 2008 for a review). Recent reports show that overall, women in college earn higher grades and graduate at higher rates than men (Conger & Long, 2010), earn a greater number of life science degrees (National Science Board, 2020; Xie & Shauman, 2003), and among students who enter intending to pursue science, persist to earn science degrees at a similar rate (King, 2016; Lord et al., 2009). However, other studies suggest that persistence in specific science majors continues to vary widely by gender (Matz et al., 2017). For example, while in general the number of women earning undergraduate degrees in the life sciences has surpassed men in recent years (e.g., Biology and Neuroscience, 72% women), the physical sciences continue to be male-dominated (e.g., Chemistry and Physics, 38% women; National Science Foundation, 2011). This lack of gender diversity has been linked to reduced innovation (Beede et al., 2011). The low participation by women in physical sciences also perpetuates gender wage disparities: graduates with a Bachelor's degree in the physical sciences have a higher median wage ($65,000) than those in the life sciences ($54,000), resulting in a $9000 wage penalty in the science degrees more commonly earned by women (Carnevale et al., 2015).

Furthermore, many studies of gender differences in STEM participation continue to treat the sciences as a single disciplinary pathway, an approach that potentially obscures continued gender disparities in enrollment between different science disciplines and retention along alternative trajectories to science (Cannady et al., 2014). While some studies locate the source of gender differences in science participation as stemming from the large gap between men and women at the point of *entering* science (Legewie & DiPrete, 2014; Ohland et al., 2008), others suggest that these differences in participation appear as a function of differences between men and women *leaving* STEM majors (Chen & Soldner, 2013). Prior work often also conceptualizes retention across science courses as primarily an issue for science majors, even though many introductory science courses are dominated by students intending to pursue prehealth or premedical studies, who may migrate into nonmedical science majors. Premedical education in the United States takes place at the undergraduate level, and typically requires a rigorous sequence of science courses across a variety of disciplines, including multiple courses in Biology, Chemistry and Physics. At our own institution, approximately two-thirds of all undergraduates in the College of Arts & Sciences enter the university with the intention to pursue medicine, and these students dominate many of the introductory science courses. The current study focuses on attitudes and retention within a physical science course sequence along the premedical pathway that has previously been shown to have pervasive gender disparities in retention (Witherspoon & Schunn, 2019).

Studies using nationally representative data have also shown that at the end of high school, women are more than twice as likely as men to express interest in pursuing careers in medicine or health (Sadler et al., 2012). Furthermore, students who enter these courses with the intent to pursue medicine are likely to have larger variation in attitudes toward particular science courses than students choosing those courses with a specific interest and intent to major in those particular disciplines (Gasiewski et al., 2012). For example, while not mutually exclusive, some premedical students may be primarily drawn to the human services aspects of a career in medicine, while others may be more interested in the salary and prestige of such professions, and others still may enter because of scientific interests. Women in particular have been shown to be more likely to choose prehealth professions with their main expressed interest being in working with people, and therefore may see introductory science courses simply as necessary stepping stones to a health profession they are pursing primarily for reasons other than their interest in science (Miller et al., 2006).

Importantly for the number of degrees earned in the sciences, while a high proportion (40–50%; Gasiewski et al., 2012; Witherspoon & Schunn, 2019) of students in introductory science courses may initially enter university intending prehealth or premedicine, relatively few students overall actually persist into those professions (less than 3%; Association of American Medical Colleges, 2020). Students intending premedicine and prehealth but who do not go on to take the medical school entrance exam have been shown to represent a larger proportion of eventual science degree earners than do students initially intending to major in science; however, women initially intending those professions are often more likely than men to switch into non-STEM fields entirely (Witherspoon & Schunn, 2019). A nationally representative study showed that of all students who initially intended to pursue medicine at the doctoral level, by 2 years later 53.6% of women had switched out of both doctoral-track medicine and STEM majors, while only 36.1% of men had switched (Morgan et al., 2013). Therefore, negative experiences in introductory science courses required for prehealth and premedical studies may be particularly important contributors to differential attrition of women from those courses, and act as pathways away from, rather than toward, other science-related majors and careers (Barr, 2010).

1.1 | Sources of gender differences and the primacy of competency beliefs

There is overwhelming research from psychology to suggest that observed gender differences in science participation are socially constructed, rather than derived from biological or cognitive sex differences. Small differences in science ability by sex that have been found are often in opposite directions such that the means for science ability are just as often higher for women. Furthermore, all differences in ability by sex fall far below a threshold that would explain the large gender differences in participation that are observed (Hyde, 2005).

Therefore, current work focuses more on the underlying and concurrent social and psychological factors. For example, evidence suggests that women are often socialized at a young age to participate in different science-related activities than men, leading to disparities in prior experience that can contribute to differences in both performance and beliefs about ability at later stages (Jones et al., 2000; Vincent-Ruz & Schunn, 2017). Furthermore, particularly in disciplines where women are underrepresented and where negative stereotypes about women's abilities exist, even high-performing women are more likely than men to attribute difficulties in the subject to inherent skill, rather than the difficulty of the content or aspects of the learning environment (Beyer & Bowden, 1997; LaCosse et al., 2016). Particularly in higher education research, separately understanding the effects of differences in prior experience and preparation as well as the more immediate effects of differences in men and women's experiences of undergraduate science learning environments can help inform targeted interventions in these processes at the undergraduate level.

Specifically, students' competency beliefs (i.e., beliefs about their abilities within certain domains) have been shown across a number of studies to be a significant predictor of gender differences in both grades and persistence in STEM fields more broadly, and the physical sciences in particular (Cromley et al., 2016; Debacker & Nelson, 2000; Huang, 2013; Sawtelle, Brewe, & Kramer, 2012). Well-researched motivational frameworks such as expectancy-value theory (EVT) and social cognitive career theory (SCCT) suggest that beliefs about ability may be more predictive of students achievement outcomes (e.g., grades), while more distal choices (e.g., course, major and career) are instead influenced indirectly thought the impact of ability beliefs on other aspects of students motivation like their value of or interest in a particular domain (Lent et al., 2017, 2018; Multon et al., 1991; Perez et al., 2019).

While a number of studies indicate that differences in competency beliefs may be an important factor in gender participation gaps that appear in introductory science courses (see Huang, 2013), less is known about the particular sources of competency beliefs for premedical students in these courses, and whether these are differential by gender. One study of premedical students shows that self-efficacy beliefs in Chemistry partially mediate gender differences in decisions to persist to the second Organic Chemistry course, even for men and women who earn the same high grade (i.e., an A or B) in the first Organic Chemistry course (Witherspoon et al., 2019). However, this effect should be replicated in other required physical science courses along the premedical pathway. Furthermore, understanding how men and women in premedical physical science courses come to develop different beliefs about their abilities can be important to informing interventions that help to mitigate gender differences in competency beliefs, performance, and persistence in these courses. Thus, the current study will seek to replicate the central role of competency beliefs in grades and retention in an algebra-based physics course, which enrolls a high proportion of premedical students.

1.2 | Sources of competency beliefs

There is a large body of extant research on possible sources of competency beliefs (see Usher & Pajares, 2008 for a review). Bandura (1986) identified four primary sources of competency beliefs: *mastery experiences* (i.e., interpretations of prior academic performance), *vicarious experiences* (i.e., students' perceptions and comparison of the activities of relevant others in the class), *verbal and social persuasions* (i.e., the feedback that students receive from others regarding their abilities), and *physiological and affective states* (i.e., the physical and mental feelings experienced in a course, and interpretations of those experiences). Each of these could be connected to women's experiences in physical science courses (Zeldin & Pajares, 2000).

Much of the research on sources of competency beliefs have primarily been concerned with identifying which of these psychological factors are most strongly correlated with competency beliefs overall (Usher & Pajares, 2008). However, by gender there remains uncertainty in the literature if men and women experience differences in sources of competency beliefs. Some studies claim to find no gender differences in sources of competency beliefs (Matsui et al., 1990; Stevens et al., 2007) while others propose that competency beliefs are necessarily interpreted in context, thereby suggesting that an interaction between the individual and a particular domain determines the sources of competency beliefs that students attend to (Lent et al., 1996; Pajares et al., 2007; Trujillo & Tanner, 2014). For example, studies in science suggest that women are more likely to generate competency beliefs from vicarious experiences and verbal and social persuasions, while men are more likely to draw on mastery experiences (Sawtelle, Brewe, & Kramer, 2012). This suggests that while interpretation of their own prior experiences and performance (i.e., their own *mastery experiences*) may be the primary source of competency beliefs for men, the behaviors of others (i.e., interpretation of others' *vicarious experiences*, explicit *social and verbal persuasions*) greatly influence how women experience science classrooms. While there are relatively few studies that have tested sources of competency beliefs within specific science domains, this earlier work can provide a foothold for identifying constructs that could contribute to the differential development of competency beliefs in science courses by gender.

1.2.1 | Sense of belonging

Recent studies have documented how the lack of a "sense of belonging" (i.e., feelings of membership and acceptance in a group) can lead to lower rates of persistence for women in male-dominated physical science domains like Engineering and Physics (Lewis et al., 2017; Walton et al., 2015). Research examining sense of belonging in a general premedical context suggest it may mediate interest in continued premedical study by gender (Rosenthal et al., 2013), while studies in a Calculus-based Physics course have shown belonging to be equally associated with higher grades for both men and women (Stout et al., 2013). Understanding the effects of sense of belonging on grades and persistence may therefore be particularly important in introductory physical sciences courses for nonmajors, as these courses are often required for premedicine, typically enroll a higher proportion of women, and yet exhibit greater attrition for high-performing women (Witherspoon et al., 2019).

Early research has also established a temporal ordering in the link between students' feelings of belonging and subsequent academic competency beliefs, through primarily these studies have been conducted with middle school students (McMahon & Wernsman, 2008; Roeser et al., 1996). While studies in higher education have also shown a high correlation between belonging and competency beliefs, the directionality of these effects is less clear; some drawing from social cognitive career theory instead model belonging and competency beliefs both as predictors of subsequent interest (Tellhed et al., 2017).

Particularly in an introductory physical science course for nonmajors, sense of belonging may be related to two key sources of competency beliefs for women: *vicarious experiences*, and *social and verbal persuasions*. Uncertainty about belonging for underrepresented groups is believed to function in part through inducing stereotype threat (i.e., anxiety about confirming negative stereotypes about one's group), which may unconsciously interfere with cognition and inhibit performance (Steele & Aronson, 1995; Walton & Cohen, 2007). This mechanism also can contribute to gender differences in the interpretation of relative effort in domains like the physical sciences, where women may be more likely to experience belonging uncertainty. If stereotype threat leads to difficulties in performance, women may interpret this mismatch of effort and performance as an indicator of lower abilities (Smith et al., 2013). Furthermore, students' beliefs about their abilities may be particularly likely to develop through both explicit feedback and comparison to others during new and transitional experiences in science such as the first course in college, as internalized metrics of successful performance have not yet formed and are thus relatively unknown during these periods (Bandura, 1997; Eccles et al., 1984). Thus, the current study will examine how the processes described above predict changes in competency beliefs for both men and women within introductory physics.

1.2.2 | Implicit theories of intelligence

Another separate but related element of the learning environment that may contribute to gender differences in the development of ability beliefs are implicit theories of intelligence that students hold about a particular domain (e.g., theories of intelligence in physics). Adopting theories about the nature of ability in a field as "fixed" (i.e., ability is innate and unchangeable), rather than "malleable" (i.e., ability is changeable and can to be developed through effort), is detrimental to students' future achievement and persistence (Blackwell et al., 2007). This mechanism is thought to operate through influencing students' perception of effort; students who believe that ability is linked to effort see challenges as part of growth, while students who believe that ability is fixed see effort as an indicator of lack of ability (Dweck & Leggett, 1988). Therefore, the development of a more fixed belief about the nature of intelligence in a field could provide a lens through which *vicarious experiences* and *social persuasions* are interpreted; students with a fixed mindset may be more likely to interpret their effort relative to others as a lack of inherent ability, or negative feedback about their ability as something that cannot be changed.

Empirical support for this hypothesized link remains mixed, with some studies finding no correlation between implicit theories of intelligence and ability beliefs (Cury et al., 2006; Hong et al., 1999). However, other recent studies of students in middle and high school science courses support this hypothesis; latent profile analyses show that students with more malleable beliefs about ability are more likely to accept competency beliefs feedback from multiple sources than those with more fixed mindsets, and as a result showed larger gains in competency beliefs (Chen & Usher, 2013). Importantly, to the extent that these beliefs are implicitly or explicitly perpetuated by faculty or other students in the class, fixed mindsets may be particularly salient and detrimental to women in historically male-dominated fields like the physical sciences. Studies have shown that university faculty perceptions of domains that require "innate brilliance" are negatively correlated with the number of women in those fields, with many of the physical sciences being the most extreme along those dimensions (Leslie et al., 2015).

Furthermore, students' implicit theories of intelligence may contribute separately to their ability beliefs, or in combination with other factors like sense of belonging, in ways that produce gender differences in competency beliefs. For example, research in mathematics suggests that sense of belonging for women may be particularly low if they believe negative stereotypes about women's mathematics ability, and if they believe that mathematics ability is a natural inherent ability that is difficult to change (Good et al., 2012). Similarly, women in the physical sciences could also internalize beliefs about their abilities through a combination of messaging courses: the extent to which the particular domain is explicitly presented as requiring an innate set of skills, as well as the extent to which they implicitly feel accepted as full members of that domain. However, we are not aware of any studies that that have empirically tested the link between theories of intelligence, belonging and competency beliefs in the context of an undergraduate Algebra-based Physics course (see Eddy & Brownell, 2016 for a review). Thus, the current study will also examine the role theories of intelligence in physics play in predicting changes in competency beliefs for both men and women.

1.3 | Current study overview

The current study addresses gaps in the current literature by utilizing a combination of historical institutional data and intensive course-level survey data to examine gender differences in the development of competency beliefs, grades, and retention to the next course, an introductory Algebra-based Physics course sequence for non-Physics majors. This course was selected as it is, along with Organic Chemistry, one of the required introductory physical science courses for premedical students that often show large differences in attrition for underrepresented groups (Barr et al., 2010; Witherspoon et al., 2019). Physics in particular is one of the core science domains that has continued to have difficulty attracting and retaining women (Cheryan et al., 2016; Matz et al., 2017). Women in Physics tend to underperform relative to their male peers on Physics assessments, as well as have lower beliefs about their ability to perform in Physics courses (Marshman et al., 2017). This could in part be related to differences in the perception of the importance of mathematics ability in the physical sciences, and differences in perceived mathematics ability by gender. For example, performance on the SAT Mathematics section may be more strongly related to lower beliefs about Chemistry ability for women than men (Vincent-Ruz et al., 2018). Relatedly, in mathematics-related fields like Physics which are often disproportionately male, men are more likely to overestimate their performance, which can contribute to women perceiving their abilities in these courses to be artificially low relative to these male peers (Bench et al., 2015). These attributions have been related to differences in retention decisions, even when women earn the same grades as their male peers (Sanabria & Penner, 2017). However, much of the research on gender differences in Physics participation focus on how these differences appear among Physical science or Engineering majors, and therefore focus on the sequence of calculus-based courses.

Understanding gender differences in participation in algebra-based Physics courses, typically taken by premedicine or prehealth majors, can help to corroborate prior mechanisms or uncover new ones, which explain how discipline-specific academic and attitudinal gender differences manifest along this large pathway to science-related careers. Furthermore, unlike Calculus-based Physics, Algebra-based Physics are typically more equally populated by gender, reducing the effects of numeric underrepresentation in the classroom for women, and focusing instead on beliefs about the discipline. Therefore, in this way the current study builds on prior work conducted in Calculus-based Physics classrooms that examines the impact of belonging on gendered grades and participation in Physics (Lewis et al., 2016; Stout et al., 2013), and expands on it in an Algebra-based Physics course by incorporating measures of students' beliefs about the fixed or malleable nature of Physics ability, and contrasting the effects of these broader attitudes about physics with concurrent measures of content knowledge and indicators of prior experience.

Importantly, we also include the longitudinal development of self-efficacy as a key outcome that has been shown to contribute to gender differences in participation in physical sciences courses along the premedical track. Such analyses allowed us to test not just cross-sectional associations in mean levels of competency beliefs, but also the relationship between changes in these variables over time, and the mediation of the stability of ability beliefs by gender.

Therefore, the current study addresses the following research questions in an introductory, Algebra-based Physics sequence for nonmajors:

1. What are the relative contributions of prior knowledge and prior competency beliefs to grades and retention by gender?
2. What are the sources of the competency beliefs that are developed over the duration of the Physics course, and do these sources differ by gender?
3. To what extent are differences in the development of competency beliefs mediated by broader beliefs about belonging and inherent Physics ability?

2 | METHODS

2.1 | Sample

Analyses were conducted with a sample of ($N = 416$) undergraduate students ($M_{\text{age}} = 21$, $SD_{\text{age}} = 1.3$) enrolled in an Algebra-Based Introductory Physics course sequence at a large research-intensive public institution in the northeastern United States (henceforth, "the University"), a course which historically comprises over 50% premed students, and 60% women, and yet where women have slightly lower retention (66% women retained vs. 70% men). All course sections were taught using a similar mix of whole-class lecture and small group work. Data were gathered through both in-course surveys, as well as from institutional warehouse data provided by the University, both with Institutional Review Board approval.

The University is broadly representative of similar public research institutions with a relatively selective admission rate (approximately 60%): it offers over 100 undergraduate majors, a high percentage (60%) of in-state students, and a small percentage (5%) of international students. While there is large variability in family income ($SD = \$122, 000$), students tend to come from upper middle-income brackets ($Mdn = \$111{,}000$). The sample was predominantly female (58%), which is typical in the Algebra-Based Introductory Physics courses at this institution, a course which enrolls primarily premedical students.[1] The racial and ethnic diversity of the sample roughly mirrored that of the University as a whole; students were predominantly White (67%), with Asian (19%), with smaller proportion of Black (5%), Hispanic (4%), and multiracial (6%) students.

2.2 | Measures

2.2.1 | Attitudes toward physics

All students enrolled in this course were given surveys assessing domain-specific attitudes toward Physics, including competency beliefs, their sense of belonging, and theories of intelligence (see Supporting Information, Table A1 in Appendix S1 for sample items). All of the instruments were previously developed and validated (Marshman et al., 2017) with introductory physics students at the University through an iterative qualitative (cognitive interviews) and quantitative process (e.g., factor analyses). Item response theory analyses validated the use of means with these Likert-based scales and assured there was no differential scale functioning by gender.

In the current sample, no mean was near the max or min of each scale, which would have limited usefulness in analysis (see Table 1). *N*s varied due to variable attendance during recitation for pre- and post-measurements. In terms of discriminability, the highest correlation among these attitudinal variables was between competency beliefs and belonging, showing a moderate correlation of $r = 0.55$ (see Table 1). Importantly, the correlations between attitudes and ability in physics were relatively small (i.e., are separable aspects of performance with likely separate drivers). To ensure our scales maintained the same factor structure as had been observed in prior studies, we conducted a confirmatory factor analysis (CFA) of our 17 items. Results suggested that a three-factor model remained an acceptable fit to the current study data, using conventional cutoffs as thresholds for good model fit (i.e., CFI/TLI > 0.95, RMSEA <0.06, SRMR <0.08; L. T. Hu & Bentler, 1999; see Supporting Information, Figure A1 in Appendix S1.)

2.2.2 | Competency beliefs

The primary attitudinal outcome variable for this study was students' competency beliefs in Physics. It was measured both at the beginning of the semester, and at the end of the first semester, prior to the final exam. This scale contained discipline-specific items measuring students' beliefs about their ability to perform well on physics assessments and to understand physics concepts (e.g., "If I wanted to, I could do well on a physics test") rated on a four-point Likert scale (e.g., 1 = NO! to 4 = YES!). This scale ($\alpha = 0.79$) was computed as the mean across five items, with reliability similar to what has been observed in prior studies in this same context (see Marshman et al., 2017).

2.2.3 | Sense of belonging

Belonging items measured the extent to which students felt as though they were a member of the particular Physics class (e.g., "Sometimes I worry that I do not belong in this physics class") rated along a 5-point Likert scale (e.g., 1= "Not at all true" to 5 = "Completely true"). This scale ($\alpha = 0.83$) was computed as a mean across five items, and it was measured at the end of the semester.

2.2.4 | Implicit theories of intelligence

Theory of intelligence items measured the extent to which students agreed with statements that described growth mindset (e.g., "Anyone can become good at solving Physics problems through hard work") or fixed mindset (e.g., "To really excel in Physics, a person needs to have a natural ability in Physics") in relation to Physics, rated along a 4-point Likert scale (e.g. 1 = "Strongly Disagree" to 4 = "Strongly Agree"). This scale ($\alpha = 0.79$) was computed as a mean across seven items, with fixed growth items reverse coded, to generate an overall scale where higher numbers reflect higher endorsement of growth mindset, and lower numbers represent a more fixed mindset. It was measured at the end of the semester.

TABLE 1 Scale descriptives and Pearson intercorrelations among key academic (top) and motivational (bottom) variables

				Intercorrelations								
Variables	***N***	***M***	***SD***	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1. Physics 2 Retention	416	78%	—	1								
2. Physics 1 GPA	416	3.0	1.0	0.42***	1							
3. Conceptual Understanding Pre	392	37%	17%	0.07	0.32***	1						
4. Conceptual Understanding Post	386	56%	19%	0.20***	0.50***	0.67***	1					
5. SAT Math	323	669	68	0.15**	0.44***	0.45***	0.51***	1				
6. Competency Beliefs Pre	386	2.6	0.6	0.25***	0.42***	0.36***	0.35***	0.29***	1			
7. Competency Beliefs Post	416	2.8	0.6	0.15**	0.23***	0.33***	0.25***	0.23***	0.46***	1		
8. Belonging	414	3.5	0.9	0.24***	0.38***	0.28***	0.33***	0.27***	0.55***	0.40***	1	
9. Growth Mindset	416	2.8	0.5	0.22***	0.14**	0.20***	0.21***	0.21***	0.49***	0.37***	0.47***	1

**$p < 0.01$, ***$p < 0.001$.

2.2.5 | Academic performance and persistence variables

Students' academic performance was operationalized as the grade in their Physics course, measured along a 4-point GPA scale. A measure of students' prior academic performance was also collected from the University data warehouse, focusing on the measure that is most relevant for predicting performance in introductory Physics courses: students highest recorded score on the Math section of the Scholastic Aptitude Test (SAT; or the converted score from the math subscore of the ACT).

Separately, as another indicator of prior mastery of course content, student's Conceptual Understanding was measured using the Force Concept Inventory, a commonly-used, research-validated measure of students' deep conceptual understanding of first introductory Physics course content (Savinainen & Scott, 2002). By contrast, course grades involve a mixture of performance and task completion (e.g., homework points). Conceptual Understanding scores were gathered at the beginning and end of the semester during recitation as a percentage of correct items.

Finally, to examine students' persistence along the Physics course sequence, retention to the second algebra-based Introductory Physics course was measured with a single binary variable from the institutional data warehouse showing whether the student enrolled in Physics 2 at any point after the semester they took Physics 1, with 1 = continued to Physics 2 and 0 = did not continue to Physics 2 (see Table 2). While it is important to note that the significant differences in retention by gender suggested by the historical data for this course were not observed here, the wide 95% CIs (representing a possible range of gender difference from −6% to 10%) reflect that this binary measure was relatively noisy in this relatively small survey sample. However, as the current path analyses are more concerned with intermediary steps involving students' motivations than a direct gender difference in retention (which has been shown by other studies, and would require a larger sample), this does not alter the ability to draw conclusions regarding the gender moderation research questions posed in the current study.

TABLE 2 Scale descriptives by gender, with mean differences (as Cohen's *d*[a]) and 95% CI shown

	Female			Male				95% CI	
Variables	***N***	***M***	***SD***	***N***	***M***	***SD***	**Diff. (*d*)**	**Lower**	**Upper**
1. Physics 2 Retention	252	77%	—	164	79%	—	0.06	−0.14	0.25
2. Physics 1 GPA	252	2.9	1.0	164	3.2	0.9	0.30**	0.10	0.50
3. Conceptual Understanding Pre	243	32%	14%	149	45%	19%	0.83***	0.61	1.04
4. Conceptual Understanding Post	234	50%	16%	152	65%	20%	0.85***	0.63	1.06
5. SAT Math	192	653	61	132	692	70	0.60***	0.37	0.82
6. Competency Beliefs Pre	239	2.5	0.6	147	3.0	0.5	0.72***	0.51	0.93
7. Competency Beliefs Post	252	2.7	0.6	164	2.9	0.5	0.65***	0.44	0.85
8. Belonging	251	3.3	0.9	163	3.7	0.8	0.50***	0.29	0.70
9. Growth Mindset	252	2.6	0.5	164	2.9	0.6	0.55***	0.35	0.75

[a]Cohen's *d* can be interpreted as small (0.20) medium (0.50) and large (0.80) differences in means between male and female students (Cohen, 1992).

**$p < 0.01$, ***$p < 0.001$.

2.3 | Analyses

There was a small amount of missingness (all <8%) in the attitudinal survey data; a brief correlational analysis showed that a missingness indicator variable was not significantly associated with any of the predictors of interest, and so a maximum likelihood estimator was used to account for missingness in the following analyses. Examining intraclass correlations of all variables showed a large proportion of the variance for one variable (Physics 1 GPA; $ICC = 0.52$) was at the recitation level; therefore, robust standard errors were calculated to allow correlated errors to be nested within recitation. We report our findings below as standardized regression coefficients (β), using the conventional thresholds of small (0.10) medium (0.30) and large (0.50) effect sizes for regression coefficients as reported by Cohen (1992).

The structural equation modeling (SEM) package in Stata 16 was used for all analyses. SEM enables a simultaneous regression analysis of the different attitudes and academic performance outcomes, while accounting for their intercorrelations. This produces a more robust analysis of gender effects within the correlated pathways of attitudes and prior academic factors as drivers of Physics grades and retention. SEM also uses a maximum likelihood estimator, which produces less biased standard errors in the presence of missing data on the indicator variables; however, the small amount of missingness in this dataset is not likely to contribute to a significant source of bias in these analyses. Model fit were assessed using a variety of fit statistics, including CFI, TLI, RMSEA and SRMR, using conventional cutoffs as thresholds for good model fit (i.e., CFI/TLI > 0.95, RMSEA <0.06, SRMR <0.08; L. T. Hu & Bentler, 1999).

The initial path model was specified based on temporal precedence of the measured variables, with paths from prior academic performance to content knowledge, performance in the course, and to initial self-efficacy, as has been shown in prior studies (Vincent-Ruz & Schunn, 2017). Consistent with social-cognitive theory and prior work on sources of self-efficacy, students' mastery experiences (i.e., conceptual exams) and social experiences of the course (i.e., theories of intelligence, sense of belonging) were hypothesized to be related to self-efficacy at the end of the course (Usher & Pajares, 2008), and from these experiences to final course grades and persistence in the discipline

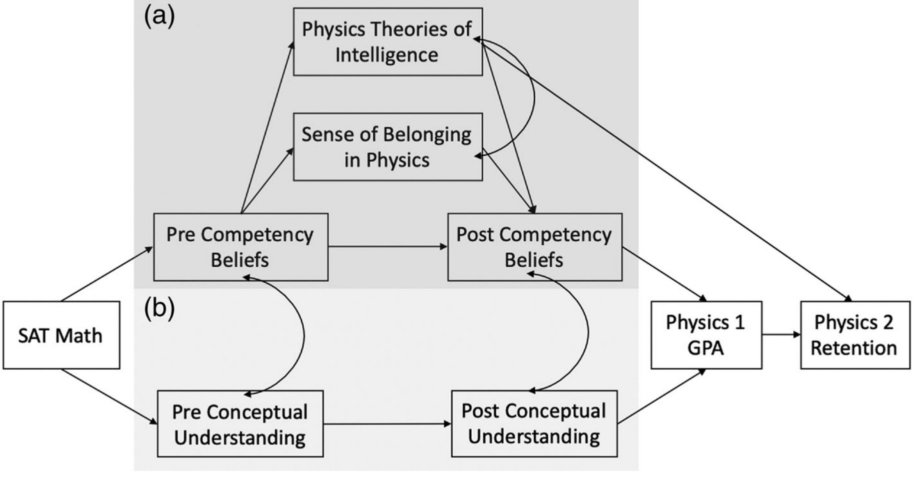

FIGURE 1 The proposed structural path model, testing the contribution of SAT math to physics performance and retention through (a) attitudinal factors (i.e., competency beliefs, belonging, theories of intelligence), and (b) conceptual understanding (i.e., FCI), by gender

(i.e., course retention; Lent et al., 2017). The final model evaluated in this study therefore simultaneously tested two hypothesized sources of students' competency beliefs in Physics: (A) a model focused on social influences and student beliefs about physics, in which changes in students' competency beliefs were mediated by their sense of belonging and growth mindset in Physics, and (B) a performance feedback model in which changes in competency beliefs were explained through prior academic performance and current Physics knowledge, as measured by the SAT math section and the conceptual understanding tests (see Figure 1). For all pathways, we tested a moderated-mediation hypothesis, to understand whether one or both of the pathways were differential strength by students' gender (e.g., did sense of belonging or growth mindset matter more for one gender?).

Model building progressed by first running a fully unconstrained model, both for all students and then moderated by gender as a multigroup SEM. Next, all paths that were significant for both men and women were constrained to equality, and a Lagrange multiplier test (Sörbom, 1989) was applied to evaluate if any of these constraints should be relaxed and be freely estimated by gender. Finally, this model was compared to a model with all path coefficients constrained to equality, and a model with intercepts constrained, to see if a model with those equal for men and women showed a better fit.

3 | RESULTS AND DISCUSSION

Fit statistics suggest that compared to the overall model across all students, a multigroup model by gender showed equally good fit ($\Delta\chi^2(18) = 15.6$, $p = 0.63$), providing motivation to continue exploration of more specific moderation by gender. Next, all path coefficients that were significant for both men and women were constrained to equality across gender, to test a more parsimonious model. Of those, a Lagrange multiplier test identified that model fit would improve if a single parameter were unconstrained and freely estimated by gender, the path from SAT Math to initial Conceptual Understanding ($\chi^2(1) = 5.9$, $p < 0.05$), and so the constraint by gender was removed for this one path. The resulting partially constrained model showed no decline in fit and a slight improvement in other fit statistics from the fully unconstrained model ($\Delta\chi^2(9) = 7.2$, $p = 0.62$). To test if the remaining parameters were indeed different by gender, we also tested the fit of a model with all path coefficients constrained to equality, and a fully constrained model. Both showed a significant decline in fit (see Table 3), suggesting a model allowing for the remaining paths to differ by gender produced the best fit to the data. Furthermore, many paths showed substantial moderation by gender, including cases in which a path was only significant for one gender. Therefore, this model was selected as the final model (see Figure 2).

3.1 | Prior knowledge and competency beliefs as predictors of retention

3.1.1 | Direct effects

Building from the final outcome (Physics 2 retention), path coefficients showed that for both men and women, grade in Physics 1 is the strongest predictor of retention to Physics 2 ($\beta = 0.43$, $p < 0.001$). Interestingly, there is a small direct effect of Theory of Intelligence in physics on retention, which is significant for women ($\beta = 0.16$, $p < 0.05$) but not for men ($\beta = 0.10$, $p = 0.23$; see Figure 2). This moderated path could explain why prior longitudinal work with a large multicohort sample (Witherspoon et al., 2019) found that women were less likely to persist to the second algebra-based Physics course,

TABLE 3 Preliminary fit statistics for the proposed SEM models (acceptable values in bold)

Model	*df*	RMSEA	CFI	TLI	SRMR	χ^2	$\Delta\chi^2$	Δdf	*p*
Overall	18	0.06	**0.98**	0.95	**0.07**	37.2	—	—	—
Constraints by Gender									
Unconstrained	36	0.06	**0.98**	0.95	**0.07**	52.7	15.5	18	0.625
Significant Paths[a]	45	**0.05**	**0.98**	**0.96**	**0.07**	56.0	7.2	9	0.615
All Paths	51	0.06	**0.97**	0.95	0.09	73.7	15.9	6	0.032
All Paths and Intercepts	70	0.10	0.86	0.86	0.14	163.3	89.6	19	0.000

Note: The bold values indicate acceptable model fit, not statistical significance. These values should not have asterisks.
[a]Final reported model.

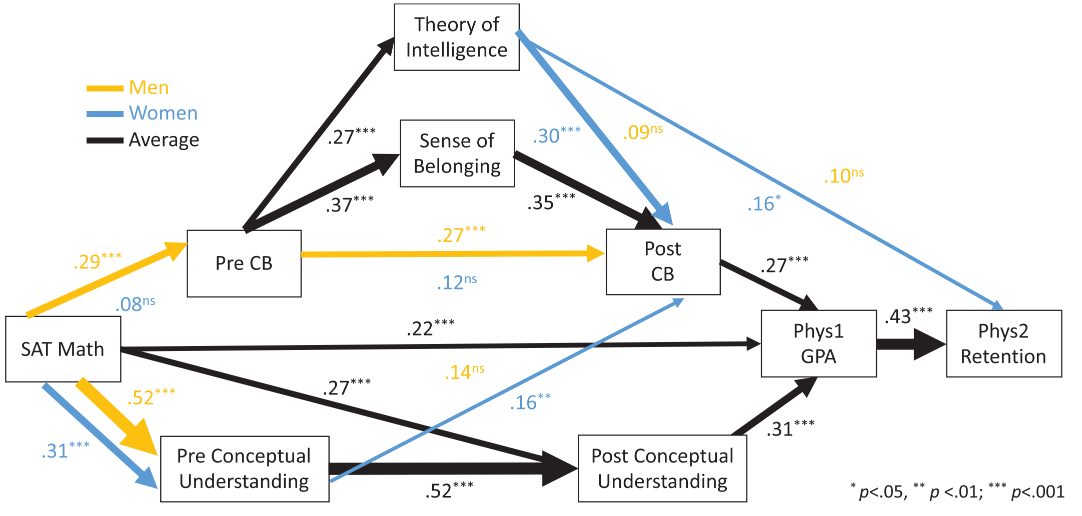

FIGURE 2 The partially constrained path model, with standardized beta coefficients shown. Color paths show gender-specific connection strengths, and black show connection strengths averaged across gender. Line thickness indicate relative strength of effect, dashed lines are nonsignificant. Estimated nonsignificant paths and covariances ($Cov_{(SoB,ToI)} = 0.42$***; $Cov_{(PreCU,PreCB)} = .15^{ns}$; $Cov_{(PostCU,PostCB)} = .08^{ns}$) are not shown in the diagram for interpretability

even when they have equivalent grades. To test for the existence of suppression effects and to examine an alternative mechanism that exists in the literature (Perez et al., 2014; Sawtelle, Brewe, & Kramer, 2012), we also examined a model testing a direct path from Competency Beliefs to Physics retention with Theories of Intelligence not included in the model (see Supporting Information, Figure B1 in Appendix S1); the paths were not significant for men ($\beta = 0.04$, $p = 0.09$) or women ($\beta = 0.09$, $p = 0.07$; see Supporting Information, Table B1 in Appendix S1 for full model output).

3.2 | Prior knowledge and competency beliefs as predictors of physics grade

3.2.1 | Direct effects

Turning to the penultimate outcome (Physics 1 GPA), for both men and women, there were similarly sized direct associations of grade with knowledge on the conceptual Physics exam ($\beta = 0.31$,

$p < 0.001$) and with end-of-course competency beliefs ($\beta = 0.27$, $p < 0.001$; see Figure 2). Prior performance on the SAT Math also contributed to all students' Physics GPA both directly ($\beta = 0.22$, $p < 0.001$), and indirectly as a predictor of Conceptual Understanding at both pre and post. This supports prior work suggesting that Physics is a deeply mathematical science, in which mathematics can act as an important resource for exam performance and conceptual understanding of quantitative laws (Meltzer, 2002).

3.2.2 | Indirect effects

The total indirect effects from initial Conceptual Understanding to Physics 1 grade were significant for both men ($\beta = 0.23$, $p < 0.001$) and women ($\beta = 0.18$, $p < 0.001$); however, for women, a larger proportion of that effect (20% vs. 14%) was through a significant cross-lagged effect of preconceptual understanding on end of course Competency Beliefs. This aspect may reflect perceptions of how easily the content was mastered. Such experiences about effort and perceived success are inherently ambiguous regarding source, and women may be more likely to attribute the amount of effort they exert to their inherent ability, rather than the general difficulty of the content to be mastered (Beyer & Bowden, 1997; LaCosse et al., 2016).

Initial competency beliefs also significantly contributed to Physics 1 GPA indirectly, for both men ($\beta = 0.08$, $p < 0.01$) and women ($\beta = 0.08$, $p < 0.01$). Overall, these significant direct and indirect effects of pre and post competency beliefs on final grades in this course underscores the importance of attending to sources of competency beliefs in this model of student performance, and retention (which is driven by performance; see Table 4 for a summary of these findings).

3.3 | Sources of competency beliefs

3.3.1 | Direct effects

Initial Competency Beliefs was significantly associated with SAT Math score for men ($\beta = 0.29$, $p < 0.001$), but for women this direct effect was not significant ($\beta = 0.08$, $p = 0.27$; see Figure 2). Turning to sources of Competency Beliefs at post, students' broader beliefs about physics were shown to differentially contribute to their end-of-course competency beliefs by gender. In particular, for both men and women, their sense of belonging in Physics at the end of the course was a significant contributor to their end-of-course Competency Beliefs ($\beta = 0.35$, $p < 0.001$). By contrast, students' Theory of Intelligence in physics was only a significant contributor to end-of-course Competency Beliefs for women ($\beta = 0.30$, $p < 0.001$), and not for men ($\beta = 0.09$, $p = 0.30$; see Figure 2).

Interestingly, for both men and women, the direct covariations of Competency Beliefs and Conceptual Understanding were small and not statistically significant at pre ($\beta_{\text{Men}} = 0.18$, $p = 0.06$; $\beta_{\text{Women}} = 0.12$, $p = 0.10$) or post ($\beta_{\text{Men}} = -0.02$, $p = 0.85$; $\beta_{\text{Women}} = 0.13$, $p = 0.08$). Overall, with the exception of the one cross-lagged connection for women from pre Conceptual Understanding to post Competency Beliefs, Competency Beliefs and Conceptual Understanding were largely independent. The moderately-sized concurrent Pearson correlations appear therefore to reflect indirect relations through prior academic experiences and performance: they are

TABLE 4 Summary of direct and indirect results of the path model, reported as standardized betas

Path			Gender differences		No gender differences
Direct effects			***Men***	***Women***	
SAT Math		→ CB (Pre)	0.29***	0.08[ns]	
SAT Math		→ FCI (Pre)	0.52***	0.31***	
CB (Pre)		→ TOI			0.27***
CB (Pre)		→ BEL			0.37***
BEL		→ CB (Post)			0.35***
TOI		→ CB (Post)	0.09[ns]	0.30***	
FCI (Pre)		→ CB (Post)	0.14[ns]	0.16**	
FCI (Pre)		→ FCI (Post)			0.52***
SAT Math		→ FCI (Post)			0.27***
FCI (Post)		→ Physics 1 GPA			0.31***
CB (Post)		→ Physics 1 GPA			0.27***
SAT Math		→ Physics 1 GPA			0.22***
Physics 1 GPA		→ Physics 2 Retention			0.43***
TOI		→ Physics 2 Retention	0.10[ns]	0.16*	
Indirect effects			*Men*	*Women*	
SAT Math[a]		→ Physics 1 GPA			0.16**
FCI (Pre)[a]		→ Physics 1 GPA	0.23***	0.18***	
CB (Pre)[a]		→ CB (Post)	0.27***	0.12[ns]	
CB (Pre)	→ TOI	→ CB (Post)	0.02[ns]	0.08**	
CB (Pre)	→ BEL	→ CB (Post)			0.13**

Abbreviations: CB, competency beliefs; TOI, theories of intelligence; BEL, sense of belonging; FCI, physics conceptual understanding; ns, not significant.
[a]Total indirect effects.
**p* < 0.05, ***p* < 0.01, ****p* < 0.001.

useful for doing well and act as sources of initial higher Competency Beliefs, but the Competency Beliefs are not based in any kind of veridical, direct self-assessment.

3.3.2 | Mediation effects by gender

Overall, when including these measures of students' broader beliefs in Physics as mediators, the association between initial to end-of-course competency beliefs remained significant for men ($\beta = 0.27$, $p < 0.001$), while these factors partially explain the association of pre- to post-competency beliefs for women, such that the direct association between pre- and post-competency beliefs was no longer significant ($\beta = 0.12$, $p = 0.053$). In other words, broader beliefs about Physics mediated only 30% of the total effect of initial competency beliefs on end-of-course competency beliefs for men, while those same factors mediated 59% of the total effect for women. Importantly, results suggest that while sense of Belonging is important for both men and women (mediating 26% and 43%, respectively) Theories of Intelligence in physics were only significantly important for maintaining women's competency beliefs, mediating 24% of the total effect on end-of-course competency beliefs for women, compared to only 8% for men (see Table 4 for a summary of these findings).

4 | GENERAL DISCUSSION

Despite recent reports that suggest growing gender *equality* in science (i.e., equal numbers of men and women earning degrees overall), attending to *equity* in science (i.e., equal opportunities for men and women to earn any degree) must address gender differences in retention across specific science domains, and in particular the physical sciences, where large gender differences in participation persist. In this study, we examine one potential source of continued inequity in a particularly large and gender diverse population of students enrolling in introductory physical science courses: undergraduate premedical students enrolled in an introductory Algebra-based Physics course. Our results show that retention in this physical science course for both men and women is primarily driven by prior grades; however, that direct effect is heavily driven by competency beliefs, centering its role in both grades and retention.

The finding that actual understanding of physics content is poorly aligned with students' competency beliefs in physics further draws attention to the importance of competency beliefs as its own target of intervention (i.e., it is not enough to just improve understanding of the content; students need to feel competent, too). Interestingly, our findings suggest that competency beliefs contribute indirectly to retention through their effect on grades, supporting some prior findings in the literature (Perez et al., 2014; Sawtelle, Brewe, Goertzen, & Kramer, 2012; Kalender et al., 2020; Li & Singh, 2021) while failing to replicate the results found with high performing students in the context of an Organic Chemistry course (Witherspoon et al., 2019). It may be for this group of high-performing women, the reduced variance in performance feedback made internal attributions of success like perceived effort and ability beliefs more salient than grades, while this was not the case in a sample with a broader performance distribution (Bar-Tal & Frieze, 1977; Beyer & Bowden, 1997). While our sample here was underpowered to perform this type of analysis by performance level, future research with larger samples could further test whether these attitudinal mechanisms differ as a function of the particular discipline, course timing, achievement level, or some combination of these factors.

Just as importantly in the current study, competency beliefs were found to have both overlapping and nonoverlapping sources by gender. Sense of belonging appeared as an important source of ability beliefs for all students (although women were less likely to have a high sense of belonging). However, only for women, the extent to which they believe Physics intelligence to be fixed or malleable acts as a strong predictor of changes in competency beliefs from the beginning to the end of the course.

These results are particularly informative in an introductory physical science course for nonmajors, where stereotype threats for women may be especially salient. Recent studies have suggested that for upper-secondary students, spending additional years in a science program was associated with theories of intelligence becoming more fixed (Jonsson & Beach, 2017). Students who enroll in introductory algebra-based Physics courses with the intent to study premedicine, a larger proportion of whom are women, may enter with less exposure to these disciplines relative to students who enter with an express intent in pursuing science degrees (e.g., Physics or Engineering). Therefore, the enculturation and shifts in beliefs about the nature of intelligence in these disciplines may be more salient for premed students, in ways that are detrimental to their competency beliefs. Furthermore, in a predominantly male Physics environment, it may be possible that women who enter those undergraduate courses were able to persist to that point because they have been well-supported and encouraged through earlier educational experiences by high school teachers based on their relatively high aptitude in Mathematics and Science (Seymour, 1995). However, upon entering a more selective pool of students and confronting more difficult content in college, individuals with high initial competency beliefs may experience a "Little Fish, Big Pond" effect (Marsh et al., 2008), where their prior assessment of their abilities relative to their peers is recalibrated.

For women in an historically male-dominated discipline like Physics, this can lead to attribution biases, where women are more likely to attribute a decline in performance to their own lack of ability, whereas men may get their ability beliefs from a broader variety of sources (Beyer, 1990; LaCosse et al., 2016). If sense of belonging and theories of intelligence mediate the stability of competency beliefs for women, and lower competency beliefs are associated with lower grades, attitudes and broader beliefs about Physics can lead to a self-fulfilling prophecy that operates separately from ability. That is, women who feel like they do not belong in Physics and believe that negative performance feedback is a stable indication of their ability could experience declining grades, further solidifying these attitudes and beliefs. At the same time, men who do not experience this mediation effect may experience a relative boost in self-efficacy, which can also continue to perpetuate stereotypes of male-dominance in Physics.

It is interesting that our findings did not show gender differences in effects of belonging on retention, supporting prior work with similar results in Calculus-based Physics (Stout et al., 2013). It may be that as a later course along the premed sequence, there has already been a significant "weeding out" process of students with low belonging, who do not persevere to the point of taking Physics. Alternatively, work examining gender differences in retention has proposed a relative strengths hypothesis for high performing students, suggesting that not lack of ability, but instead high ability in other areas, may better predict women's choices to pursue alternative majors and career trajectories outside of STEM (Te Wang et al., 2013). Similarly, future work could test a hypothesis about "relative" belonging, to see if this framing demonstrates gender differences not in overall belonging, but instead relative belonging for men and women in the physical sciences, as compared to their concurrent sense of belonging in other fields (Thoman et al., 2014). Indeed recent work has suggested that sense of belonging for women in physics is context-dependent and particularly relevant majors in higher-level physics classes (Hazari, et al., 2020), potentially explaining the lack of an effect for premedical students who are less likely to identify as members of the physics community.

4.1 | Limitations

It is important to acknowledge certain limitations of the current study. First, because the data collected and analyses conducted were correlational, the level of casual inference that can be drawn from the relations among our variables found here is reduced. Interventions with a control group targeting particular points within the model would provide a stronger casual test of the hypotheses derived from the current findings. However, the longitudinal data collected and the structural modeling method used do allow for some inference about the temporal ordering of these effects. For example, it is not possible for the directionality of an effect of later competency beliefs on prior performance on the SAT to be reversed, or that later persistence decisions could affect theories of intelligence measured during a previous course. Therefore, this study builds empirical support for the directionality of these effects.

Second, we selected available measures to act as operationalizations of sources of self-efficacy beliefs (i.e., mastery experiences; vicarious experiences and social persuasions). However, because these measures do not directly assess student's perceptions, they make assumptions about students' perceptions of their experiences. For example, it is possible that students have other sources than the SAT and course exams that contribute to their perception of mastery in a domain. However, as a practically important and widely accepted measure of academic performance, it is likely that students' are aware of their performance on their math SATs and that their interpretation of this performance is likely to provide strong feedback to them on their relative mastery of a mathematically-oriented science.

Third, while we were able to demonstrate that the selected model was a good fit to the data, structural equation modeling allows for a near infinite number of alternative model specifications, some of which might also adequately fit the data and would therefore provide alternative theoretical explanations. Our model building process addressed and rejected the most obvious alternative explanations for our particular research questions; future work testing similar models will help to corroborate these results. For example, though not reported here, some preliminary analyses conducted to test the influence of recitation-level factors (i.e., variation of the proportion of women in a recitation section) showed inconclusive results, possibly due to a sample size which is lower than required to do these kinds of robust multilevel analyses by gender. However, understanding the effects of variation in course and recitation factors such as gender composition or relative attitudes and achievement could be a fruitful next step in understanding sources of students' development of competency beliefs within these introductory courses.

Finally, the current study was conducted with a particular population of mostly premedical students in an Algebra-based Physics course, and data was collected at a single institution, limiting the potential generalizability of the results. Therefore, replication at other similar institutions and within other premed science courses should be conducted to see if these conditions alter the results found here. In addition, a multi-institutional approach would allow for intentional selection of variation of elements of the learning environment (i.e., proportion of women faculty, size of course sections, selectivity of the institution) to see if these also offer alternative explanations for the results found.

4.2 | Implications for practice

Practically, our results provide support for interventions in introductory physical science courses at the undergraduate level that are designed to address both students' sense of belonging, and their implicit theories of intelligence about the nature of ability for prehealth and premedical students in the physical sciences. While prior preparation does play a role in premedical students' grades (and thereby retention) in physical science courses, our findings suggest that the attitudes and beliefs about these domains that are developing concurrent during undergraduate introductory courses contribute an equal amount to these factors. This is important in that it demonstrates that the responsibility of improving achievement and retention for undergraduate students in physical science courses does not solely fall on primary and secondary educators, but also will require input from higher education institutions in creating environments that support all students in developing and maintaining attitudes that contribute to their success in those courses.

For example, instructors of the physical sciences in higher education institutions should consider attending to not only developing understanding of the content, but also the development of competency beliefs in the content, while being aware that these may develop independently from actual content understanding, and differentially for men and women. Specifically, our results show that interventions targeting improving a sense of belonging are likely to contribute to the development of competency beliefs for all students; examples of such interventions exist, and have demonstrated positive effects on grades and attitudes (Blackwell et al., 2007; Walton et al., 2015). However, this alone is unlikely to address the persistence gender gap found in the physical sciences; interventions may need to be tailored to the unique goals and concerns of nonmajors in the physical sciences. For example, the large population of prehealth and premedical students who enter these classes is a particularly important group to consider because interventions targeting them could provide larger effects and contribute to a larger number of students, and women in particular, who choose to continue in the sciences rather than leave to pursue alternative careers outside of STEM. Instead, providing experiences that contribute to a greater sense of belonging in the physical sciences, in combination with demonstrating that ability in the physical sciences is not fixed but instead can be improved through effort, could be especially impactful for the development of competency beliefs and improved retention for women in premedical Physics courses.

5 | CONCLUSIONS

Domain-specific investigations of gendered attrition in STEM point out the dangers of the "STEM pipeline" metaphor that is often used to describe the issue of underrepresentation in STEM fields. The pipeline metaphor does not accurately model or locate the discipline-specific sources of differences in grades, ability beliefs, and retention for women. In addition, depicting the issue in undergraduate education as simply a deficit of incoming preparation of underrepresented STEM majors can have the unintended effect of prescribing solutions such as increasing applicants, which do not address the broader range of social and environmental deterrents women experience during their undergraduate education (Blickenstaff, 2005). Longitudinal studies of the interactive effects of grades and motivations for underrepresented groups can help to pinpoint the timing and character of these negative experiences and suggest interventions that may alleviate some of the nonacademic deterrents for women in entering various STEM careers.

Furthermore, it is important that research continues to gather data on contextual factors in particular sources of students' sense of belonging and theories of intelligence in the physical sciences. Qualitative studies may be needed to uncover more nuanced understanding of other motivational aspects beyond those measured here that may influence subsequent grades and persistence (Seymour, 1995; Seymour & Hunter, 2019). For example, studies show that undergraduate teaching faculty are more likely to rate fields with fewer women as domains that require "innate brilliance," with Physics as one of those domains with both the highest professor ratings of "brilliance required," and the fewest women who persist to earn degrees (Leslie et al., 2015). Historical underrepresentation of women in these fields can lead to a cementing of these discriminatory associations between talent and gender, and heightened stereotype threat for women in those courses. Importantly, interventions in social belonging and theories of intelligence have shown these beliefs to be highly malleable (Blackwell et al., 2007; Walton et al., 2015); therefore confirmation of a gendered link between these and self-efficacy provide a promising avenue for interventions focused on increasing women's self-efficacy beliefs in Physics classrooms. Finally, rather than continuing to target only improvement in performance, which places students as the primary source of change, further research in this area can help to understand the role faculty (Zohar & Bronshtein, 2005) and institutions can play in mitigating the "chilly climate" (Walton et al., 2015) that deters women from pursuing the physical sciences at the undergraduate level.

ACKNOWLEDGMENTS

This study was supported by research grant Division of Undergraduate Education (DUE) 1524575 from the National Science Foundation. The authors would like to thank Elizabeth Votruba-Drzal, David I. Miller, and Kevin R. Binning for their feedback and suggestions on earlier drafts of this work.

ORCID

Eben B. Witherspoon https://orcid.org/0000-0003-0987-3679

ENDNOTE

[1] Calculus-Based Introductory Physics courses, not included in these analyses, typically enroll engineering students and are often predominantly (~70%) male.

REFERENCES

(2020). *AAMC facts table (tab. A-1)*. Association of American Medical Colleges. https://www.aamc.org/system/files/2019-11/2019_FACTS_Table_A-1.pdf

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

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company.

Barr, D. A. (2010). *Questioning the premedical paradigm: Enhancing diversity in the medical profession a century after the Flexner report*. Johns Hopkins University Press.

Barr, D. A., Matsui, J., Wanat, S. F., & Gonzalez, M. E. (2010). Chemistry courses as the turning point for premedical students. *Advances in Health Sciences Education*, *15*(1), 45–54. https://doi.org/10.1007/s10459-009-9165-3

Bar-Tal, D., & Frieze, I. H. (1977). Achievement motivation for males and females as a determinant of attributions for success and failure. *Sex Roles*, *3*(3), 301–313. https://doi.org/10.1007/BF00287617

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. (2011). *Women in STEM: A gender gap to innovation* (Issue Brief 04-11, U.S. Department of Commerce, ESA). https://doi.org/10.2139/ssrn.1964782

Bench, S. W., Lench, H. C., Liew, J., Miner, K., & Flores, S. A. (2015). Gender gaps in overestimation of math performance. *Sex Roles*, *72*(11–12), 536–546. https://doi.org/10.1007/s11199-015-0486-9

Beyer, S. (1990). Gender differences in the accuracy of self- evaluations of performance. *Journal of Personality and Social Psychology*, *59*(5), 960–970. https://doi.org/10.1037/0022-3514.59.5.960

Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perceptions: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin*, *23*(2), 157–172. https://doi.org/10.1177/0146167297232005

Blackwell, K. L., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, *17*(4), 369–386. https://doi.org/10.1080/09540250500145072

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching*, *45*(9), 971–1002. https://doi.org/10.1002/tea.20241

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education*, *98*(3), 443–460. https://doi.org/10.1002/sce.21108

Carnevale, A. P., Cheah, B., & Hanson, A. R. (2015). *The economic value of college majors*. Washington, DC: Georgetown University, Center on Education and the Workforce.

Chen, X. (2013). STEM attrition: College Students' path into and out of STEM fields (NCES 2014-001), 1–52. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, *24*, 11–21. https://doi.org/10.1016/j.lindif.2012.11.002

Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2016). Why are some STEM field more gender balanced than others? *Psychological Bulletin*, *142*(206), 1–131. https://doi.org/10.1037/bul0000052

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Conger, D., & Long, M. C. (2010). Why are men falling behind? Gender gaps in college performance and persistence. *The Annals of the American Academy of Political and Social Science*, *627*(1), 184–214. https://doi.org/10.1177/0002716209348751

Cromley, J. G., Perez, T., & Kaplan, A. (2016). Undergraduate STEM achievement and retention: Cognitive, motivational, and institutional factors and solutions. *Policy Insights From the Behavioral and Brain Sciences*, *3*(1), 4–11. https://doi.org/10.1177/2372732215622648

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *90*(4), 666–679. https://doi.org/10.1037/0022-3514.90.4.666

Debacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: Differences related to gender, class type, and ability. *Journal of Educational Research*, *93*(4), 245–254. https://doi.org/10.1080/00220670009598713

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment. In J. G. Nicholls (Ed.), *The development of achievement motivation* (Vol. 3, pp. 238–331). JAI Press, Inc.

Eddy, S. L., & Brownell, S. E. (2016). Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, *12*(2), 1–20. https://doi.org/10.1103/PhysRevPhysEducRes.12.020106

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, *53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

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*(4), 700–717. https://doi.org/10.1037/a0026659

Hong, Y.-Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–299. https://doi.org/10.1037/0022-3514.77.3.588

Hu, L., & 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), 1–55. https://doi.org/10.1080/10705519909540118

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in Students' experiences, interests, and attitudes toward science and scientists. *Science Education*, *84*, 180–192. https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<180::AID-SCE3>3.0.CO;2-X

Jonsson, A. C., & Beach, D. (2017). The influence of subject disciplinary studies on students' implicit theories of intelligence and achievement goals in one Swedish upper-secondary school. *Education Inquiry*, *8*(1), 50–67. https://doi.org/10.1080/20004508.2016.1275182

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2020). Damage caused by women's lower self-efficacy on physics learning. *Physical Review Physics Education Research*, *16*(1), 010118.

King, B. (2016). Does postsecondary persistence in STEM vary by gender? *AERA Open*, *2*(4), 233285841666970. https://doi.org/10.1177/2332858416669709

LaCosse, J., Sekaquaptewa, D., & Bennett, J. (2016). STEM stereotypic attribution bias among women in an unwelcoming science setting. *Psychology of Women Quarterly*, *40*(3), 378–397. https://doi.org/10.1177/0361684316630965

Legewie, J., & DiPrete, T. A. (2014). The high school environment and the gender gap in science and engineering. *Sociology of Education*, *87*(4), 259–280. https://doi.org/10.1177/0038040714547770

Lent, R. W., Ireland, G. W., Penn, L. T., Morris, T. R., & Sappington, R. (2017). Sources of self-efficacy and outcome expectations for career exploration and decision-making: A test of the social cognitive model of career self-management. *Journal of Vocational Behavior*, *99*, 107–117. https://doi.org/10.1016/j.jvb.2017.01.002

Lent, R. W., Lopez, F. G., Brown, S. D., & Gore, P. A. (1996). Latent structure of the sources of mathematics self-efficacy. *Journal of Vocational Behavior*, *49*(3), 292–308. https://doi.org/10.1006/jvbe.1996.0045

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., & Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: A meta-analytic path analysis of the social–cognitive choice model by gender and race/ethnicity. *Journal of Counseling Psychology*, *65*(1), 17–35. https://doi.org/10.1037/cou0000243

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 23–34. https://doi.org/10.1081/E-EWS

Lewis, K. L., Stout, J. G., Finkelstein, N. D., Pollock, S. J., Miyake, A., Cohen, G. L., & Ito, T. A. (2017). Fitting in to move forward: Belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychology of Women Quarterly*, *41*(4), 420–436. https://doi.org/10.1177/0361684317720186

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*(2), 1–10. https://doi.org/10.1103/PhysRevPhysEducRes.12.020110

Li, Y., & Singh, C. (2021). Effect of gender, self-efficacy, and interest on perception of the learning environment and outcomes in calculus-based introductory physics courses. *Physical Review Physics Education Research*, *17*(1), 010143.

Lord, S. M., Camacho, M. M., Layton, R. A., Long, R. A., Ohland, M. W., & Wasburn, M. H. (2009). Who's persisting in engineering? A comparative analysis of female and male Asian, black, Hispanic, native American, and white students. *Journal of Women and Minorities in Science and Engineering*, *15*(2), 167–190. https://doi.org/10.1615/JWomenMinorScienEng.v15.i2.40

Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology*, *100*(3), 510–524. https://doi.org/10.1037/0022-0663.100.3.510

Marshman, E., Kalender, Z. Y., Schunn, C., Nokes-Malach, T., & Singh, C. (2017). A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics*, *96*(4), 391–405. https://doi.org/10.1139/cjp-2017-0185

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, *37*(2), 225–238. https://doi.org/10.1016/0001-8791(90)90042-Z

Matz, R. L., Koester, B. P., Fiorini, S., Grom, G., Shepard, L., Stangor, C. G., Weiner, B., & McKay, T. A. (2017). Patterns of gendered performance differences in large introductory courses at five research universities. *AERA Open*, *3*(4), 1–12. https://doi.org/10.1177/2332858417743754

McMahon, S. D., & Wernsman, J. (2008). The relation of classroom environment and school belonging to academic self-efficacy among urban fourth- and fifth-grade students. *The Elementary School Journal*, *109*(3), 267–281. https://doi.org/10.1086/592307

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics*, *70*(12), 1259–1268. https://doi.org/10.1119/1.1514215

Miller, P. H., Blessing, J. S., & Schwartz, S. (2006). Gender differences in high-school students' views about science. *International Journal of Science Education*, *28*(4), 363–381. https://doi.org/10.1080/09500690500277664

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, *42*(4), 989–1005. https://doi.org/10.1016/j.ssresearch.2013.03.008

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30–38. https://doi.org/10.1037/0022-0167.38.1.30

National Science Board. (2020). Science and Engineering Indicators 2020: The State of U.S. Science and Engineering (NSB-2020-1). Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/pubs/nsb20201/.

National Center for Science and Engineering and Statistics (2021). Women, Minorities, and Persons with Disabilities in Science and Engineering: 2021, (Special Report NSF 21-321). Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/wmpd.

Ohland, M. W., Sheppard, S. D., Lichtenstein, G., Eris, O., Chachra, D., & Layton, R. A. (2008). Persistence, engagement, and migration in engineering programs. *Journal of Engineering Education*, *97*(3), 259–278. https://doi.org/10.1002/j.2168-9830.2008.tb00978.x

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*, *42*(1), 104–120.

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology*, *106*(1), 315–329. https://doi.org/10.1037/a0034027

Perez, T., Wormington, S. V., Barger, M. M., Schwartz-Bloom, R. D., Lee, Y., & Linnenbrink-Garcia, L. (2019). Science expectancy, value, and cost profiles and their proximal and distal relations to undergraduate science, technology, engineering, and math persistence. *Science Education*, *2018*, 264–286. https://doi.org/10.1002/sce.21490

Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, *88*(3), 408–422. https://doi.org/10.1037/0022-0663.88.3.408

Rosenthal, L., Levy, S. R., London, B., Lobel, M., & Bazile, C. (2013). In pursuit of the MD: The impact of role models, identity compatibility, and belonging among undergraduate women. *Sex Roles*, *68*(7–8), 464–473. https://doi.org/10.1007/s11199-012-0257-9

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, *96*(3), 411–427. https://doi.org/10.1002/sce.21007

Sanabria, T., & Penner, A. (2017). Weeded out? Gendered responses to failing calculus. *Social Science*, *6*(2), 1–14. https://doi.org/10.3390/socsci6020047

Savinainen, A., & Scott, P. (2002). The force concept inventory: A tool for monitoring student learning. *Physics Education*, *37*(1), 45–52. https://doi.org/10.1088/0031-9120/37/1/306

Sawtelle, V., Brewe, E., Goertzen, R. M., & Kramer, L. H. (2012). Identifying events that impact self-efficacy in physics learning. *Physical Review Special Topics - Physics Education Research*, *8*(2), 1–18. https://doi.org/10.1103/PhysRevSTPER.8.020111

Sawtelle, V., Brewe, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching*, *49*(9), 1096–1121. https://doi.org/10.1002/tea.21050

Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education*, *79*(4), 437–473. https://doi.org/10.1002/sce.3730790406

Seymour, E., & Hunter, A.-B. (2019). *Talking about leaving revisited persistence, relocation, and loss in undergraduate STEM education*. Springer.

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard Isn't natural. *Personality and Social Psychology Bulletin*, *39*(2), 131–143. https://doi.org/10.1177/0146167212468332

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*(3), 371–384.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stevens, T., Wang, K., Olivárez, A., & Hamman, D. (2007). Use of self-perspectives and their sources to predict the mathematics enrollment intentions of girls and boys. *Sex Roles*, *56*(5–6), 351–363. https://doi.org/10.1007/s11199-006-9180-2

Stout, J. G., Ito, T. A., Finkelstein, N. D., & Pollock, S. J. (2013). How a gender gap in belonging contributes to the gender gap in physics participation. *AIP Conference Proceedings*, *1513*(1), 402–405. https://doi.org/10.1063/1.4789737

Te Wang, M., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science*, *24*(5), 770–775. https://doi.org/10.1177/0956797612458937

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED majors. *Sex Roles*, *77*(1–2), 86–96. https://doi.org/10.1007/s11199-016-0694-y

Thoman, D. B., Arizaga, J. A., Smith, J. L., Story, T. S., & Soncuya, G. (2014). The grass is greener in non-science, technology, engineering, and math classes: Examining the role of competing belonging to undergraduate Women's vulnerability to being pulled away from science. *Psychology of Women Quarterly*, *38*(2), 246–258. https://doi.org/10.1177/0361684313499899

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring Students' self-efficacy, sense of belonging, and science identity. *CBE—Life Sciences Education*, *13*(1), 6–15. https://doi.org/10.1187/cbe.13-12-0241

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. https://doi.org/10.3102/0034654308321456

Vincent-Ruz, P., Binning, K., Schunn, C. D., & Grabowski, J. (2018). The effect of math SAT on women's chemistry competency beliefs. *Chemistry Education Research and Practice*, *19*, 342–351. https://doi.org/10.1039/C7RP00137A

Vincent-Ruz, P., & Schunn, C. D. (2017). The increasingly important role of science competency beliefs for science learning in girls. *Journal of Research in Science Teaching*, *54*(6), 790–822. https://doi.org/10.1002/tea.21387

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96. https://doi.org/10.1037/0022-3514.92.1.82

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "Chilly Climate" transform Women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*(2), 468–485. https://doi.org/10.1037/a0037461

Witherspoon, E. B., & Schunn, C. D. (2019). Locating and understanding the largest gender differences in pathways to science degrees. *Science Education*, *104*(2), 1–20. https://doi.org/10.1002/sce.21557

Witherspoon, E. B., Vincent-Ruz, P., & Schunn, C. D. (2019). When making the grade Isn't enough: The gendered nature of premed science course attrition. *Educational Researcher*, *48*(4), 1–12. https://doi.org/10.3102/0013189X19840331

Xie, Y., & Shauman, K. A. (2003). *Women in science: Career processes and outcomes*. Harvard University Press.

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal*, *37*(1), 215–246. https://doi.org/10.3102/00028312037001215

Zohar, A., & Bronshtein, B. (2005). Physics teachers' knowledge and beliefs regarding girls' low participation rates in advanced physics classes. *International Journal of Science Education*, *27*(1), 61–77. https://doi.org/10.1080/0950069032000138798

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Witherspoon, E. B., & Schunn, C. D. (2022). Sources of gender differences in competency beliefs and retention in an introductory premedical science course. *Journal of Research in Science Teaching*, *59*(5), 695–719. https://doi.org/10.1002/tea.21741