Featured in Physics

Students' sense of belonging in introductory physics course for bioscience majors predicts their grade

Sonja Cwik and Chandralekha Singh
Department of Physics and Astronomy University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

(Received 11 October 2021; accepted 16 March 2022; published 31 May 2022)

Student sense of belonging in physics classes may not only play a key role in shaping course outcomes but also influence student persistence and future career aspirations. Prior research has shown that women have a lower sense of belonging than men in calculus-based introductory physics courses. However, prior research has generally not investigated students' sense of belonging in introductory physics courses in which women are not underrepresented. We administered a validated survey to investigate the sense of belonging of 814 students and how it predicts student grades in a mandatory introductory physics course primarily for bioscience majors. In particular, we investigated how students' sense of belonging predicts female and male students' grade at the end of the mandatory physics course for bioscience majors using structural equation modeling. We found that women had a lower sense of belonging and grade than men in this course and that the students' sense of belonging played a major role in predicting students' grade in the course. In addition, while men's sense of belonging significantly increased from the beginning to the end of the physics course, women's sense of belonging did not significantly change by the end of the course.

DOI: 10.1103/PhysRevPhysEducRes.18.010139

I. INTRODUCTION AND THEORETICAL FRAMEWORK

Motivational beliefs in academic domains can influence students' continuation in related courses, majors, and careers [1–10]. In particular, students' sense of belonging has been shown to be important in their academic outcomes and future careers [11,12]. Students' sense of belonging in their college or university has been shown to predict students' intentions to persist through college [13,14]. In science, technology, engineering, and math (STEM) fields, studies have shown that belonging is linked to students' engagement in the courses [15], grade [16], persistence [17,18], and lower expectations of dropping out of their STEM major [19]. For example, in one study, students' sense of belonging in a math class was found to predict their intention to pursue math in the future [20]. However, it has also been found that women have a lower sense of belonging than men, potentially due to stereotypes related to who belongs and can excel in STEM fields [20,21].

These stereotypes about who can excel in STEM courses could impact women even in physics courses in which they are not underrepresented, e.g., mandatory college physics courses for bioscience majors. One common stereotype is that genius and brilliance are important factors to succeed in physics [22]. However, genius is often associated with boys [23], and girls from a young age tend to shy away from fields associated with innate brilliance or genius [24]. Moreover, as these students get older, norms in the science curriculum hold less relevance for girls since the existing curricula tend not to represent the interests and values of girls as much [25]. Furthermore, teachers and school counselors often pay more attention to male students and counselors give gendered advice to students regarding which high school physics and math courses to take and majors to pursue when in college. All these factors which are associated with societal stereotypes about who belongs in physics and can excel in it can influence female students' perceptions about their ability to do physics even before they enter the college classroom. Also, it is possible that although women are the majority in algebra-based physics courses primarily for bioscience majors, these societal stereotypes can still influence their sense of belonging and grades in the physics class unless instructors make an explicit effort to create an equitable and inclusive learning environment.

One mechanism by which societal stereotypes and biases negatively affect female students is via lowering their sense of belonging in fields in which they are underrepresented. For example, in one study, women STEM graduate students had the perception that they exerted more effort than their peers to succeed, and this perception predicted their sense of belonging [26]. Another study showed that female participants in a Physics Olympiad competition who

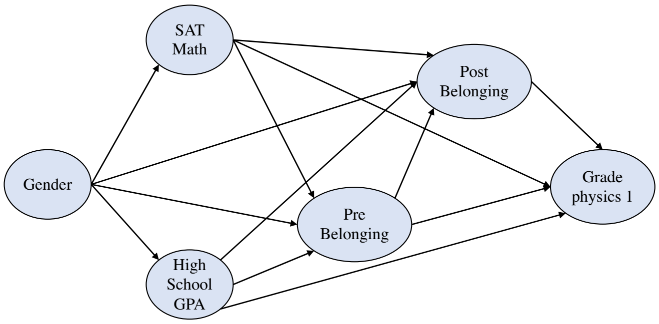

FIG. 1. Schematic representation of the path analysis of the model and how sense of belonging mediates the relation between gender and grade in the physics 1 course for bioscience majors in which women are not underrepresented. From left to right, all possible paths were considered (including the one from gender to grade in physics 1), however, only some of the paths are shown here for clarity.

endorsed negative stereotypes about female talent for physics felt a lower sense of belonging in physics [27]. In addition, when active learning is implemented in STEM courses in which learning environments are not equitable or inclusive, men have been shown to dominate responding to questions from instructors in classes [28]. Students' sense of belonging has been shown to correlate with their retention and self-efficacy in school [29–31]. In addition, it is shown to be a predictor of students' physics identity for senior physics majors [32]. Several prior studies have shown performance gaps between men's and women's grades and scores on conceptual tests in calculus-based and algebra-based physics courses [33–38] that some hypothesize may be due to gender gap in prior preparation or motivational beliefs, such as sense of belonging [37–47]. However, there has not been research conducted on students' sense of belonging and how it predicts students' grades in physics courses in which women are not underrepresented.

This study examined the difference between male and female students with regard to the role of their sense of belonging in predicting performance outcomes at the end of a mandatory first semester of an algebra-based introductory physics course sequence for bioscience majors, controlling for their high school GPA, SAT math scores, and sense of belonging at the start of the class. We controlled for students' past performance in high school in order to isolate the effect of students' sense of belonging on their course outcomes. A visual representation of our final model is shown in Fig. 1. All paths were considered from left to right in our model. However, only some of the paths are shown for clarity. Our research questions are as follows:

RQ1 Are there gender differences in students' high school GPA and SAT math scores, as well as grades and sense of belonging at the beginning and end of the mandatory physics course?

RQ2 To what extent does physics sense of belonging predict physics course grade in the model shown schematically in Fig. 1?

II. METHODOLOGY

In this study, a validated survey covering sense of belonging and other motivational constructs (not discussed here) was administered to students at a large public research university in the U.S. The survey was given at the beginning (pre) and end (post) of the first semester in the introductory algebra-based physics classes over the course of two years that were taught in the traditional in-person lecture style before the pandemic. This course is primarily taken by junior or senior bioscience majors for whom a two-semester physics course sequence is mandatory. We analyzed the data for 814 students who completed the survey in the introductory physics 1 class. The university provided demographic information such as age, gender, and ethnic or racial information using an honest broker process by which the research team received the information without knowledge of the identities of the participants. From the university data, the participants were 36% male and 64% female. We recognize that gender is fluid and not a binary construct; however, the data collected by the institution is in binary terms. We use binary data provided by the university in this study. Less than 1% of the students did not choose male or female and thus were not included in the study.

We note that there were six sections and four instructors of the course within the investigated time period. However, we found that the intraclass correlation coefficient (ICC), which measures the proportion of the variance between instructors, e.g., in the student sense of belonging to be 4%. This is below 10%, the usual threshold cited for warranting the use of multi-level models and thus we grouped all the students together, regardless of instructor.

A. Instrument validity

This study measured the physics sense of belonging of students enrolled in the first of the two introductory algebra-based physics courses primarily for bioscience majors. The survey with multiple motivational constructs was administered to the students [48], however, we focus on the sense of belonging questions here. The validated survey involving several motivational beliefs, including the sense of belonging items, was adapted from previous research [49,50]. Re-validation of the survey at our institution involved conducting one-on-one student interviews [10], exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and calculation of the Pearson correlations and Cronbach's alpha. The sense of belonging items on the survey measured whether students felt like they belonged in the introductory physics class [30,50]. The questions in the study were designed on a Likert scale

TABLE I. Survey items for students' sense of belonging along with factor loadings (Lambda) from the confirmatory factor analysis result for all students ($N = 814$). The rating scale for the items was not at all true, a little true, somewhat true, mostly true, and completely true. All p values (of the significance test of each item loading) are $p < 0.001$.

Physics belonging items	Lambda
I feel like I belong in this class.	0.77
I feel like an outsider in this class.	0.77
I feel comfortable in this class.	0.77
Sometimes I worry that I do not belong in this class.	0.75

of 1 (low endorsement) to 5 (high endorsement) [51]. A lower score is indicative of a lower sense of belonging while a higher score is indicative of a higher sense of belonging. Two of the items were reverse coded. The survey items for the sense of belonging factor and factor loadings for each item from the CFA, performed on the entire survey, are given in Table I. Cronbach's alpha was used to measure the internal consistency of the items. Cronbach's alpha is 0.85 for the sense of belonging items, which is considered reasonable [52].

B. Analysis

Initially, we analyzed the descriptive statistics by comparing female and male students' mean scores on the belonging questions and students' grades for statistical significance using t tests and investigated the effect size using Cohen's d [53]. Cohen's d is $d = (\mu_m - \mu_f)/\sigma_{\text{pooled}}$, where μ_m is the average score of male students, μ_f is the average score of female students, and σ_{pooled} is the pooled standard deviation (or weighted standard deviation for men and women) for all students. In general, $d = 0.20$ indicates a small effect size, $d = 0.50$ indicates a medium effect size, and $d = 0.80$ indicates a large effect size [53].

To quantify the statistical significance and relative strength of our framework's path links, we used structural equation modeling (SEM) as a statistical tool by using R (lavaan package) with a maximum likelihood estimation method [54]. SEM is a statistical method consisting of two parts that are completed together; a measurement part which consists of CFA and a structural part that consists of path analysis. Path analysican be considered an extension of multiple regression analysis, but it allows one to conduct several multiple regressions simultaneously between variables in one estimation model and allows us to predict multiple outcomes simultaneously. SEM also allows us to calculate the overall goodness of fit and for all estimates to be standardized simultaneously so there can be a direct comparison between different structural components. Thus, we are able to test more complicated models than we would with multiple regression analysis. A full SEM model combines this path analysis with CFA, allowing researchers to test the validity of their constructs (using CFA) and the connections between these constructs (using path analysis) in a single model with a single set of fit indices. We report the model fit for SEM by using the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residuals (SRMR). Commonly used thresholds for goodness of fit are as follows: CFI and TLI > 0.90, and RMSEA and SRMR < 0.08 [55].

Initially, we performed gender moderation analysis by conducting multigroup SEM, i.e., the model estimates were performed separately for men and women to check whether any of the relations between variables show differences across gender by using lavaan [56]. In particular, our moderation analysis was similar to our mediation model in Fig. 1 but there was no link from gender, instead, multigroup SEM was performed separately for women and men simultaneously.

In order to explain what moderation analysis means, we start with a simple moderation analysis example. In a simple moderation analysis involving the predictive relation between only two variables, the predictive relationship (the regression path) between those two variables is tested for two or more different groups (e.g., men and women) simultaneously. If the predictive relationship is different for the groups [i.e., the values of the regression coefficients (β) are not the same for the correlation between the two constructs for different groups], then there is a moderation effect in the model. For example, in a study focusing on how smoking predicts lung cancer, if there was a moderation effect by gender, the predictive relation (regression coefficient) between smoking and lung cancer would be different for women and men. However, if the regression coefficients for how smoking predicts lung cancer were exactly the same for women and men, then there is no moderation by gender and one can just focus on mediation analysis by gender (in other words, we need not separately calculate the regression coefficients for women and men since they are equal, and we can introduce gender as an additional categorical variable in the model to do mediation by gender).

When the model is more complex than the preceding example of smoking and lung cancer as in our SEM model (which has a measurement part involving CFA and a structural part involving path analysis), checking to make sure there are no gender moderation effects involves checking that there are no gender moderation effects for both the measurement and structural parts. For the measurement part, to check for measurement invariance in each step of gender moderation analysis, we fixed different elements of the measurement part of the model to equality across gender and compared the results to the previous step when they were allowed to vary between groups (i.e., for women and men) separately using the likelihood ratio test [56]. A nonsignificant p value at each step indicates that the fit of this model is not appreciably worse than that of the model in the previous step, so the more restrictive invariance hypothesis (when the parameters are set to the same values for women and men) is retained. Therefore, setting those different elements of the measurement part of the model to equality across gender is valid, which means that estimates are not statistically significantly different across groups (i.e., women and men).

First, we tested for "weak" measurement invariance, which determines if survey items have similar factor loadings for men and women. We compared two models, one in which the factor loadings (which represent the correlation between each item and its corresponding construct) for women and men were predicted independently, and the other in which the factor loadings were forced to be equal between the groups (i.e., for women and men). Next, we tested for "strong" measurement invariance, which determines if survey items have similar factor loadings as well as similar intercepts (which represent the expected value of an observed variable when its associated latent variable is equal to zero) for men and women. Similar to weak invariance testing, we compared the models in which these factors were allowed to vary between groups separately for women and men and when they were set equal for women and men. If measurement invariance passes the weak and strong invariance test, i.e., there is no statistically significant difference between models when those parameters for women and men are set equal, then we must check for differences in the path analysis part, i.e., regression coefficients (β) among different latent variables in the model between women and men. This is because differences between the groups could occur at the factor (latent variable) level in regression coefficients (β).

Similar to "weak" and "strong" measurement invariance for the measurement part, when testing moderation effect in path analysis, the predictive relationship (regression path) between two variables is tested for the two groups (e.g., women and men) simultaneously. If the predictive relationship is different for the groups [i.e., the values of the regression coefficients (β) are not the same for the predictive relationship between the two constructs for women and men], then there is a gender moderation effect in the model. If moderation does not show differences by gender in any of these steps (measurement invariance holds and testing for regression coefficients shows that they can be set equal for women and men), we can utilize a gender mediation model (see Fig. 1). In other words, we can interpret our model the same way for both men and women, and any gender differences can be modeled using a separate gender variable.

In our multi-group SEM model, we found a nonsignificant p value in each step, and thus measurement invariance holds and the regression coefficients for women and men can be set equal, i.e., there are no moderation effects by gender (for men and women) in our models. Thus, we concluded that our SEM model can be interpreted similarly for men and women and we can use gender mediation analysis (instead of doing moderation by gender). Therefore, we tested the theoretical model in mediation analysis, using gender as a variable (1 for male and 0 for female) directly predicting items to examine the resulting structural paths between constructs (a schematic representation of the path analysis for the gender mediation model is shown in Fig. 1). In the mediation analysis, if there are paths from gender to any of the constructs as we found in our results (Fig. 2) discussed in the next section, it implies that women and men did not have the same average value for those constructs controlling for all constructs to the left. However, it is important to note that all of the item factor loadings and regression coefficients between the constructs are the same for women and men (as found from the gender moderation analysis which preceded the mediation analysis).

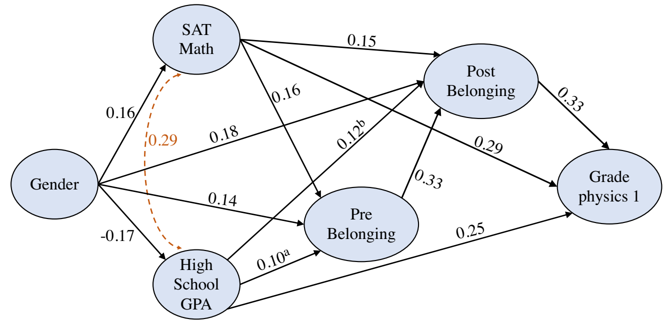

FIG. 2. Result of the path analysis part of the SEM between gender and physics 1 grade through students' sense of belonging. The line thickness indicates the relative magnitude of β values. The dashed line indicates covariances between constructs. The gender variable was coded as 1 for men and 0 for women, so paths from gender with $\beta > 0$ indicate a higher mean for men while $\beta < 0$ indicate a higher mean for women in the predicted variable. All p values are indicated by no superscript for $p < 0.001$, superscript "a" for $p = 0.014$, and superscript "b" for $p = 0.001$. Gender does not directly predict the physics grade.

III. RESULTS

A. Gender differences in predictors and outcomes

We find that women had statistically significantly lower mean values in their SAT math scores, physics 1 grade, prebelonging, and postbelonging (see Table II) while men had statistically significantly lower mean values in high school GPA (see Table II). In addition, the gender gap in men and women's sense of belonging increased from the beginning (pre: Cohen's $d = 0.29$) to the end (post: Cohen's $d = 0.44$) of the course. We also investigated the difference between men's and women's sense of belonging from the pretest to the post-test. We found that women's sense of belonging was not statistically significantly different from the beginning of the class to the end of

TABLE II. Mean predictor and outcome values by gender as well as statistical significance (p values) and effect sizes (Cohen's d) by gender.

Predictors and outcomes (score range)	Mean			
	Male	Female	p value	Cohen's d
High school GPA (0–5)	3.96	4.13	<0.001	−0.34
SAT math (200–800)	682	656	<0.001	0.38
Prebelonging (1–5)	3.52	3.28	<0.001	0.29
Postbelonging (1–5)	3.74	3.34	<0.001	0.44
Physics 1 Grade (0–4)	3.24	3.06	0.003	0.22

the class (Cohen's $d = 0.06$, $p = 0.340$) while men's increase in their sense of belonging was statistically significant (Cohen's $d = 0.27$, $p = 0.001$). We note that Cohen's d indicates a small to medium effect size. However, small and medium effect sizes can still be important for instructors and researchers to focus on considering this is in a physics course in which women make up two-thirds of the class and women had higher high school GPAs than men. While the mean values and effect sizes (Cohen's d) indicate that women have a lower sense of belonging than men, Table II cannot be used to make inferences about the relationships between various factors in the table. Therefore, we used SEM in the next section to show the relationships between the constructs.

Additionally, in the Appendix, we provide the percentage of men and women who selected each response to the items. This provides a sense of how students shifted their answers from the beginning to the end. From Table III in the Appendix, we can see that in general, the percentage of women who selected 1 or 5 increased from pre to post and the percentage of men who selected 5 increased from pre to post, while the percentage of men who selected 3 decreased. In addition, more women than men selected choices 1 (not at all true) and 2 (a little true), while men were more likely to select the answers 4 (mostly true) and 5 (completely true).

B. SEM path model

We used SEM to investigate the relationships between the constructs and to unpack the sense of belonging contribution in explaining the physics 1 grade of women and men. We initially conducted gender moderation analysis between variables using multigroup SEM to investigate if any of the relationships between the constructs were different across gender. There were no group differences at the levels of weak and strong measurement invariance and the level of regression coefficients. Thus, there was no moderation effect by gender. Therefore, we proceeded to gender mediation to investigate the extent to which gender differences in physics 1 grade at the end of the course were mediated by differences in students' precollege academic measures (high school GPA and SAT Math) and sense of belonging in physics.

The result of the path analysis part of the SEM is presented visually in Fig. 2. The model fit indices indicate a good fit to the data: $\mathrm{CFI} = 0.934$ (>0.90), $\mathrm{TLI} = 0.903$ (0.90), $\mathrm{RMSEA} = 0.078$ (<0.08), and $\mathrm{SRMR} = 0.033$ (<0.08). Gender had direct connections to SAT math ($\beta = 0.16$), high school GPA ($\beta = -0.17$), prebelonging ($\beta = 0.14$) and postbelonging ($\beta = 0.18$). However, gender did not have any direct connections to the grade in physics 1. Instead, gender was mediated through students' high school factors and their sense of belonging in the course. To expand further, the statistically significant path from gender to pre belonging means that men are predicted to have higher mean values in their sense of belonging than women controlling for their SAT math scores and high school GPA. The reason that there is no direct path from gender to students' grades in physics 1 is that the gender differences in students' grades (Table II) are statistically nonsignificant when controlling for the other constructs in the model (SAT math, high school GPA, pre- and postbelonging).

Students' grade in physics 1 had direct connections to their high school factors of high school GPA ($\beta = 0.25$) and SAT math ($\beta = 0.29$) as well as postsense of belonging in physics ($\beta = 0.33$). To clarify, a standardized regression coefficient of $\beta = 0.33$ between postsense of belonging and physics grade in Fig. 2 implies that for one standard deviation increase in the postsense of belonging, the grade would show 0.33 standard deviation change controlling for gender, SAT scores, high school GPA, and prebelonging at the beginning of the course. Women also have lower scores in sense of belonging than men at the start and end of the physics course. Since the students' sense of belonging can predict students' grade at the end of the course, it is important for instructors to improve students' sense of belonging in the course by making their classes equitable and inclusive.

IV. SUMMARY AND DISCUSSION

In this research involving both descriptive and inferential quantitative analyses, we find gender gaps in physics grades and sense of belonging (Table II) disadvantaging women in the mandatory introductory physics 1 course for bioscience majors in which women are not outnumbered by men. This is supported by a statistically significant value of Cohen's d between women and men for both sense of belonging and grade with small to medium effect sizes (Table II). Our SEM model shows that students' sense of belonging plays an important role in predicting students' grades at the end of the physics course. Gender also directly predicts students' SAT math scores, high school GPA, prebelonging, and postbelonging. In addition, while men's sense of belonging statistically significantly increases from the beginning to the end of the course, women's sense of belonging does not statistically significantly increase by the end of the course despite women outnumbering men in the course. Since students' sense of belonging can have important implications not only for students' grades in the class but also for their retention in their major and future career trajectories, instructors teaching these physics courses must create an equitable and inclusive learning environment and increase women's sense of belonging.

Instructors have a central role to play in increasing students' sense of belonging in the introductory physics courses. In order to increase students' sense of belonging particularly for students who are underrepresented, some recommendations researchers have made for physics instructors are to send messages that concerns about belonging are normal and fade with time, identify and temper cues that perpetuate the "geeky" scientist stereotype, and openly endorse effort and hard work over brilliance [57]. One study showed that teacher practices, including their encouragement of cooperative activities in an inclusive environment, were related to students' engagement in the course and a high sense of belonging in the classroom [58]. Instructors could provide time for students to work in groups during class or recitations in a supportive environment, making sure that all student voices are heard equally while discussing problems. These types of pedagogical approaches could in turn help increase students' sense of belonging in the classroom if they observe that their ideas are respected by other students in the classroom. This is especially important when implementing active learning pedagogies in the classroom. If not implemented using teaching strategies that are equitable and inclusive, men have been shown to dominate responding to questions from instructors in STEM classes and women have reported lower scientific self-efficacy [28].

In addition, values-affirmation interventions that typically have students write about core personal values can have positive effects on students in academic settings [59]. The values-affirmation interventions may be more beneficial for students from underrepresented groups [60]. Moreover, belonging interventions that focus on improving the sense of belonging of underrepresented students in STEM courses have been found effective. They have been found to create an inclusive learning environment that can mitigate their doubts about social belonging in college and improve the grades of racial and ethnic minority students [61] and women [62,63] in STEM fields. Lastly, inclusive mentoring, as well as contact with experts and peers who share similar demographic characteristics in academic fields, may be able to protect people, especially individuals who are underrepresented in the field, from the negative effects of stereotype threat and has the potential to increase their sense of belonging [64].

ACKNOWLEDGMENTS

This work was supported by the University of Pittsburgh Grant No. NSF DUE-152457. We thank all students who participated in this research and Dr. Robert Devaty for his constructive feedback on this manuscript.

APPENDIX: PERCENTAGES OF MALE AND FEMALE STUDENTS WHO SELECTED EACH CHOICE FOR EACH SURVEY ITEM

Below, we provide the percentages of men and women who selected each answer choice for each sense of belonging item in the pre and post survey in physics 1. This distribution provides a sense of how students shifted their answers from pre to post survey. From Table III, we can see that in general, the percentage of women who selected 1 or 5 increased from pre to post, and the percentage of men who selected 5 increased from pre to post while the percentage of men who selected 3 decreased. In addition, more women than men selected the answer choices 1 (not at all true) or 2 (a little true), while men were more likely to select the answer choices 4 (mostly true) and 5 (completely true).

TABLE III. Percentages of 521 women and 293 in physics 1 who responded to each sense of belonging item by the options they selected with 1 being the low value (not at all true) and 5 being the high value (completely true). The rating scale for the sense of belonging items was: not at all true, a little true, somewhat true, mostly true, and completely true.

	Women's sense of belonging distribution across choices									
	pre test					post test				
Question	1	2	3	4	5	1	2	3	4	5
1	5%	18%	45%	25%	7%	11%	24%	33%	23%	9%
2	3%	12%	34%	33%	18%	4%	9%	18%	34%	35%
3	3%	18%	44%	31%	4%	13%	24%	31%	23%	9%
4	6%	13%	32%	33%	16%	6%	12%	24%	28%	30%
	Men's sense of belonging distribution across choices									
1	3%	15%	41%	28%	13%	3%	12%	33%	36%	16%
2	3%	12%	31%	27%	27%	2%	4%	18%	29%	47%
3	1%	9%	48%	33%	9%	5%	16%	31%	35%	13%
4	2%	10%	26%	39%	23%	3%	6%	18%	31%	42%

[1] S. J. Correll, Gender and the career choice process: The role of biased self-assessments, Am. J. Soc. **106**, 1691 (2001).

[2] S. J. Correll, Constraints into preferences: Gender, status, and emerging career aspirations, Am. Soc. Rev. **69**, 93 (2004).

[3] J. S. Eccles, Understanding women's educational, and occupational choices: Applying the Eccles, *et al.* model of achievement-related choices, Psychol. Women Q. **18**, 585 (1994).

[4] M.-T. Wang and J. Degol, Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields, Dev. Policy Rev. **33**, 304 (2013).

[5] S. L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, J. Res. Sci. Teach. **45**, 955 (2008).

[6] C. P. Benbow and L. L. Minor, Mathematically talented males and females and achievement in the high school sciences, Am. Educ. Res. J. **23**, 425 (1986).

[7] J. E. Stets, P. S. Brenner, P. J. Burke, and R. T. Serpe, The science identity and entering a science occupation, Soc. Sci. Res. **64**, 1 (2017).

[8] M. M. Chemers, E. L. Zurbriggen, M. Syed, B. K. Goza, and S. Bearman, The role of efficacy and identity in science career commitment among underrepresented minority students, J. Soc. Issues **67**, 469 (2011).

[9] K. A. Robinson, T. Perez, J. H. Carmel, and L. Linnenbrink-Garcia, Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit, Contemp. Educ. Psychol. **56**, 180 (2019).

[10] E. M. Marshman, Z. Y. Kalender, C. Schunn, T. Nokes-Malach, and C. Singh, A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences, Can. J. Phys. **96**, 391 (2018).

[11] R. W. Roeser, C. Midgley, and T. C. Urdan, Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging, J. Educ. Psychol **88**, 408 (1996).

[12] T. L. Strayhorn, *College Students' Sense of Belonging: A Key to Educational Success for All Students* (Routledge, London, 2012).

[13] L. R. Hausmann, J. W. Schofield, and R. L. Woods, Sense of belonging as a predictor of intentions to persist among African American and White first-year college students, Res. High. Educ. **48**, 803 (2007).

[14] L. D. Pittman and A. Richmond, Academic and psychological functioning in late adolescence: The importance of school belonging, J. Exp. Educ. **75**, 270 (2007).

[15] D. Wilson, D. Jones, F. Bocell, J. Crawford, M. J. Kim, N. Veilleux, T. Floyd-Smith, R. Bates, and M. Plett, Belonging and academic engagement among undergraduate STEM students: A multi-institutional study, Res. High. Educ. **56**, 750 (2015).

[16] G. Trujillo and K. D. Tanner, Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity, CBE Life Sci. Educ. **13**, 6 (2014).

[17] M. Ong, J. M. Smith, and L. T. Ko, Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success, J. Res. Sci. Teach. **55**, 206 (2018).

[18] K. L. Lewis, J. G. Stout, N. D. Finkelstein, S. J. Pollock, A. Miyake, G. L. Cohen, and T. A. Ito, Fitting in to move forward: Belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM), Psychol. Women Q. **41**, 420 (2017).

[19] B. London, L. Rosenthal, S. R. Levy, and M. Lobel, The influences of perceived identity compatibility and social support on women in nontraditional fields during the college transition, Basic Appl. Soc. Psychol. **33**, 304 (2011).

[20] C. Good, A. Rattan, and C. S. Dweck, Why do women opt out? Sense of belonging and women's representation in mathematics, J. Pers. Soc. Psychol. **102**, 700 (2012).

[21] J. G. Stout, T. A. Ito, N. D. Finkelstein, and S. J. Pollock, How a gender gap in belonging contributes to the gender gap in physics participation, AIP Conf. Proc. **1513**, 402 (2013).

[22] S.-J. Leslie, A. Cimpian, M. Meyer, and E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines, Science **347**, 262 (2015).

[23] S. Upson and L. F. Friedman, Where are all the female geniuses?, Sci. Am. Mind **23**, 63 (2012).

[24] L. Bian, S.-J. Leslie, and A. Cimpian, Gender stereotypes about intellectual ability emerge early and influence children's interests, Science **355**, 389 (2017).

[25] L. Archer, J. Moote, B. Francis, J. DeWitt, and L. Yeomans, The "exceptional" physics girl: A sociological analysis of multimethod data from young women aged 10–16 to explore gendered patterns of post-16 participation, Am. Educ. Res. J. **54**, 88 (2017).

[26] J. L. Smith, K. L. Lewis, L. Hawthorne, and S. D. Hodges, When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns, Pers. Soc. Psychol. Bull. **39**, 131 (2013).

[27] A. Ladewig, M. Keller, and U. Klusmann, Sense of belonging as an important factor in the pursuit of physics: Does it also matter for female participants of the German Physics Olympiad?, Front. Psychol. **11**, 548781 (2020).

[28] S. M. Aguillon, G.-F. Siegmund, R. H. Petipas, A. G. Drake, S. Cotner, and C. J. Ballen, Gender differences in student participation in an active-learning classroom, CBE Life Sci. Educ. **19**, 1 (2020).

[29] R. Masika and J. Jones, Building student belonging and engagement: Insights into higher education students' experiences of participating and learning together, Teaching Higher Educ. **21**, 138 (2016).

[30] C. Goodenow, Classroom belonging among early adolescent students: Relationships to motivation and achievement, J. Early Adolesc. **13**, 21 (1993).

[31] S. Zumbrunn, C. McKim, E. Buhs, and L. R. Hawley, Support, belonging, motivation, and engagement in the college classroom: A mixed method study, Instr. Sci. **42**, 661 (2014).

[32] Z. Hazari, D. Chari, G. Potvin, and E. Brewe, The context dependence of physics identity: Examining the role of performance/competence, recognition, interest, and sense

of belonging for lower and upper female physics undergraduates, J. Res. Sci. Teach. **57**, 1583 (2020).
[33] R. H. Tai and P. M. Sadler, Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA, Int. J. Sci. Educ. **23**, 1017 (2001).
[34] R. Henderson, G. Stewart, J. Stewart, L. Michaluk, and A. Traxler, Exploring the gender gap in the conceptual survey of electricity and magnetism, Phys. Rev. Phys. Educ. Res. **13**, 020114 (2017).
[35] A. T. Wee, B. E. Baaquie, and A. C. Huan, Gender differences in undergraduate physics examination performance and learning strategies in Singapore, Phys. Educ. **28**, 158 (1993).
[36] N. I. Karim, A. Maries, and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics?, Eur. J. Phys. **39**, 025701 (2018).
[37] Z. Hazari, R. H. Tai, and P. M. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, Sci. Educ. **91**, 847 (2007).
[38] L. E. Kost, S. J. Pollock, and N. D. Finkelstein, Characterizing the gender gap in introductory physics, Phys. Rev. Phys. Educ. Res. **5**, 010101 (2009).
[39] A. M. L. Cavallo, W. H. Potter, and M. Rozman, Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors, School Sci. Math. **104**, 288 (2004).
[40] R. M. Felder, G. N. Felder, M. Mauney, C. E. Hamrin Jr, and E. J. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Eng. Educ. **84**, 151 (1995).
[41] A. C. Strenta, R. Elliott, R. Adair, M. Matier, and J. Scott, Choosing and leaving science in highly selective institutions, Res. High. Educ. **35**, 513 (1994).
[42] P. M. Sadler and R. H. Tai, Success in introductory college physics: The role of high school preparation, Sci. Educ. **85**, 111 (2001).
[43] A. Madsen, S. B. McKagan, and E. C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, Phys. Rev. Phys. Educ. Res. **9**, 020121 (2013).
[44] M. Lorenzo, C. H. Crouch, and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).
[45] G. E. Hart and P. D. Cottle, Academic backgrounds and achievement in college physics, Phys. Teach. **31**, 470 (1993).
[46] M. G. Jones, A. Howe, and M. J. Rua, Gender differences in students' experiences, interests, and attitudes toward science and scientists, Sci. Educ. **84**, 180 (2000).
[47] S. L. Eddy and S. E. Brownell, Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines, Phys. Rev. Phys. Educ. Res. **12**, 020106 (2016).
[48] Y. Li and C. Singh, Effect of gender, self-efficacy, and interest on perception of the learning environment and outcomes in calculus-based introductory physics courses, Phys. Rev. Phys. Educ. Res. **17**, 010143 (2021).
[49] S. M. Glynn, P. Brickman, N. Armstrong, and G. Taasoobshirazi, Science motivation questionnaire II: Validation with science majors and nonscience majors, J. Res. Sci. Teach. **48**, 1159 (2011).
[50] PERTS Academic Mindsets Assessment, https://www.perts.net/orientation/ascend.
[51] R. Likert, A technique for the measurement of attitudes, Arch. Psychol. **22**, 55 (1932), https://www.worldcat.org/title/technique-for-the-measurement-of-attitudes/oclc/812060.
[52] L. J. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).
[53] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Routledge, London, 2013).
[54] R. C. Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2013).
[55] R. C. MacCallum, M. W. Browne, and H. M. Sugawara, Power analysis and determination of sample size for covariance structure modeling, Psychological Methods **1**, 130 (1996).
[56] A. J. Tomarken and N. G. Waller, Structural equation modeling: Strengths, limitations, and misconceptions, Annu. Rev. Clin. Psychol. **1**, 31 (2005).
[57] K. L. Lewis, J. G. Stout, S. J. Pollock, N. D. Finkelstein, and T. A. Ito, Fitting in or opting out: A review of key social-psychological factors influencing a sense of belonging for women in physics, Phys. Rev. Phys. Educ. Res. **12**, 020110 (2016).
[58] D. Solomon, V. Battistich, D.-i. Kim, and M. Watson, Teacher practices associated with students' sense of the classroom as a community, Soc. Psychol. Educ. **1**, 235 (1996).
[59] G. L. Cohen and D. K. Sherman, The psychology of change: Self-affirmation and social psychological intervention, Ann. Rev. Psychol. **65**, 333 (2014).
[60] G. L. Cohen, J. Garcia, N. Apfel, and A. Master, Reducing the racial achievement gap: A social-psychological intervention, Science **313**, 1307 (2006).
[61] G. M. Walton and G. L. Cohen, A question of belonging: Race, social fit, and achievement, J. Pers. Soc. Psychol. **92**, 82 (2007).
[62] G. M. Walton, C. Logel, J. M. Peach, S. J. Spencer, and M. P. Zanna, Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering, J. Educ. Psychol **107**, 468 (2015).
[63] K. R. Binning, N. Kaufmann, E. M. McGreevy, O. Fotuhi, S. Chen, E. Marshman, Z. Y. Kalender, L. Limeri, L. Betancur, and C. Singh, Changing social contexts to foster equity in college science courses: An ecological-belonging intervention, Psychol. Sci. **31**, 1059 (2020).
[64] N. Dasgupta, Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model, Psychological inquiry **22**, 231 (2011).