

Effect of gender, self-efficacy, and interest on perception of the learning environment and outcomes in calculus-based introductory physics courses

Yangqiuting Li¹ and Chandralekha Singh¹

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA



(Received 17 February 2021; accepted 10 May 2021; published 23 June 2021)

Students' motivational beliefs about physics can influence their learning outcomes as well as retention in their majors and career choices. Moreover, due to societal stereotypes and biases about who belongs in physics and can succeed in physics, women often have lower motivational beliefs about physics than men. The expectancy-value theory emphasizes the importance of self-efficacy and value in predicting students' short- and long-term academic and professional outcomes, but there are few studies focusing on how the learning environment shapes these motivational beliefs of women and men. Investigating how the perception of learning environment in introductory physics courses for the engineering, physical science, and mathematics majors in their first year of college predicts the motivational beliefs of women and men can be useful in making the learning environments equitable and inclusive so that the underrepresented students, e.g., women, are not disadvantaged. In this study, we adapt prior identity framework to investigate how the learning environment (including sense of belonging, perceived peer interaction, and perceived recognition) predicts students' physics self-efficacy, interest, and identity by controlling for their self-efficacy and interest at the beginning of a calculus-based introductory physics course. We surveyed 1203 students, 35% of whom identified as women. We found signatures of inequitable and noninclusive learning environment in that not only were female students' physics self-efficacy and interest lower than male students' at the beginning of the course, but the gender gaps in these motivational constructs became even larger by the end of the course. Analysis revealed that the decrease in students' physics self-efficacy and interest were mediated by the learning environment and predicted students' physics identity. We find that the perceived recognition played a major role in predicting students' physics identity, and students' sense of belonging in physics played an important role in explaining the change in students' physics self-efficacy.

DOI: [10.1103/PhysRevPhysEducRes.17.010143](https://doi.org/10.1103/PhysRevPhysEducRes.17.010143)

I. INTRODUCTION AND THEORETICAL FRAMEWORK

Several prior studies have focused on underrepresented groups such as women in science, technology, engineering, and mathematics (STEM) courses, majors and careers [1–20]. Prior research suggests that individuals' career enrollment and achievement in STEM can be influenced by their domain specific motivational beliefs such as self-efficacy, interest, and identity [1–3,11,21–36]. For students from underrepresented groups, these motivational characteristics might be undermined due to lack of encouragement, negative stereotypes, and inadequate prior preparation, leading to withdrawal from STEM fields [37–47]. Hence, investigating students' motivational characteristics is critical to understanding and addressing diversity, equity, and inclusion in STEM disciplines.

Prior research suggests that self-efficacy is an important motivational characteristic of students in order for them to excel in a domain [4,11,13,32]. In particular, self-efficacy is the belief in one's capability to be successful in a particular task, course, or subject area [48,49], and it has been shown to influence students' engagement and performance in a given domain [27,28,30,50,51]. Students with high self-efficacy in a domain often enroll in more difficult courses in that domain than those with low self-efficacy because they perceive difficult tasks as challenges rather than threats [29].

Another motivational characteristic is interest, which is defined by positive emotions accompanied by curiosity and engagement in a particular discipline [52,53]. Studies have shown that interest can also influence students' learning [26,27,53–55]. For example, one study shows that making science courses more relevant to students' lives and transforming curricula to promote interest in learning can improve students' achievement [56]. In addition, studies have shown that students' interest is not independent from self-efficacy [27,57].

According to Eccles' expectancy-value theory (EVT) [57,58], interest is paired well with self-efficacy as connected constructs that predict students' academic outcome

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International* license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

expectations and career aspirations. In this theory, students' persistence and engagement in a task or field can be influenced by their expectancy of success and by their estimation of the task's value. The expectancy here refers to learners' belief in their ability to succeed in the given task [57], which is closely related to self-efficacy. Value in this theory includes four components: intrinsic value, attainment value, utility value, and cost [57]. Intrinsic value represents students' personal interest in the task or field. Attainment value refers to how important students themselves feel it is for them to develop mastery and do a good job in the field [57]. Utility value pertains to whether this task can help them succeed in various fields [57]. The last value component is cost, which corresponds to the negative aspect of engagement such as the amount of anxiety or opportunity cost due to the time spent on the task [57]. In the expectancy-value theory, people's learning goal, academic performance, and persistence in the field are impacted by their expectancy and the four components of value [57].

Science identity is another important motivational characteristic that influences students' career decisions and outcome expectations [1–3,59–63]. Students' identity in a specific field such as physics is related to whether they see themselves as a physics person [1–3,59,60,63]. Some studies have found that female students are less likely to see themselves as a physics person than male students [2,64]. In prior research, in general, the challenges women face in developing physics identity are related to societal biases and stereotypes [65–67]. For example, negative societal stereotypes and generalizations about who can succeed in physics and other STEM disciplines can lower women's sense of belonging and undermine their experiences, so that they often have lower STEM identity than men [37,68,69]. Thus, studying STEM identity may help us to understand the gender difference in participation in STEM. The well-known science identity framework developed by Carlone and Johnson [1] includes three dimensions: competence ("I think I can"), performance ("I am able to do"), and recognition ("I am recognized by others"). Hazari *et al.* adapted this model to physics and added interest to this model [3]. They investigated whether the relation between gender and physics identity was mediated by interest, competency belief, and perceived recognition from other people [70,71]. These two studies reveal that individuals' internal identity in science is not only impacted by their own motivational characteristics but also by their perceived recognition from others.

Several studies have shown that female students did not feel that they were recognized appropriately even before they entered college [65,72,73]. We have conducted interviews with women in physics courses and it is clear from these interviews that they felt that in general men were recognized more by the instructors than women and it impacted their self-efficacy, interest, and identity as a

physics person [74,75]. One of the stereotypical views of science is that it is for high achievers or naturally gifted students [65]. In general, due to societal stereotypes, being a genius or exceptionally smart is associated with boys [76]. In one investigation, boys and girls were externally exposed to these fixed intelligence views starting from early ages, which influenced the development of their science identity [73]. In addition, one study indicated that elementary and high school boys and girls interested in science were treated differently by parents, teachers, and friends. While boys received admiration and encouragement for their interests, responses to girls were often characterized by ambivalence, lack of encouragement, or suggestions that their goals were inappropriate [77]. Studies showed that these stereotypes and biases also exist in university context [78,79]. For example, one study found that science faculty members in biological and physical sciences exhibit biases against female students by rating male students significantly more competent [78]. Our prior study also found that in introductory physics course, there is a significant gender difference in perceived recognition from TAs and instructors [79,80]. There is often a feedback loop, e.g., between recognition and identity. For example, the experiences of not being recognized as a science person may further increase the stereotype threat, and these gender-based biases may accumulate over time and become a detriment to female students' science identity.

Moreover, students' interest and self-efficacy have also been found to be connected to their interaction with other people and recognition by them [49,53]. In Hidi and Renninger's four stages model of interest development [53,54,81], people's interest in a field is triggered and maintained by external factors first and then becomes an individual interest. In addition, according to Bandura's social cognitive theory, individual's self-efficacy can be shaped by verbal encouragement from others [82,83]. Kalender *et al.*'s physics identity framework [79] showed that students' perceived recognition not only strongly predicts their physics identity, but also predicts their physics interest and self-efficacy, which suggests that the gender difference in students' perceived classroom experience may partially explain the gender differences in students' physics self-efficacy, interest, and identity at the end of the course. However, it is not clear how much of self-efficacy and interest changed from the beginning to the end of the course, and what role was played by perception of learning environment in this change.

In addition to perceived recognition, some studies have shown that students' sense of belonging and their interaction with peers are also important constructs of learning environment [7,68,69,84–87]. For example, if students feel more secure in their belonging in school, they may approach others in the academic environment more and with more positive attitudes, building better interaction and higher perceived recognition [88]. However, there are very

few quantitative studies about the effect of learning environment on students' motivational beliefs and what role is played by each factor in the learning environment. To better understand how the learning environment influences students' motivational outcomes and how to develop an inclusive and equitable learning environment, further study is needed.

Our conceptualization of inclusive excellence and equity in learning includes three pillars: equitable access and opportunity to learn, equitable and inclusive learning environment, and equitable outcomes. Thus, by inclusive excellence and equity in learning, we mean that not only should all students have equitable opportunities and access to resources, they should also have an equitable and inclusive learning environment with appropriate support and mentoring so that they can engage in learning in a meaningful and enjoyable manner and the learning outcomes should be equitable. By equitable learning outcomes, we mean that students from all demographic groups (e.g., regardless of their gender identity or race or ethnicity) who have the pre-requisites to enroll in courses have comparable learning outcomes. This conceptualization of equitable outcome is consistent with Rodriguez *et al.*'s equity of parity model [89]. The learning outcomes include student performance in courses as well as evolution in their motivational beliefs such as self-efficacy, etc., because regardless of performance, students' motivational beliefs can influence their short- and long-term retention in their major and careers. In other words, an equitable and inclusive learning environment should provide guidance, support, and mentoring to all students as appropriate and ensure that students from all demographic groups have equal sense of belonging regardless of their prior preparation so long as they have the prerequisite basic knowledge and skills. An equitable and inclusive learning environment

would also ensure that students from all demographic groups and prior preparation embrace challenges as learning opportunity instead of being threatened by them and enjoy learning. Equitable learning outcomes also include the ability of the courses to empower students from all demographic groups and make them passionate to pursue further learning and careers in related areas. We note that equitable access and opportunity to learn, equitable and inclusive learning environment, and equitable outcomes are strongly entangled with each other. For example, if the learning environment is not equitable and inclusive in a particular course, the learning outcomes are unlikely to be equitable.

To improve equity and inclusion in physics classes, we conducted a study to investigate the effect of the perception of learning environment (including sense of belonging, perceived peer interaction, and perceived recognition) on students' self-efficacy, interest and identity by controlling for students' gender and their self-efficacy and interest at the beginning of a calculus-based introductory physics course. We note that the learning environment mentioned here is not necessarily the classroom environment: it also includes students' experiences outside the class. For example, students may work together on their homework after class, and they could also ask for help during instructors' office hours. As shown in Fig. 1, the total nine constructs are divided into three groups: what we control for, perception of learning environment, and outcomes. Students' gender, pre-self-efficacy (Pre SE) and pre-interest measured at the beginning of the course are constructs that we control for, which are related to students' beliefs about physics based on their prior experience. Outcomes include students' physics post-self-efficacy (Post SE), post-interest, and identity at the end of the course. Perceived recognition (Perceived Recog), perceived

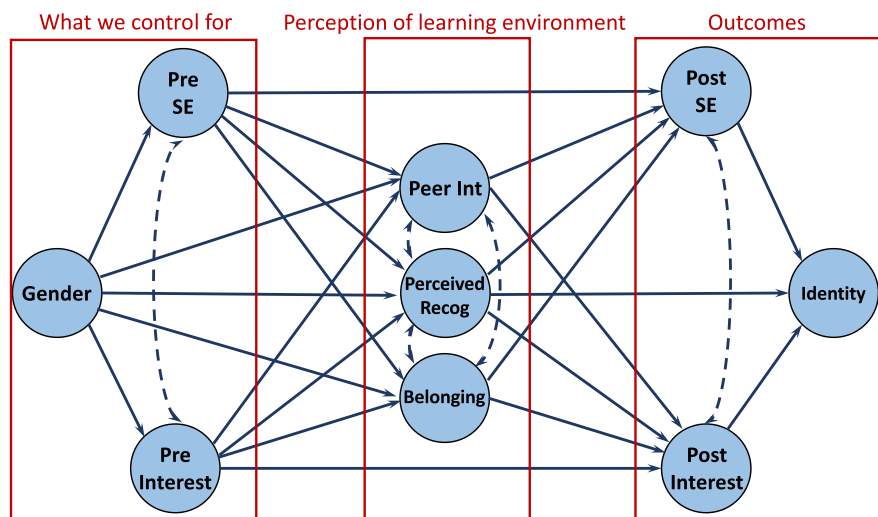


FIG. 1. Schematic representation of the theoretical framework regarding physics identity. From left to right, all possible regression paths were considered, but only some of the paths are shown here.

peer interaction (Peer Int) and sense of belonging (Belonging) constitute the perception of learning environment. For convenience, perceived peer interaction is shortened to peer interaction in the rest of the paper. We note that out of the three components of the pre-physics identity, pre-self-efficacy, and pre-interest have been controlled for. However, perceived recognition is part of the perception of learning environment and thus is not controlled for at the beginning of the course (i.e., students have no experience interacting with their instructors and TAs).

In our study, peer interaction, perceived recognition, and sense of belonging were measured at the end of the course. This is because only after the course can students answer these survey questions based on their real experience in the course such as their interaction with peers, TAs, and instructors. Because pre- and post-responses are actually students' responses to the same questions at two different time points, it is not surprising if students' pre-self-efficacy and pre-interest partially predict their post-self-efficacy and postinterest. However, if students' self-efficacy and interest changed from pre to post, we want to study whether the perception of learning environment helps to explain the changes and what role is played by each construct in the perception of learning environment.

In this study, we first calculated female and male students' mean scores for each motivational construct. Then we studied how much of students' self-efficacy and interest changed from pre- to post- and how much of these changes can be explained by the perception of learning environment. We used structural equation modeling (SEM) to study the effect of learning environment on students' post-self-efficacy, interest, and identity by controlling for gender and their pre-self-efficacy and interest. To better understand the role played by each perception of learning environment construct, we first considered a model with perceived recognition as the only perception of learning environment construct to see how much variance in students' self-efficacy, interest and identity was explained by the model. Then, we added peer interaction and sense of belonging into this model one by one to see whether the adding constructs help to explain extra variance in students' motivational outcomes.

II. RESEARCH QUESTIONS

Our research questions regarding the calculus-based introductory physics course 1 at a large state-related university are as follows. This course is generally taken by engineering, physical science, and mathematics majors in the first year of undergraduate studies.

RQ1. Are there gender differences in students' motivational characteristics and do they change from the beginning to the end of the course (i.e., from pre to post)?

RQ2. How do the components of the perception of learning environment (including sense of belonging, peer interaction, and perceived recognition) predict students'

physics identity as well as post-self-efficacy and post-interest controlling their gender, pre-self-efficacy, and pre-interest?

RQ3. Does gender moderate the relationship between any two constructs in the model? (i.e., Does the strength of relationship given by the standardized regression coefficient between any two constructs in the model differ for women and men?)

RQ4. If gender does not moderate any path in the model, how does gender mediate

- the factors that were controlled for?
- the perception of learning environment after controlling for pre-self-efficacy and pre-interest?
- the motivational outcomes after controlling for everything in the model?

RQ5. What unique role is played by each of the three components we have included in the learning environment in predicting students' motivational outcomes?

RQ6. Based on the aspects of the perception of learning environment that explain most of the variance in the outcomes, which model is most productive for creating the inclusive environment?

III. METHODOLOGY

A. Participants

In this study, we collected motivational survey data at the beginning and end of the semester from students who took the introductory calculus-based physics 1 course in two consecutive fall semesters. This course is taken mostly by students majoring in engineering, physical science, and mathematics. This course is a traditional lecture-based course (4 h per week) with recitations (1 h per week) in which students typically work collaboratively on physics problems. The paper surveys were handed out and collected by TAs in the first and last recitation class of a semester. Course instructors were encouraged to give students course credit or extra credit for completing the survey, and the completion rates are typically 80%–90%. We named the data collected at the beginning of the semester as predata and that collected at the end of the semester as postdata. Finally, we combined the two semesters' data and put them into two categories, pre and post. The demographic data of students—such as gender—were provided by the university. Students' names and IDs were de-identified by an honest broker who provided each student with a unique new ID (which connected students' survey responses with their demographic information). Thus, researchers could analyze students' data without having access to students' identifying information.

There were 1364 students participating in the presurvey and 1219 students participating in the postsurvey. There were 1052 students participating both pre- and postsurvey. Students may miss the first or the last recitation class for many possible reasons. In this study, we only considered

students who are in our postdataset even though some of them might not take the presurvey. This is because our focus is on students' motivational outcomes at the end of the course, and what role is played by the perception of learning environment in predicting students' motivational outcomes by controlling for their premotivational scores. In addition, SEM can provide a better estimation with the more completed outcome data [90]. Thus, it is reasonable to keep students' response to the postsurvey rather than the presurvey. To handle the missing data, we used the full information maximum likelihood method to estimate the model by using all available information [91]. In this method, the population parameters are estimated that would most likely produce the estimates from the sample data that is analyzed [91].

In all, there were 1219 students participating in the postsurvey including both semesters. In our final dataset, we kept 1203 students (including 427 female students and 776 male students) because the other 16 students did not provide their gender information. We recognize that gender identity is not a binary construct. However, because students' gender information was collected by the

university which offered binary options, we did the analysis with the binary gender data in this study. 1.3% of the students who did not provide this information were not included in this analysis.

B. Survey instruments

In this study, our analysis includes six motivational constructs—physics self-efficacy, interest, peer interaction, perceived recognition, sense of belonging, and identity. The questions for each construct are listed in Table I. The survey questions were adapted from the existing motivational research [92–98] and were revalidated in our prior work [11,99–102]. The validation and refinement of the survey involved use of one-on-one student interviews with both introductory and advanced students [11,74,102,103], exploratory and confirmatory factor analyses (EFA and CFA) [104], Pearson correlation between different constructs, and Cronbach's alpha [105,106]. The peer interaction scale was added later, and therefore additional validation (CFA and Cronbach's alpha results) for that scale is reported in Table I.

TABLE I. Survey items for each of the motivational scales. The Cronbach's alpha and CFA item loadings (Lambda and p values of the significance test for each item loading) shown here were calculated with the postdata. † The response options for this question are "Never, once a month, once a week, every day". ‡ The response options for this question are "Very boring, boring, interesting, very interesting."

Construct and item	Lambda	p value
<i>Physics identity</i>		
I see myself as physics person.	1.000	< 0.001
<i>Physics self-efficacy</i> (Cronbach's $\alpha = 0.80$)		
I am able to help my classmates with physics in the laboratory or in recitation.	0.722	< 0.001
I understand concepts I have studied in physics.	0.726	< 0.001
If I study, I will do well on a physics test.	0.727	< 0.001
If I encounter a setback in a physics exam, I can overcome it.	0.669	< 0.001
<i>Physics interest</i> (Cronbach's $\alpha = 0.82$)		
I wonder about how physics works [†]	0.664	< 0.001
In general, I find physics [‡]	0.795	< 0.001
I want to know everything I can about physics.	0.796	< 0.001
I am curious about recent physics discoveries.	0.693	< 0.001
<i>Physics perceived recognition</i> (Cronbach's $\alpha = 0.86$)		
My family sees me as physics person.	0.902	< 0.001
My friends see me as physics person.	0.899	< 0.001
My physics TA and/or instructor see me as physics person.	0.693	< 0.001
<i>Physics sense of belonging</i> (Cronbach's $\alpha = 0.86$)		
I feel like I belong in this class.	0.831	< 0.001
I feel like an outsider in this class.	0.697	< 0.001
I feel comfortable in this class.	0.830	< 0.001
I feel like I can be myself in this class.	0.616	< 0.001
Sometimes I worry that I do not belong in this physics class.	0.712	< 0.001
<i>Physics peer interaction</i> (Cronbach's $\alpha = 0.91$)		
<i>My experience and interaction with other students in this class...</i>		
made me feel more relaxed about learning physics.	0.717	< 0.001
increased my confidence in my ability to do physics.	0.910	< 0.001
increased my confidence that I can succeed in physics.	0.928	< 0.001
increased my confidence in my ability to handle difficult physics problems.	0.846	< 0.001

Physics self-efficacy represents students’ belief about whether they can excel in physics. In our survey, we had four items for self-efficacy (Cronbach’s alpha = 0.69 for pre-self-efficacy and Cronbach’s alpha = 0.8 for the post-self-efficacy [105]). These items had response scale of “NO!, no, yes, YES!”, which is a 4-point Likert scale (1–4). We also had four items for physics interest (Cronbach’s alpha = 0.75 for the pre-interest, Cronbach’s alpha = 0.82 for the postinterest). The question “I wonder about how physics works” had temporal response options: “Never, once a month, once a week, every day.” whereas the question “In general, I find physics:” had response options “Very boring, boring, interesting, very interesting”. The remaining two items were answered on the “NO!, no, yes, YES!” scale. By choosing the four options, students will get score from 1 to 4 accordingly. For example, if a student finds physics very boring, they will get one point for this item. And the more interest a student has in physics, the higher score the student will get for this item. It is noteworthy that Cronbach’s alpha is always higher in our postdata than in the predata. This may be because students had a clearer judgment about their self-efficacy and interest after one semester of learning so that they could answer those questions in a more consistent way at the end of the semester than at the beginning of the semester.

There is one item for physics identity in this survey (I see myself a physics person). Physics identity corresponds to students’ belief about whether they designate themselves as a physics person [3]. This item involved a four-point Likert response on the scale: “strongly disagree, disagree, agree, and strongly agree” and they correspond to 1 to 4 points [107].

In addition, perceived recognition, peer interaction, and sense of belonging are the other three motivational constructs in our study. Unlike self-efficacy, interest and identity, these three constructs are directly related to students’ experience in the course. Perceived recognition included three items which represent whether a student thinks other people see them as a physics person [2,3,59] (Cronbach’s alpha = 0.86). Peer interaction including four items represents whether students have a productive and enjoyable experience when working with peers (Cronbach’s alpha = 0.91). Sense of belonging is about

students’ feelings of whether they belonged in the physics class [86], and it included five items that were scored on a 5-point Likert scale: “not at all true, a little true, somewhat true, mostly true and completely true” (Cronbach alpha = 0.86). Two sense of belonging items (“I feel like an outsider in this class” and “Sometimes I worry that I do not belong in this physics class”) were reverse coded, which means that a higher score in these two items represents a lower sense of belonging. Students’ score of each construct is the average score of all items in this construct.

C. Quantitative analysis of survey data

First, we calculated the mean score for each construct for each student. Then we used a regular two-sample *t*-test [108,109] to compare students’ pre and postscore and to compare responses for female and male students. Finally, we used structural equation modeling (SEM) [110] to study the effect of the perception of learning environment on students’ motivational outcomes by controlling for gender and their pre-self-efficacy and interest. The SEM includes two parts: confirmatory factor analysis (CFA) and path analysis.

To validate the items on our survey, we performed the CFA for each construct. The model fit is good if the fit parameters are above certain thresholds. In CFA, comparative fit index (CFI) > 0.9, Tucker-Lewis index (TLI) > 0.9, root mean square error of approximation (RMSEA) < 0.08, and standardized root mean square residual (SRMR) < 0.08 are considered as acceptable and RMSEA < 0.06 and SRMA < 0.06 are considered as a good fit [90]. In our study, CFI = 0.943, TLI = 0.934, RMSEA = 0.05, and SRMR = 0.041 which represents a good fit. Thus, there is additional quantitative support for dividing the constructs as proposed. In addition, as show in Table I, all of the CFA item loadings are above 0.5 and most of them are above 0.7, which means that our constructs extract sufficient variance from the items [111]. The CFA results allowed us to perform the path analysis.

Before performing the path analysis, we calculated the Pearson correlation coefficients pairwise between each two constructs [106]. As shown in Table II, even though these constructs have strong correlations with each other, the

TABLE II. Zeroth order correlation coefficients of the constructs in the mediation model.

Observed variable	1	2	3	4	5	6	7	8
1. Physics identity
2. Pre-self-efficacy	0.55
3. Pre-interest	0.60	0.66
4. Post-self-efficacy	0.68	0.61	0.45
5. Postinterest	0.70	0.49	0.87	0.63
6. Perceived recognition	0.84	0.46	0.59	0.70	0.68
7. Peer interaction	0.51	0.37	0.31	0.67	0.47	0.49
8. Sense of belonging	0.62	0.46	0.38	0.83	0.55	0.63	0.68	...

correlations are not so high that the constructs could not be separately examined in SEM [112]. It is noteworthy that in Table II there are three very strong correlations. The correlation coefficient between the pre-interest and post-interest is 0.87, which means that students' interest at the end of the course is highly related to their interest at the beginning of the course. The correlation coefficient between physics identity and perceived recognition is 0.84, which is consistent with Godwin *et al.* and Kalender *et al.*'s prior work [63,79] finding that perceived recognition is the largest predictor of physics identity. Another large coefficient is between students' post-self-efficacy and sense of belonging. According to prior work done by Kalender *et al.*, these two constructs are indeed strongly correlated with each other [72].

To analyze the relations among the constructs, we performed the full SEM. Apart from CFA, the path analysis in SEM gives regression coefficients β for paths between each pair of constructs and the value of each β is a measure of the strength of that relationship. Compared with a multiple regression model, the advantage of SEM is that we can estimate all of the regression links for multiple outcomes and factor loadings for items simultaneously, which improves the statistical power. The level of SEM model fit can also be represented by CFI, TLI, RMSEA, and SRMR. We first analyzed the saturated SEM model that includes all of the possible links between different constructs, and then we used the modification indices to improve the model fit. We kept path links which were statistically significant in SEM path analysis. Before performing gender mediation analysis, we first tested the gender moderation relations between each pair of constructs using multigroup SEM (to investigate any interaction effects with gender), which includes testing of factor loadings, indicator intercepts, residual variances, and regression coefficients. Results showed that in all of our models strong measurement invariance holds and there is no difference in any regression coefficients by gender, which allowed us to perform the gender mediation analysis using SEM (see the Appendix A for detailed multigroup SEM analysis results).

One advantage of SEM is that it shows not only the direct regression relation between two constructs but also all of the indirect relations mediated through other constructs,

which allowed us to calculate the total regression effect by adding the direct and indirect regression coefficients up. In this study, we first considered a model with perceived recognition as the only perception of learning environment construct to see how students' physics self-efficacy, interest, and identity were predicted by it. Then, we added peer interaction or sense of belonging as additional constructs in the perception of learning environment. Finally, our model included all of the perception of learning environment constructs. We analyzed the variance in each motivational outcome denoting students' motivational characteristics explained by each model to understand the unique role played by each perception of learning environment component and to determine if all three perception of learning environment components are productive.

IV. RESULTS

A. Gender differences in motivational characteristics

The results of the regular two-sample t test are shown in Tables III and IV. We also used a paired sample t test for matched pairs to compare students' pre- and post-self-efficacy and interest, and the results are similar to those using the regular t test (see Appendix B for results of the paired-sample t test for matched pairs of pre- and post-responses). As shown in Table III, female students had significantly lower average interest and self-efficacy scores in both pre- and postsurvey than male students, and these gender differences increased by the end of the semester. The effect size given by Cohen's d [109] of gender difference in physics interest increased from 0.54 to 0.60, and the effect size of gender difference in self-efficacy increased from 0.32 to 0.53. In addition, even though students' interest and self-efficacy dropped generally from pre to post (see Table III), female students' interest and self-efficacy dropped ($d = 0.52$ for self-efficacy and $d = 0.30$ for interest) even more than male students' ($d = 0.27$ for self-efficacy and $d = 0.19$ for interest). Table IV shows the average scores on the other constructs (perception of peer interaction, perceived recognition, sense of belonging and identity) in the postsurvey. As shown in Table IV, female students also had significantly lower average scores in all of the four constructs than male students, and the effect sizes are all in the medium range [109].

TABLE III. Descriptive statistics of pre- and post-self-efficacy and interest for female and male students.

Gender	Pre-interest	Postinterest	Statistics		Pre-SE	Post-SE	Statistics	
	(1-4)	(1-4)			(1-4)	(1-4)		
	Mean	Mean	p value	Cohen's d	Mean	Mean	p value	Cohen's d
Male	3.19	3.08	<0.001	0.19	3.12	2.98	<0.001	0.27
Female	2.89	2.70	<0.001	0.30	2.96	2.68	<0.001	0.52
p value	<0.001	<0.001			<0.001	<0.001		
Cohen's d	0.54	0.60			0.32	0.53		

TABLE IV. Descriptive statistics of peer interaction, perceived recognition, sense of belonging, and identity for female and male students.

	Post-peer Int	Post-perceived recog	Postsense of belonging	Post-identity
Gender	(1-4)	(1-4)	(1-5)	(1-4)
Male	2.97	2.60	3.73	2.63
Female	2.68	2.24	3.33	2.17
<i>p</i> value	<0.001	<0.001	<0.001	<0.001
Cohen's <i>d</i>	0.44	0.49	0.46	0.56

B. Perception of learning environment mediation models

In this section, we will show the predictive relationships among the constructs using SEM models. Because many studies have shown that perceived recognition is a strong predictor of students' motivational beliefs and identity [79,113–115], all of the models shown in the main text of this paper include perceived recognition as one of the perception of learning environment constructs (see Appendix C for results of other SEM models). First, perceived recognition was the only perception of learning environment construct in the model. Then we added peer interaction or sense of belonging to the perception of learning environment one by one to analyze how each helped to predict students' identity, post-self-efficacy, and interest. Finally, we included all of the three constructs in our model and studied how these constructs mediated the outcomes together and what role was played by each of them.

Figure 2 shows the path analysis results of the SEM model in which perceived recognition is the only

perception of learning environment construct. The model fit indices suggest a good fit to the data: CFI = 0.939 (>0.90), TLI = 0.927 (>0.90), RMSEA = 0.055 (<0.08), and SRME = 0.043 (<0.08). The solid lines represent regression paths, and the numbers on the lines are regression coefficients (β values), which represent the strength of the regression relations. As we can see in Fig. 2., there are two paths going from pre-self-efficacy to post-self-efficacy. One path goes from pre to post directly, and the other path goes through the perceived recognition. The regression coefficient of the direct path from pre-self-efficacy to post-self-efficacy is 0.35. The regression coefficient of the indirect path can be calculated by multiplying the regression coefficients from pre-self-efficacy to perceived recognition ($\beta = 0.28$) and the regression coefficient from perceived recognition to post-self-efficacy ($\beta = 0.50$), which gives us $0.28 \times 0.50 = 0.14$. Similarly, the direct effect from pre-interest to postinterest is $\beta = 0.72$, and the indirect effect is $0.42 \times 0.25 = 0.11$.

We then added peer interaction to the model. The results of the SEM model are presented visually in Fig. 3. This model also fit the data very well. CFI = 0.951 (>0.90), TLI = 0.943 (>0.90), RMSEA = 0.048 (<0.08), and SRMR = 0.040 (<0.08). The results show that peer interaction directly predicts post-self-efficacy and interest. By comparing this model with the one only including perceived recognition, we can see that after adding peer interaction to the model, the direct effect of pre-self-efficacy on post-self-efficacy became weaker (the direct β value dropped from 0.35 to 0.29). In addition, the direct effect of perceived recognition on post-self-efficacy also decreased (the β value from perceived recognition to post-self-efficacy dropped from 0.50 to 0.35). This is due to the shared variance between peer interaction and the other

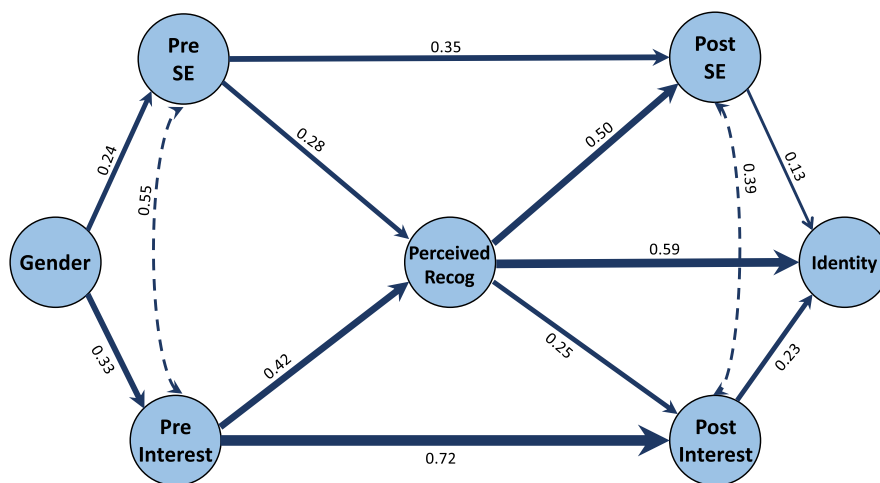


FIG. 2. Schematic diagram of the path analysis part of the structural equation modeling between gender and being a physics person through self-efficacy, interest, and perceived recognition. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of β value (standardized regression coefficient). All β values are significant with $p < 0.001$.

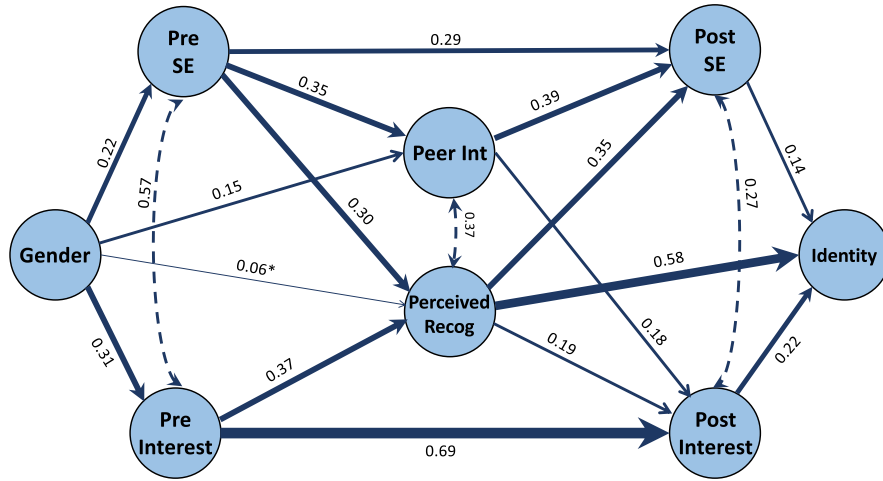


FIG. 3. Schematic diagram of the path analysis part of the structural equation modeling between gender and being a physics person through self-efficacy, interest, perceived recognition, and peer interaction. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of β value (standardized regression coefficient) with $0.01 < p < 0.05$ indicated by *. Other regression lines show relations with $p < 0.001$.

direct predictors (pre-self-efficacy and perceived recognition) of post-self-efficacy. The regression coefficient from a predictor to an outcome represents the expected changes in the outcome as a result of changes in the predictor in standardized deviation units while controlling for the correlated effects of other predictors [116]. When peer interaction was added to the model, the direct effect of pre-self-efficacy and perceived recognition on post-self-efficacy decreased because more correlated effect had been controlled for. Similarly, the direct effect of pre-interest and perceived recognition on postinterest also decreased because the effect that correlated with peer interaction had been controlled for. It is noteworthy that the direct effect of perceived recognition, postinterest and

post-self-efficacy on identity did not change after adding peer interaction. This is because peer interaction does not predict identity directly.

Then, we analyzed a SEM model which only includes perceived recognition and sense of belonging as shown in Fig. 4. The model also fits the data well [CFI = 0.930 (>0.90), TLI = 0.919 (>0.90), RMSEA = 0.054 (<0.08) and SRMR = 0.043 (<0.08)]. Figure 4 shows that sense of belonging is predicted by pre-self-efficacy and interest, and it directly predicts post-self-efficacy and interest. In addition, there is a strong correlation between sense of belonging and perceived recognition. Thus, there is more correlated effect being controlled for when estimating the regression coefficients from the predictors to the outcomes.

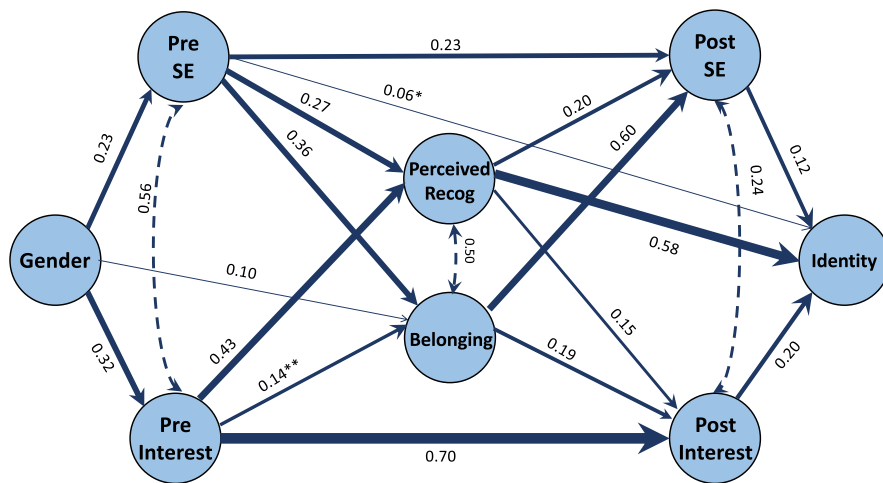


FIG. 4. Schematic diagram of the path analysis part of the structural equation modeling between gender and being a physics person through self-efficacy, interest, perceived recognition, and peer interaction. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of the β value (standardized regression coefficient) with $0.01 < p < 0.05$ indicated by * and $0.001 < p < 0.01$ indicated by **. Other regression lines show relations with $p < 0.001$.

For example, compared to the model only including perceived recognition (Fig. 2), adding sense of belonging decreased the direct regression coefficients from pre-self-efficacy, pre-interest, and perceived recognition to post-self-efficacy and postinterest. It is noteworthy that even though pre-self-efficacy directly predicts identity in this model as shown in Fig. 4, because the regression coefficient of this path is very small ($\beta = 0.06$), the regression coefficients from post-self-efficacy, postinterest, and perceived recognition to identity are still almost the same as in the first two models discussed earlier.

Finally, we analyzed a SEM model which includes all of the three perception of learning environment constructs. Figure 5 shows the results visually. This model also fits the data very well [CFI = 0.941 (>0.90), TLI = 0.932 (>0.90), RMSEA = 0.049 (<0.08) and SRMR = 0.042 (<0.08)]. As shown in Fig. 5, all three perception of learning environment constructs directly predict post-self-efficacy and postinterest, and the direct effect from pre-self-efficacy to post-self-efficacy and from pre-interest to postinterest are weaker than any models discussed earlier. In other words, as we added more constructs to the perception of learning environment, the strength of the direct paths from pre to post decreased. This result indicates that the perception of learning environment is mediating the effect of students' pre-self-efficacy and pre-interest on their motivational outcomes. Because there is no other construct predicting identity directly apart from perceived recognition, post-self-efficacy and postinterest, the regression coefficients from the three predictors to identity are almost the same in all of the above models. This result is consistent with Godwin *et al.* and Kalender *et al.*' prior work on physics identity framework [63,79]. In addition, as shown in Fig. 2–5, as more perception of learning environment

constructs are added, the residual covariance between post-self-efficacy and postinterest decreases while the residual covariance between pre-self-efficacy and pre-interest almost stays the same. This is because the residual covariance represents covariance between the constructs that is not explained by their predictors. This means that as we added more perception of learning environment constructs, more variance in post-self-efficacy and postinterest was explained.

Figure 5 shows that gender not only directly predicts the three perception of learning environment constructs but also indirectly predicts them through pre-self-efficacy and pre-interest. For example, the direct effect of gender on peer interaction is 0.15. The indirect effect can be calculated by multiplying the coefficient 0.22 (gender → pre-self-efficacy) and coefficient 0.35 (pre-self-efficacy → peer interaction). Thus, the total effect of gender on peer interaction is $0.15 + 0.22 \times 0.35 = 0.23$. Similarly, the total effect of gender on perceived recognition is $0.06 + 0.22 \times 0.32 + 0.31 \times 0.35 = 0.24$, and the total effect of gender on sense of belonging is $0.14 + 0.22 \times 0.45 = 0.24$. These results are consistent with the descriptive statistics shown in Table IV, which shows that there are statistically significant gender differences in all three perception of learning environment constructs and the effect sizes are all in the medium range [109].

Although there were large gender differences in students' pre- and post-self-efficacy and interest, Fig. 5 shows gender mediation and clarifies that gender only directly predicts pre-self-efficacy, pre-interest, and the three learning environment constructs. Thus, Fig. 5 reveals that the gender differences in students' post-self-efficacy, postinterest, and identity shown in Table III were mediated by the different components of the perception of the learning environment.

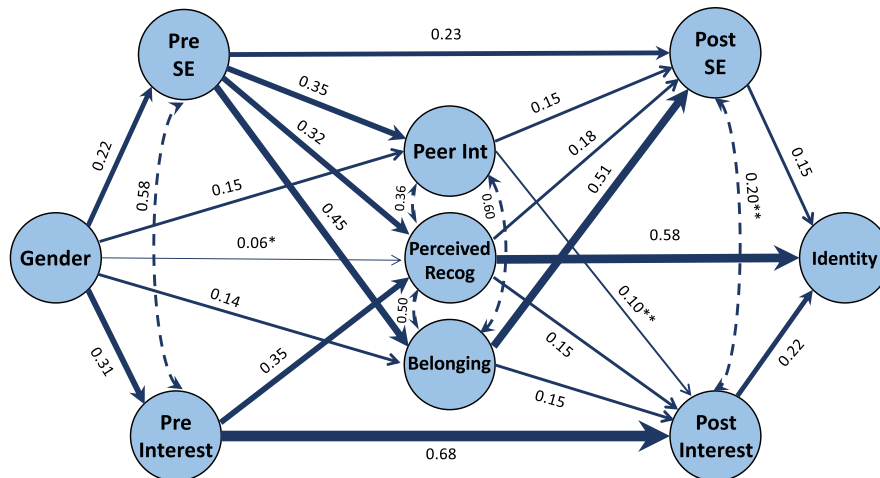


FIG. 5. Schematic diagram of the path analysis part of the structural equation modeling between gender and being a physics person through self-efficacy, interest, perceived recognition, peer interaction, and sense of belonging. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of the β value (standardized regression coefficient) with $0.01 < p < 0.05$ indicated by * and $0.001 < p < 0.01$ indicated by **. Other regression lines show relations with $p < 0.001$.

We note that due to reasons such as societal stereotypes and biases about who belongs in physics and can excel in it, women already had lower self-efficacy and interest at the beginning of the course. Moreover, Table III shows that the gender differences in students' self-efficacy and interest actually increased from the beginning to the end of the course. Thus, our results indicate that the current learning environment disadvantages female students more than male students and is not equitable and inclusive.

In this study, we investigated how students' perception of learning environment predicts their motivational outcomes by controlling for gender, self-efficacy, and interest at the beginning of the course. Even though two out of the three components of pre-identity (pre-self-efficacy and pre-interest) have already been controlled for, it would still be helpful to check the model in which pre-identity is also controlled for. However, because the identity construct was added to our survey at the end of the first fall semester we studied, we do not have the pre-identity data for that semester. Thus, we checked the model including pre-identity with only the second fall semester's data (see Appendix D for results of the path analysis part of this SEM model). We find that the direct effect from pre-identity to postidentity is very small and there are no qualitative changes with regard to how the perception of learning environment predicts outcomes compared with the model without pre-identity.

To summarize how the outcome constructs were predicted through both direct and indirect paths, we calculated the regression coefficient for each path in the model that includes all three learning environment constructs (Table V). The indirect effect of pre-self-efficacy on students' post-self-efficacy is 0.34 which is even larger than the direct effect 0.23. This means that most effects of

pre-self-efficacy on post-self-efficacy were mediated by the perception of learning environment. It is noteworthy that the effect of students' sense of belonging on post-self-efficacy ($\beta = 0.51$) is almost the same as the effect of pre-self-efficacy on post-self-efficacy ($\beta = 0.57$). Unlike post-self-efficacy, students' postinterest is mainly predicted by their pre-interest directly, and the effect of perception of learning environment constructs is small. In addition, we found that self-efficacy, interest, and perceived recognition are the three main predictors of identity, and perceived recognition is the major predictor, which is consistent with prior studies about physics identity [3,63,79].

Table V shows that the perception of learning environment is mediating the effect of students' pre-self-efficacy and pre-interest on their motivational outcomes. To further understand the role played by each learning environment construct, we calculated the coefficients of determination R^2 (fraction of variance explained) for each construct in seven different SEM models with different combinations of components of the perception of learning environment (Table VI). We found in all models, input constructs have very small R^2 (R^2 for pre-self-efficacy is 0.05 and R^2 for pre-interest is 0.10). This is because they are input constructs that we are controlling for, which are only explained by gender. The three perception of learning environment constructs are not only explained by gender but also by pre-self-efficacy and pre-interest. In all models, the variance in peer interaction is explained 17%, the variance in perceived recognition is explained 39%, and the variance in sense of belonging is explained around 24%.

R^2 for all outcome constructs are pretty high, which means that our models have explained most variance in them. It is noteworthy that R^2 values for postinterest are around 0.79 in all of the models, which means that the

TABLE V. Regression coefficients (β) of direct and indirect paths for three outcomes predicted by various predictors in the model which includes perception of learning environment (sense of belonging, peer interaction, and perceived recognition).

Predictor	Outcome	Direct	Indirect	Total (Direct + Indirect)
Pre-SE	Post-SE	0.23	0.34	0.57
Pre-SE	Postinterest	0.00	0.15	0.15
Pre-SE	Identity	0.00	0.29	0.29
Pre-Interest	Post-SE	0.00	0.06	0.06
Pre-Interest	Postinterest	0.68	0.05	0.73
Pre-Interest	Identity	0.00	0.17	0.17
Peer Int	Post-SE	0.15	0.00	0.15
Peer Int	Postinterest	0.10	0.00	0.10
Peer Int	Identity	0.00	0.04	0.04
Perceived Recog	Post-SE	0.18	0.00	0.18
Perceived Recog	Postinterest	0.15	0.00	0.15
Perceived Recog	Identity	0.58	0.06	0.64
Sense of Belonging	Post-SE	0.51	0.00	0.51
Sense of Belonging	Postinterest	0.15	0.00	0.15
Sense of Belonging	Identity	0.00	0.11	0.11

TABLE VI. Coefficient of determination (R^2) for various constructs in different models for impact of the perception of learning environment. All R^2 values are significant and p values <0.001 .

Construct	Models						
	Recog	Peer	Bel	Peer + Recog	Peer + Bel	Recog + Bel	Peer + Recog + Bel
Pre-SE	0.05	0.04	0.04	0.05	0.05	0.05	0.05
Pre-Interest	0.11	0.10	0.10	0.10	0.09	0.10	0.10
Peer Int	...	0.17	...	0.17	0.17	...	0.17
Recog	0.39	0.39	...	0.39	0.38
Belonging	0.24	...	0.25	0.24	0.25
Post-SE	0.56	0.61	0.75	0.68	0.77	0.77	0.79
Post-Interest	0.79	0.77	0.78	0.80	0.78	0.80	0.80
Identity	0.74	0.61	0.61	0.74	0.61	0.75	0.75

models which include any of the three learning environment constructs can explain 79% of the variance of postinterest. However, there is 75% of the variance of post-self-efficacy explained by the model which only includes sense of belonging, which is larger than that explained by the other two single construct models and is very close to that explained by the model including all of the three perception of learning environment constructs. In addition, we found that the models including sense of belonging always explain more variance in post-self-efficacy than the models without sense of belonging. For example, the model including perceived recognition and sense of belonging explains 77% of variance of post-self-efficacy, and the model including peer interaction and sense of belonging also explains 77% of variance of post-self-efficacy; however, the combination of peer interaction and perceived recognition only explains 68% of variance of post-self-efficacy. Similarly, the model that only includes perceived recognition explains 74% of the variance of identity, and adding peer interaction or sense of belonging does not help explain the variance in identity further. Table VI shows that both sense of belonging and perceived recognition play unique roles in the perception of learning environment in explaining students' motivational outcomes. However, peer interaction covaries with a sense of belonging and perceived recognition and uniquely explains very small percentages of the variance in the outcomes. However, this does not mean that peer interaction is not important. In contrast, the covariation actually suggests a possibility that students' sense of belonging and perceived recognition can potentially be shaped by helping students interact meaningfully with peers (which in turn can improve their motivational outcomes). Thus, we believe the model including all of the three perception of learning environment constructs is productive.

V. SUMMARY AND DISCUSSION

We studied students' physics motivational beliefs in a calculus-based introductory physics course. In particular, we studied how the perception of learning environment—

including peer interaction, perceived recognition, and sense of belonging—predicts students' physics self-efficacy, interest, and identity at the end of the course by controlling for their gender as well as their self-efficacy and interest at the beginning of the course.

We found that both male and female students' self-efficacy and interest dropped from pre to post but female students' dropped even more than male students' (Table III). In addition, we found significant gender differences disadvantaging female students in all motivational constructs in our models with the effect sizes for all of them in the medium range (Table V) [109]. The gender differences in students' perception of learning environment may partially explain the finding that the gender differences in students' self-efficacy and interest increased by the end of the course. These results show that the current learning environment influenced students' motivational belief in a negative way and was even more detrimental to female students' feelings.

To further understand how the perception of learning environment predicts students' motivational outcomes, we performed structural equation modeling to analyze predictive relationships among the constructs. Our results show that the perception of learning environment constructs directly predict students' motivational outcomes even after controlling for students' gender, pre-self-efficacy, and pre-interest. In particular, we found that in the learning environment, students' sense of belonging is the major predictor of their post-self-efficacy, and students' perceived recognition is the major predictor of their physics identity (Fig. 5).

Finally, we studied the role played by each of the three perception of learning environment components in explaining students' motivational outcomes. By comparing the fraction of variance in motivational outcomes explained by each model, we found that perceived recognition uniquely contributed most in explaining identity, and sense of belonging uniquely contributed most in explaining post-self-efficacy. Our results are consistent with Kalender *et al.*'s prior work [72,79], which showed that perceived recognition is the major predictor of identity and there is a

strong correlation between students' sense of belonging and self-efficacy. However, they did not take into account students' motivational characteristics at the beginning of the course. In our study, we found that even after controlling for students' gender, pre-interest, and pre-self-efficacy, the perception of learning environment still plays a very important role in predicting students' motivational outcomes. We note that even though peer interaction covaries with sense of belonging and perceived recognition and uniquely explains very small percentages of the variance in the outcomes, this does not mean that peer interaction is not important. Many instructors may not know how to implement strategies to improve students' sense of belonging. The covariation suggests a possibility that students' sense of belonging and perceived recognition may possibly be shaped by helping students interact meaningfully with peers (which in turn can improve their learning outcomes). Thus, we believe the model including all of the three learning environment constructs is productive.

In this study, we found that students' postinterest was mainly predicted by their pre-interest and the perception of learning environment predicted only a small amount. This means that our current learning environment barely helped students build their physics interest. Actually, both male and female students' interest in physics dropped by the end of the course. Even though our results show that the perception of learning environment had a large effect on students' post-self-efficacy, the drop in students' average self-efficacy scores after instruction shows that this overall effect of the current learning environment in introductory physics is actually negative.

Another important finding is that even though there are significant gender differences in students' post-self-efficacy, interest, and identity, gender does not directly predict them. This means that all of the gender differences in students' motivational outcomes can be explained by the gender differences in the predictors. This finding indicates that we may be able to bridge the gender gap in students' self-efficacy, interest and identity by developing an inclusive and equitable learning environment. Unfortunately, our results show that the current learning environment did not bridge the gap and actually enlarged it. Because of reasons such as societal stereotypes and biases about who belongs in physics and can excel in it, women already had lower self-efficacy and interest at the beginning of the course. The current learning environment disadvantaging female students may further impact their motivational beliefs.

Our findings suggest that an inclusive and equitable learning environment is very important for helping students improve their motivational belief in physics. An ideal situation would be that all students get full scores in self-efficacy, interest, and identity after the course; however, this also means that there would be no variance in these motivational outcomes at all and the effect of the perception of learning environment would also be zero. A

more realistic situation is that we improve the learning environment so that everyone develops a significantly higher self-efficacy, interest, and identity at the end of the course. As we found, the perception of learning environment directly predicts students' motivational outcomes at the end of the course (as shown in Fig. 5), so it is reasonable to expect that we can enlarge this effect (which means the regression coefficients from the perception of learning environment to students' post-self-efficacy, interest, and identity could be even larger). In addition, we should try to reduce the effect of gender, pre-self-efficacy, and interest on the perception of learning environment. Thus, students can equally benefit from the inclusive and equitable learning environment regardless of the gender and what their motivational beliefs were at the beginning of the course. We note that even though we find that the perception of learning environment did not explain students' interest much, this does not mean that interest cannot be changed. People's interest can be triggered and maintained by external factors [81]. For example, effective evidence-based instructional conditions or learning environments that include group work, puzzles, computers, and so on have been found to trigger situational interest [117–122]. Thus, it is important for instructors and researchers to develop an engaging evidence-based learning environment in which students can develop their interest in physics.

This investigation was conducted in a traditionally taught lecture-based course. It would definitely help to incorporate more research-validated active engagement pedagogies, but that is not enough. For example, according to prior studies [123,124], active engagement in an inequitable learning environment actually can increase the gender gap in students' performance because the stereotyped group (e.g., women) may not feel safe to participate without feeling judged or anxious if the environment is not equitable and inclusive. Instructors will need to create an equitable and inclusive learning environment keeping in mind how the societal stereotypes and biases about who belong in physics and who can excel in physics impacts the stereotyped groups. Some social-psychological interventions such as value-affirmation intervention and ecological-belonging intervention have been shown to reduce gender gaps in students' performance [125,126]. However, further studies are needed to understand whether they help to improve students' perceptions of learning environment and motivational outcomes in a course.

In our study, responses to the survey were all self-reported by students. It would be helpful to interview more students to get a deeper qualitative understanding of what they had experienced during the learning process and how the perception of learning environment influenced their learning behavior and motivational characteristics. In addition, it would also be interesting to look at the relation between the perception of learning environment and students' motivational outcomes in algebra-based physics

courses. Is the self-efficacy and interest of female students higher in algebra-based physics courses than that in calculus-based physics courses? Are the gender differences in students’ motivational beliefs smaller in algebra-based physics courses because female students are overrepresented in these courses? Even though the general societal stereotype threat still exists even for female students in algebra-based physics courses, if their motivational beliefs can be protected by a classroom where they have many female peers, this may itself be a useful finding.

ACKNOWLEDGMENTS

This work was supported by Grant No. DUE-1524575 from the National Science Foundation. We would like to thank all students whose data were analyzed and Dr. Robert Devaty for his constructive feedback on the manuscript.

APPENDIX A: MODERATION ANALYSIS

We did a moderation analysis to test whether gender moderates the relationship between any two constructs in the models (i.e., Does the strength of relationships given by the standardized regression coefficients between any two constructs in the models differ for women and men?). We used the R [127] software package “lavaan” to conduct multigroup SEM. We initially tested for measurement invariance. In other words, we looked at whether the factor loadings, intercepts, and residual variances of the observed variables are equal in the model for the latent constructs so we can confidently perform multigroup analysis. The analysis involved introducing certain constraints in steps and testing the model differences from the previous step. In each step, we compared the model to both the previous step and the freely estimated model, i.e., the model in which all parameters are freely estimated for each gender group. First, to test for “weak” or “metric” measurement invariance, we ran the model in which only factor loadings were fixed to equality across both gender groups, but intercept and errors were allowed to differ. The model was not statistically significantly different from the freely estimated

model according to a likelihood ratio test, so weak measurement invariance holds (chi-square difference $\Delta\chi^2 = 21.324$, degree of freedom difference $\Delta d.o.f. = 21$, and nonsignificant $p = 0.439$). Next, we tested for “strong” or “scalar” measurement invariance by fixing both factor loadings and intercepts to equality across gender groups. This model was not statistically significantly different from either the metric invariance model ($\Delta\chi^2 = 25.314$, $\Delta d.o.f. = 21$, $p = 0.234$) or the freely estimated model ($\Delta\chi^2 = 46.637$, $\Delta d.o.f. = 42$, $p = 0.288$), so strong measurement invariance holds. Finally, to test for “strict” measurement invariance we fixed factor loadings, intercepts, and residual variances to equality. In this step, there was a statistically significant difference from the previous models, therefore “strict invariance” did not hold when we compared to scalar measurement ($\Delta\chi^2 = 64.732$, $\Delta d.o.f. = 27$, $p < 0.001$). However, strict invariance is unlikely to hold in most situations. Therefore, since strong measurement invariance holds for this model, we continued on to perform other group comparisons.

Next, we ran a multigroup SEM in which all regression estimates were fixed to equality for female and male students in addition to the factor loadings and intercepts, and we compared this model with freely estimated model. There was no statistically significant difference between the two models, so we reported the model where regression pathways are equal for men and women. The model fit parameters for this case were acceptable (RMSEA = 0.051, SRMR = 0.057, CFI = 0.931, TLI = 0.928). The multigroup SEM results suggest that regression pathways among the constructs did not show differences across gender when we compared to freely estimated model ($\Delta\chi^2 = 77.059$, $\Delta d.o.f. = 62$, $p = 0.094$) or to the scalar model ($\Delta\chi^2 = 30.422$, $\Delta d.o.f. = 20$, $p = 0.063$). However, the means of the latent variables showed the same gender differences that have been reported in the mediation models. That is, there were large gender differences in students’ pre-self-efficacy and interest and slight differences in peer interaction, perceived recognition, and sense of belonging.

APPENDIX B: PAIRED-SAMPLE *t* TEST

TABLE VII. Paired-sample *t* test for matched pairs (matched pair refers to students who took both pre- and postsurveys) of pre- and post-self-efficacy and interest for female and male students.

	Pre-Interest		Postinterest		Pre-SE		Post-SE	
	(1-4)	(1-4)	Statistics		(1-4)	(1-4)	Statistics	
Gender	Mean	Mean	<i>p</i> value	Cohen’s <i>d</i>	Mean	Mean	<i>p</i> value	Cohen’s <i>d</i>
Male (<i>N</i> = 662)	3.19	3.07	<0.001	0.24	3.12	2.98	<0.001	0.25
Female (382)	2.89	2.72	<0.001	0.33	2.96	2.70	<0.001	0.45

APPENDIX C: SEM RESULTS OF OTHER MODELS

The figure below presents the path analysis results of the SEM models including only peer interaction, only sense of belonging, and both peer interaction and sense of belonging as learning environment constructs.

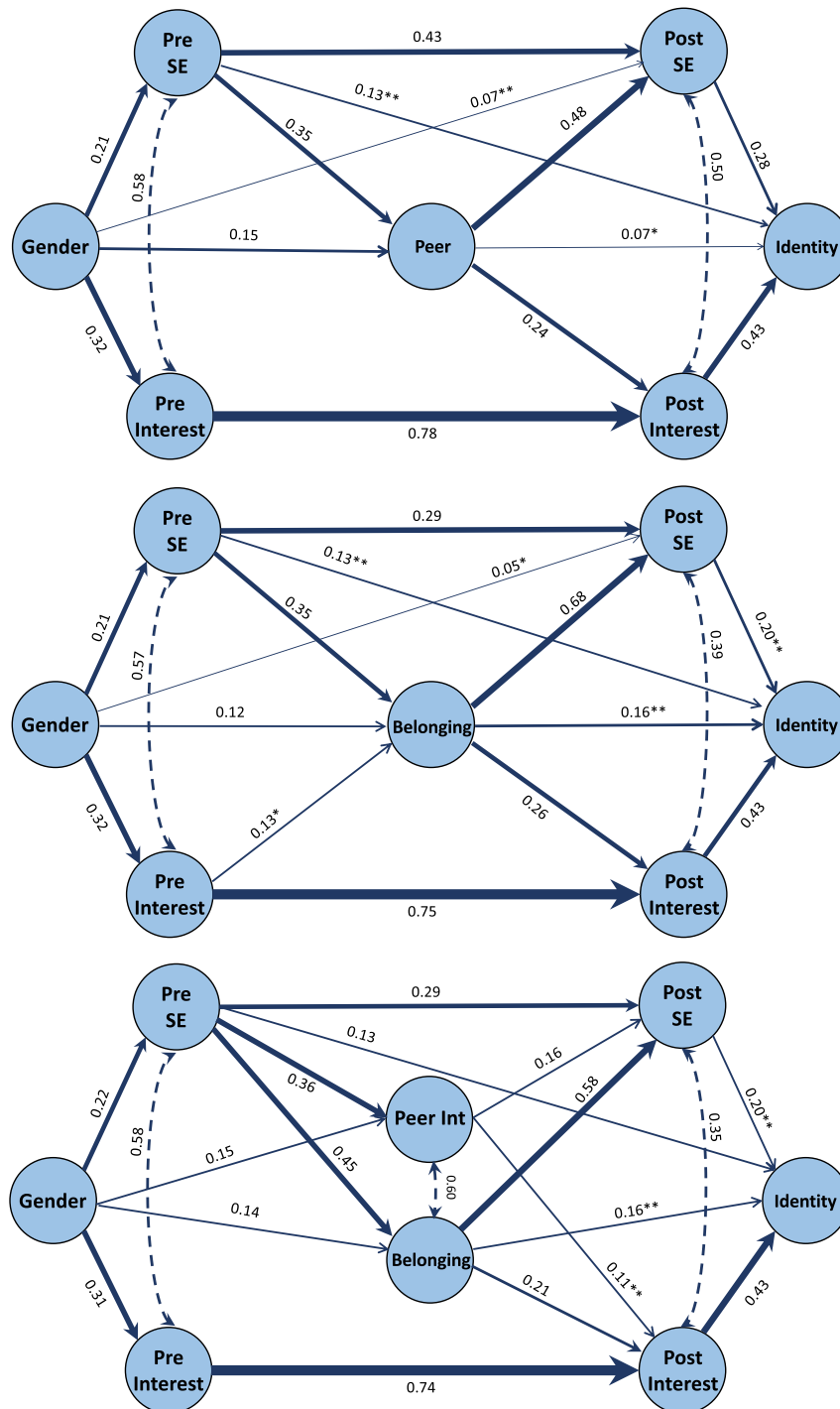


FIG. 6. Schematic diagram of the path analysis part of the structural equation modeling of perception of learning environment models including only peer interaction, only sense of belonging and both peer interaction and sense of belonging. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of β value (standardized regression coefficient) with $0.01 < p < 0.05$ indicated by * and $0.001 < p < 0.01$ indicated by **. Other regression lines show relations with $p < 0.001$.

APPENDIX D: SEM RESULTS OF THE MODEL INCLUDING PRE-IDENTITY

The figure below presents the path analysis results of the SEM model including pre-identity.

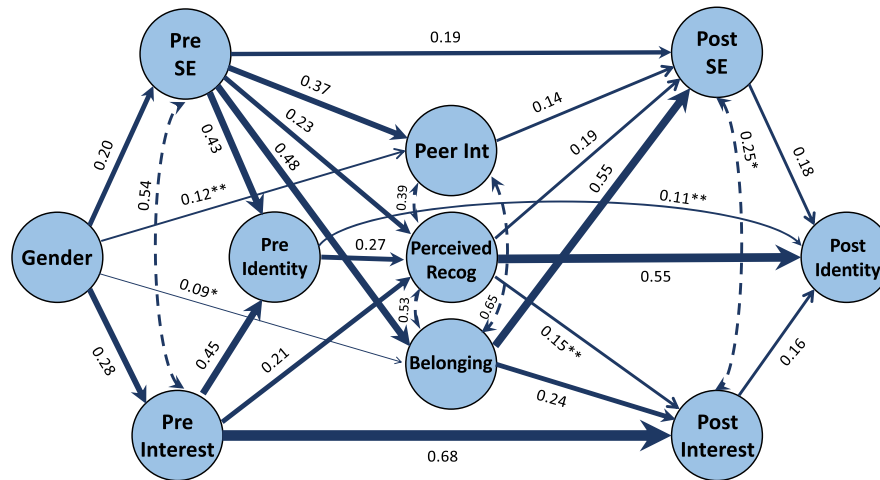


FIG. 7. Schematic diagram of the path analysis part of the SEM model including pre-identity. The solid lines represent regression paths, and the dashed lines represent residual covariances. The regression line thickness corresponds to the magnitude of the β value (standardized regression coefficient) with $0.01 < p < 0.05$ indicated by * and $0.001 < p < 0.01$ indicated by **. Other regression lines show relations with $p < 0.001$.

- [1] H. B. Carlone and A. Johnson, Understanding the science experiences of successful women of color: Science identity as an analytic lens, *J. Res. Sci. Teach.* **44**, 1187 (2007).
- [2] 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).
- [3] Z. Hazari, G. Sonnert, P. M. Sadler, and M.-C. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, *J. Res. Sci. Teach.* **47**, 978 (2010).
- [4] J. M. Nissen and J. T. Shemwell, Gender, experience, and self-efficacy in introductory physics, *Phys. Rev. Phys. Educ. Res.* **12**, 020105 (2016).
- [5] L. J. Sax, K. J. Lehman, R. S. Barthelemy, and G. Lim, Women in physics: A comparison to science, technology, engineering, and math education over four decades, *Phys. Rev. Phys. Educ. Res.* **12**, 020108 (2016).
- [6] J. Blue, A. L. Traxler, and X. C. Cid, Gender matters, *Phys. Today* **71**, 40 (2018).
- [7] M. Lorenzo, C. H. Crouch, and E. Mazur, Reducing the gender gap in the physics classroom, *Am. J. Phys.* **74**, 118 (2006).
- [8] A. L. Traxler, X. C. Cid, J. Blue, and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future, *Phys. Rev. Phys. Educ. Res.* **12**, 020114 (2016).
- [9] K. Rosa and F. M. Mensah, Educational pathways of Black women physicists: Stories of experiencing and overcoming obstacles in life, *Phys. Rev. Phys. Educ. Res.* **12**, 020113 (2016).
- [10] B. Van Dusen and J. Nissen, Equity in college physics student learning: A critical quantitative intersectionality investigation, *J. Res. Sci. Teach.* **57**, 33 (2020).
- [11] 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).
- [12] 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).
- [13] E. Marshman, Z. Y. Kalender, T. Nokes-Malach, C. Schunn, and C. Singh, Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm?, *Phys. Rev. Phys. Educ. Res.* **14**, 020123 (2018).

- [14] C. M. Steele and J. Aronson, Stereotype threat and the intellectual test performance of African Americans, *J. Pers. Soc. Psychol.* **69**, 797 (1995).
- [15] G. C. Marchand and G. Taasobshirazi, Stereotype threat and women's performance in physics, *Int. J. Sci. Educ.* **35**, 3050 (2013).
- [16] R. Ivie, S. White, and R. Y. Chu, Women's and men's career choices in astronomy and astrophysics, *Phys. Rev. Phys. Educ. Res.* **12**, 020109 (2016).
- [17] 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. ST Phys. Educ. Res.* **9**, 020121 (2013).
- [18] K. M. Whitcomb and C. Singh, For physics majors, gender differences in introductory physics do not inform future physics performance, *Eur. J. Phys.* **41**, 065701 (2020).
- [19] K. M. Whitcomb, Z. Y. Kalender, T. J. Nokes-Malach, C. D. Schunn, and C. Singh, Comparison of self-efficacy and performance of engineering undergraduate women and men, *Int. J. Eng. Educ.* **36**, 1996 (2020), www.ijee.ie/latestissues/Vol36-6/24_ijee4004.pdf.
- [20] M. Good, A. Maries, and C. Singh, Impact of traditional or evidence-based active-engagement instruction on introductory female and male students' attitudes and approaches to physics problem solving, *Phys. Rev. Phys. Educ. Res.* **15**, 020129 (2019).
- [21] 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. Rev.* **33**, 304 (2013).
- [22] 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).
- [23] E. Lichtenberger and C. George-Jackson, Predicting high school students' interest in majoring in a STEM field: Insight into high school students' postsecondary plans, *J. Career Tech. Educ.* **28**, 19 (2013).
- [24] A. Wigfield *et al.*, Motivation to succeed, in *Handbook of Child Psychology, Social Emotional and Personality Development*, edited by W. Damon, R. Lerner, and N. Eisenberg (Wiley, New York, 2006), Vol. 3.
- [25] 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).
- [26] J. L. Smith, C. Sansone, and P. H. White, The stereotyped task engagement process: The role of interest and achievement motivation, *J. Educ. Psychol.* **99**, 99 (2007).
- [27] B. J. Zimmerman, Self-efficacy: An essential motive to learn, *Contemp. Educ. Psychol.* **25**, 82 (2000).
- [28] F. Pajares and D. Schunk, The development of academic self-efficacy, in *Development of Achievement Motivation*, Vol. 7 (Academic Press, United States, 2001).
- [29] H. M. Watt, The role of motivation in gendered educational and occupational trajectories related to maths, *Educ. Res. Eval.* **12**, 305 (2006).
- [30] A. L. Zeldin, S. L. Britner, and F. Pajares, A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers, *J. Res. Sci. Teach.* **45**, 1036 (2008).
- [31] N. E. Betz and G. Hackett, Applications of self-efficacy theory to understanding career choice behavior, *J. Soc. Clin. Psychol.* **4**, 279 (1986).
- [32] V. Sawtelle, E. Brewwe, and L. H. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, *J. Res. Sci. Teach.* **49**, 1096 (2012).
- [33] V. Tinto, Classrooms as communities: Exploring the educational character of student persistence, *J. Higher Educ.* **68**, 599 (1997).
- [34] J. Clark Blickenstaff, Women and science careers: leaky pipeline or gender filter?, *Gender Educ.* **17**, 369 (2005).
- [35] S. Cwik, K. Whitcomb, and C. Singh, How the learning environment predicts male and female students' motivational beliefs in algebra-based introductory physics courses, in *Proceedings of the 2020 Physics Education Research Conference, virtual conference* (2020), p. 104, <https://doi.org/10.1119/perc.2020.pr.Cwik>.
- [36] Y. Li, K. Whitcomb, and C. Singh, How learning environment predicts male and female students' physics motivational beliefs in introductory physics courses, in *Proceedings of the 2020 Physics Education Research Conference, virtual conference* (2020), p. 284, <https://doi.org/10.1119/perc.2020.pr.Li>.
- [37] E. Seymour, Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, *Sci. Educ.* **86**, 79 (2002).
- [38] A. B. Diekman, E. K. Clark, A. M. Johnston, E. R. Brown, and M. Steinberg, Malleability in communal goals and beliefs influences attraction to STEM careers: Evidence for a goal congruity perspective, *J. Pers. Soc. Psychol.* **101**, 902 (2011).
- [39] S. G. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, *J. Eng. Educ.* **87**, 369 (1998).
- [40] S. J. Correll, Gender and the career choice process: The role of biased self-assessments, *Am. J. Sociology* **106**, 1691 (2001).
- [41] S. J. Correll, Constraints into preferences: Gender, status, and emerging career aspirations, *Am. Sociol. Rev.* **69**, 93 (2004).
- [42] R. Elliott, A. C. Strenta, R. Adair, M. Matier, and J. Scott, The role of ethnicity in choosing and leaving science in highly selective institutions, *Res. High. Educ.* **37**, 681 (1996).
- [43] E. T. Pascarella and P. T. Terenzini, *How College Affects Students: Findings and Insights from Twenty Years of Research* (ERIC, Jossey-Bass, San Francisco, 1991).
- [44] C. Hill, C. Corbett, and A. St Rose, *Why so Few? Women in Science, Technology, Engineering, and Mathematics* (ERIC, Jossey-Bass, San Francisco, 2010).
- [45] S. Ahlqvist, B. London, and L. Rosenthal, Unstable identity compatibility: How gender rejection sensitivity undermines the success of women in science, technology, engineering, and mathematics fields, *Psychol. Sci.* **24**, 1644 (2013).

- [46] 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).
- [47] S. Cheryan and V. C. Plaut, Explaining underrepresentation: A theory of precluded interest, *Sex Roles* **63**, 475 (2010).
- [48] A. Bandura, Self-efficacy, in *Encyclopedia of Psychology*, 2nd ed., edited by R. J. Corsini (Wiley, New York, 1994), Vol. 3, p. 368.
- [49] A. Bandura, Social cognitive theory of self-regulation, *Organ. Behav. Hum. Decis. Process.* **50**, 248 (1991).
- [50] S. L. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, *J. Res. Sci. Teach.* **43**, 485 (2006).
- [51] Z. Y. Kalender, E. Marshman, C. D. Schunn, T. J. Nokes-Malach, and C. Singh, Damage caused by women's lower self-efficacy on physics learning, *Phys. Rev. Phys. Educ. Res.* **16**, 010118 (2020).
- [52] J. M. Harackiewicz, K. E. Barron, J. M. Tauer, and A. J. Elliot, Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation, *J. Educ. Psychol.* **94**, 562 (2002).
- [53] S. Hidi, Interest: A unique motivational variable, *Educ. Res. Rev.* **1**, 69 (2006).
- [54] K. A. Renninger and S. Hidi, Revisiting the conceptualization, measurement, and generation of interest, *Educ. Psychol.* **46**, 168 (2011).
- [55] A. Wigfield, J. S. Eccles, U. Schiefele, R. W. Roeser, and P. Davis-Kean, Development of achievement motivation, in *Handbook of Child Psychology: Social, Emotional, and Personality Development*, Vol. 3, 6th ed., edited by N. Eisenberg, W. Damon, and R. M. Lerner (John Wiley & Sons Inc, Hoboken, NJ, US, 2006), p. 933.
- [56] P. Häußler and L. Hoffmann, An intervention study to enhance girls' interest, self-concept, and achievement in physics classes, *J. Res. Sci. Teach.* **39**, 870 (2002).
- [57] A. Wigfield and J. S. Eccles, Expectancy-value theory of achievement motivation, *Contemp. Educ. Psychol.* **25**, 68 (2000).
- [58] A. Wigfield and J. S. Eccles, The development of achievement task values: A theoretical analysis, *Dev. Rev.* **12**, 265 (1992).
- [59] Z. Hazari, G. Potvin, R. M. Lock, F. Lung, G. Sonnert, and P. M. Sadler, Factors that affect the physical science career interest of female students: Testing five common hypotheses, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020115 (2013).
- [60] Z. Hazari, P. M. Sadler, and G. Sonnert, The science identity of college students: Exploring the intersection of gender, race, and ethnicity, *J. Coll. Sci. Teach.* **42**, 82 (2013).
- [61] C. Monsalve, Z. Hazari, D. McPadden, G. Sonnert, and P. M. Sadler, Examining the relationship between career outcome expectations and physics identity, in *Proceedings of the 2016 Physics Education Research Conference, Sacramento, CA* (AIP, New York, 2016), p. 228.
- [62] G. Potvin and Z. Hazari, The development and measurement of identity across the physical sciences, in *Proceedings of the 2013 Physics Education Research Conference, Portland, OR* (AIP, New York, 2013), p. 281.
- [63] A. Godwin, G. Potvin, Z. Hazari, and R. Lock, Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice, *J. Eng. Educ.* **105**, 312 (2016).
- [64] R. M. Lock, Z. Hazari, and G. Potvin, Physics career intentions: The effect of physics identity, math identity, and gender, *AIP Conf. Proc.* **1513**, 262 (2013).
- [65] 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).
- [66] A. T. Danielsson, Exploring woman university physics students 'doing gender' and 'doing physics', *Gender Educ.* **24**, 25 (2012).
- [67] A. J. Gonsalves, Exploring how gender figures the identity trajectories of two doctoral students in observational astrophysics, *Phys. Rev. Phys. Educ. Res.* **14**, 010146 (2018).
- [68] N. M. Hewitt and E. Seymour, A long, discouraging climb, *ASEE Prism* **1**, 24 (1992).
- [69] E. Seymour, N. M. Hewitt, and C. M. Friend, *Talking about Leaving: Why Undergraduates Leave the Sciences* (Westview Press, Boulder, CO, 1997), Vol. 12.
- [70] Z. Hazari and C. Cass, Towards meaningful physics recognition: What does this recognition actually look like?, *Phys. Teach.* **56**, 442 (2018).
- [71] Z. Hazari, E. Brewe, R. M. Goertzen, and T. Hodapp, The importance of high school physics teachers for female students' physics identity and persistence, *Phys. Teach.* **55**, 96 (2017).
- [72] Z. Y. Kalender, E. Marshman, C. D. Schunn, T. J. Nokes-Malach, and C. Singh, Gendered patterns in the construction of physics identity from motivational factors, *Phys. Rev. Phys. Educ. Res.* **15**, 020119 (2019).
- [73] L. Bian, S.-J. Leslie, and A. Cimpian, Gender stereotypes about intellectual ability emerge early and influence children's interests, *Science* **355**, 389 (2017).
- [74] D. Doucette and C. Singh, Why are there so few women in physics? Reflections on the experiences of two women, *Phys. Teach.* **58**, 297 (2020).
- [75] D. Doucette, R. Clark, and C. Singh, Hermione and the Secretary: how gendered task division in introductory physics labs can disrupt equitable learning, *Eur. J. Phys.* **41**, 035702 (2020).
- [76] S. Upton and L. F. Friedman, Where are all the female geniuses?, *Scientific American Mind* **23**, 63 (2012).
- [77] K. Crowley, M. A. Callanan, H. R. Tenenbaum, and E. Allen, Parents explain more often to boys than to girls during shared scientific thinking, *Psychol. Sci.* **12**, 258 (2001).
- [78] C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman, Science faculty's subtle gender biases favor male students, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16474 (2012).
- [79] Z. Y. Kalender, E. Marshman, C. D. Schunn, T. J. Nokes-Malach, and C. Singh, Why female science, technology, engineering, and mathematics majors do not identify with

- physics: They do not think others see them that way, *Phys. Rev. Phys. Educ. Res.* **15**, 020148 (2019).
- [80] Y. Li, K. Whitcomb, and C. Singh, How perception of being recognized or not recognized by instructors as a “physics person” impacts male and female students’ self-efficacy and performance, *Phys. Teach.* **58**, 484 (2020).
- [81] S. Hidi and K. A. Renninger, The four-phase model of interest development, *Educ. Psychol.* **41**, 111 (2006).
- [82] A. Bandura, W. Freeman, and R. Lightsey, Self-efficacy: The exercise of control, *J. Cogn. Psychotherapy* **13**, 158 (1999).
- [83] A. Bandura, Self-efficacy: toward a unifying theory of behavioral change, *Psychol. Rev.* **84**, 191 (1977).
- [84] M. Meeuwisse, S. E. Severiens, and M. P. Born, Learning environment, interaction, sense of belonging and study success in ethnically diverse student groups, *Res. High. Educ.* **51**, 528 (2010).
- [85] R. Masika and J. Jones, Building student belonging and engagement: Insights into higher education students’ experiences of participating and learning together, *Teach. High. Educ.* **21**, 138 (2016).
- [86] C. Goodenow, Classroom belonging among early adolescent students: Relationships to motivation and achievement, *J. Early Adolesc.* **13**, 21 (1993).
- [87] A. W. Astin, *What Matters in College: Four Critical Years Revisited* (Jossey-Bass, San Francisco, 1993).
- [88] D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They’re not magic, *Rev. Educ. Res.* **81**, 267 (2011).
- [89] I. Rodriguez, E. Brewe, V. Sawtelle, and L. H. Kramer, Impact of equity models and statistical measures on interpretations of educational reform, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020103 (2012).
- [90] D. Hooper, J. Coughlan, and M. i Mullen, Structural equation modelling: Guidelines for determining model fit, *Electronic J. Bus. Res. Methods* **6**, 5360 (2008).
- [91] C. K. Enders and D. L. Bandalos, The relative performance of full information maximum likelihood estimation for missing data in structural equation models, *Struct. Equ. Model.* **8**, 430 (2001).
- [92] D. Hammer, Epistemological beliefs in introductory physics, *Cognit. Instr.* **12**, 151 (1994).
- [93] Activation Lab Tools: measures, and data collection instruments (2017), <http://www.activationlab.org/tools/>.
- [94] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101 (2006).
- [95] B. M. Zwickl, T. Hirokawa, N. Finkelstein, and H. J. Lewandowski, Epistemology and expectations survey about experimental physics: Development and initial results, *Phys. Rev. ST Phys. Educ. Res.* **10**, 010120 (2014).
- [96] J. Schell and B. Lukoff, Peer instruction self-efficacy instrument [Developed at Harvard University] (to be published).
- [97] S. M. Glynn, P. Brickman, N. Armstrong, and G. Taasobshirazi, Science motivation questionnaire II: Validation with science majors and nonscience majors, *J. Res. Sci. Teach.* **48**, 1159 (2011).
- [98] PERTS Academic Mindsets Assessment (2020), <https://survey.perts.net/share/dlmooc>.
- [99] Z. Y. Kalender, E. Marshman, T. Nokes-Malach, C. Schunn, and C. Singh, Motivational characteristics of underrepresented ethnic and racial minority students in introductory physics courses, in *Proceedings of the 2017 Physics Education Research Conference, Cincinnati, OH* (2017), p. 204, <https://doi.org/10.1119/perc.2017.pr.046>.
- [100] T. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn, and C. Singh, Investigation of male and female students’ motivational characteristics throughout an introductory physics course sequence, in *Proceedings of the 2017 Physics Education Research Conference, Cincinnati, OH* (2017), p. 276, <https://doi.org/10.1119/perc.2017.pr.064>.
- [101] T. J. Nokes-Malach, Z. Y. Kalender, E. Marshman, C. D. Schunn, and C. Singh, Prior preparation and motivational characteristics mediate relations between gender and learning outcomes in introductory physics, in *Proceedings of the 2018 Physics Education Research Conference, Washington, DC* (2018), <https://doi.org/10.1119/perc.2018.pr.Nokes-Malach>.
- [102] Z. Y. Kalender, E. Marshman, C. D. Schunn, T. J. Nokes-Malach, and C. Singh, Large gender differences in physics self-efficacy at equal performance levels: A warning sign?, in *Proceeding of the 2018 Physics Education Research Conference, Washington, DC* (2018), <https://doi.org/10.1119/perc.2018.pr.Kalender>.
- [103] Z. Y. Kalender, E. Marshman, T. Nokes-Malach, C. Schunn, and C. Singh, Beliefs about competence: The story of self-efficacy, gender, and physics, in *Diversity Across Disciplines: Research on People, Policy, Process and Paradigm*, edited by A. Murrell, J. Petrie-Wyman, and A. Soudi (Information Age Publishing, 2019), <https://www.infoagepub.com/products/Diversity-Across-the-Disciplines>.
- [104] B. Thompson, *Exploratory and Confirmatory Factor Analysis* (American Psychological Association, Washington, DC, 2004).
- [105] L. J. Cronbach, Coefficient alpha and the internal structure of tests, *Psychometrika* **16**, 297 (1951).
- [106] K. Pearson and F. Galton, VII. Note on regression and inheritance in the case of two parents, *Proc. R. Soc. London* **58**, 240 (1895).
- [107] R. Likert, A technique for the measurement of attitudes, *Arch. Sci. Psychol.* **140**, 5 (1932).
- [108] W. S. Gosset, The probable error of a mean, *Biometrika* **6**, 1 (1908).
- [109] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Routledge, London, 2013).
- [110] A. J. Tomarken and N. G. Waller, Structural equation modeling: Strengths, limitations, and misconceptions, *Annu. Rev. Clin. Psychol.* **1**, 31 (2005).
- [111] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis: International Version* (Pearson, New Jersey, 2010).
- [112] H. Akoglu, User’s guide to correlation coefficients, *Turkish J. Emergency Med.* **18**, 91 (2018).

- [113] K. L. Tonso, Student engineers and engineer identity: Campus engineer identities as figured world, *Cult. Stud. Sci. Educ.* **1**, 273 (2006).
- [114] J. E. Stets and P. J. Burke, A sociological approach to self and identity, in *Handbook of Self and Identity* (Guilford Press, New York, 2003), p. 128.
- [115] Z. Hazari, C. Cass, and C. Beattie, Obscuring power structures in the physics classroom: Linking teacher positioning, student engagement, and physics identity development, *J. Res. Sci. Teach.* **52**, 735 (2015).
- [116] J. B. Grace and K. A. Bollen, Interpreting the results from multiple regression and structural equation models, *Bull. Ecol. Soc. Am.* **84**, 283 (2005).
- [117] D. Cordova and M. Lepper, Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization, and choice, *J. Educ. Psychol.* **88**, 715 (1996).
- [118] S. Hidi and W. Baird, Strategies for increasing text-based interest and students' recall of expository texts, *Read. Res. Q.* **23**, 465 (1988).
- [119] S. Hidi, J. Weiss, D. Berndorff, and J. Nolan, The role of gender, instruction and a cooperative learning technique in science education across formal and informal settings, in *Interest and learning: Proceedings of the Seeon conference on interest and gender*, (IPN Kiel, Germany, 1998), p. 215.
- [120] M. R. Lepper and D. I. Cordova, A desire to be taught: Instructional consequences of intrinsic motivation, *Motivation and emotion* **16**, 187 (1992).
- [121] M. Mitchell, Situational interest: Its multifaceted structure in the secondary school mathematics classroom, *J. Educ. Psychol.* **85**, 424 (1993).
- [122] J. A. Sloboda and J. Davidson, The young performing musician, in *Musical Beginnings: Origins and Development of Musical Competence* (Oxford University Press, Oxford, 1996), p. 171.
- [123] A. Maries, N. Karim, and C. Singh, Active learning in an inequitable learning environment can increase the gender performance gap: The negative impact of stereotype threat, *Phys. Teach.* **58**, 430 (2020).
- [124] A. Maries, N. I. Karim, and C. Singh, Is agreeing with a gender stereotype correlated with the performance of female students in introductory physics?, *Phys. Rev. Phys. Educ. Res.* **14**, 020119 (2018).
- [125] 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).
- [126] A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock, G. L. Cohen, and T. A. Ito, Reducing the gender achievement gap in college science: A classroom study of values affirmation, *Science* **330**, 1234 (2010).
- [127] R. C. Team, *R: A Language and Environment for Statistical Computing* (R foundation for statistical computing, Vienna, Austria, 2013).