Damage caused by societal stereotypes: Women have lower physics self-efficacy controlling for grade even in courses in which they outnumber men

Sonja Cwik[✉] and Chandralekha Singh[✉]

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

(Received 6 April 2021; accepted 11 October 2021; published 29 November 2021)

Societal stereotypes and biases pertaining to who belongs in physics and who can excel in it can impact motivational beliefs of women in physics courses. Prior research has shown that women have lower physics self-efficacy than men in physics courses in which women are underrepresented. However, prior research has generally not investigated similar issues in physics courses in which women make up the majority of students. This study examines self-efficacy of men and women with similar performance in introductory algebra-based physics courses in which women outnumber men at a large public research university in the US. These courses are taken primarily by biological science majors many of whom are interested in health professions. Although women are not underrepresented in these physics courses, societal stereotypes and biases internalized by female students over their lifetime can impact their self-efficacy when they take any physics course. We find a gender gap in self-efficacy disadvantaging women at the beginning of the course. However, unlike courses in which women are underrepresented, in which the self-efficacy gender gap often increases from the beginning to the end of the courses, we find that the self-efficacy gender gap for students who received a certain grade either remained constant or decreased somewhat. Moreover, except for the students who received an A grade, the average self-efficacy of most of the other student groups decreased from the beginning to the end of the semester. Additionally, we find that most of the self-efficacy gender gap is due to students' biased perceptions about their capability rather than the performance difference between women and men.

DOI: [10.1103/PhysRevPhysEducRes.17.020138](https://doi.org/10.1103/PhysRevPhysEducRes.17.020138)

I. INTRODUCTION AND THEORETICAL FRAMEWORK

Recently, there has been significant focus on the underrepresentation of women in many science, technology, engineering, and math (STEM) fields in order to increase their representation and participation in these fields [1–11]. Some of these research studies have focused on comparison of degree attainment by women and men or tracked processes of change in STEM undergraduate education [1–3]. Others focus on how women are still marginalized in these disciplines and strategies to improve the situation based upon a variety of research findings [4–7], or ways to improve the retention of women via classroom interventions [8,9]. Yet other research studies have focused on the root cause of women and other undergraduates leaving STEM majors [10,11]. Motivational beliefs such as identity and self-efficacy in different STEM domains can influence students' continuation in related STEM courses, majors,

and careers [12–21]. For example, students were more likely to take courses or pursue a career in science if they had higher competency belief or self-efficacy [12–16], display higher interest in science [17], or have a higher science identity [18–20]. Moreover, a gender gap favoring men in motivational factors [21–23] and conceptual tests [24] has been studied in STEM courses. Specifically, in the context of physics, several studies in the courses in which women are underrepresented have focused on understanding and addressing the low representation of women in these courses [25]. However, few studies on self-efficacy have focused on introductory algebra-based physics courses for biology and health science (bioscience) majors in which women tend to outnumber men. In a review paper that summarizes gender differences in physics courses by Madsen *et al.*, there is an explicit call to investigate the gender gap in various constructs in physics courses where women are not underrepresented [25].

Self-efficacy is a person's belief that they can succeed in a particular activity or course [26,27]. A student's self-efficacy in academic courses may be influenced by the classroom environment [28–30] and different teaching strategies [31–33]. In addition, self-efficacy has been shown to impact students' engagement, learning, and achievement in science courses [15,34–40]. For example,

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

students' higher self-efficacy was found to correlate with higher course achievement [34–37]. In addition, in other studies, students with higher self-efficacy were more likely to persist on more challenging tasks [38], less likely to second guess themselves and switch to an incorrect answer [39], and persist in STEM courses and careers [15,40]. In addition, self-efficacy is one of the factors that contributes to students' science identity. Self-efficacy contributing to a students' identity is important because identity in a particular domain has been shown to play an important role not only in students' in-class participation and performance but also in their choices of future courses and careers [41–46]. Additionally, identity frameworks show that identity is predicted by students' self-efficacy, interest, and recognition by others in physics [46–48].

Students with higher self-efficacy are more likely to embrace challenging goals, take on difficult tasks, and use effective learning strategies in their classes [49]. Additionally, students' self-efficacy in different STEM courses can lead to long-term effects by forcing them into a positive or negative feedback loop. When students complete short-term goals as they work towards completing their major (e.g., finishing introductory physics courses, etc.) they receive feedback on their performance. Students with high self-efficacy in a domain are less likely to have anxiety that can rob them of their cognitive resources while learning and test taking, since the working memory during problem solving has limited capacity [50]. They are also more likely to employ better academic strategies, such as goal setting, self-monitoring, and self-evaluation and their self-efficacy can be a predictor of their performance in that domain [49]. Thus, the domain specific self-efficacy has the potential to create a feedback loop with performance, e.g., when students with low or high self-efficacy start receiving feedback on graded assignments and exams in their classes. In particular, a student with greater self-efficacy may perform better (than a student with similar knowledge but lower self-efficacy), which can strengthen their self-efficacy further. On the other hand, a student with lower self-efficacy may perform worse, which can further weaken their self-efficacy. Since self-efficacy can increase students' persistence in science majors and careers [51] and a gender gap in self-efficacy has been observed in physics courses in which women are underrepresented [21,52–54], it is important to investigate whether there is a gender gap in physics self-efficacy in introductory physics courses in which women are the majority.

Even though women outnumber men in the introductory algebra-based physics courses for bioscience majors, pervasive societal stereotypes and biases about who belongs and can succeed in physics that women internalize from a young age can impact women's physics self-efficacy. One common stereotype is that one needs to be a genius or brilliant to succeed in physics [55]. Genius is often associated with boys [56], and girls from a young age tend

to shy away from fields associated with innate brilliance or genius [57]. Additionally, by high school, girls are less likely to believe that physics is for them and that their teacher thinks they are good at physics [58]. Furthermore, the environment in many science classrooms, including a pedagogy that favors male students' interests, a “chilly climate” for women, and a lack of female role models can negatively impact women [59]. All these factors can influence female students' perception of their ability to do physics before they enter the classroom. Therefore, although women are not underrepresented in algebra-based physics courses, these societal stereotypes can still impact their motivational beliefs in those courses.

The study presented here investigated gender difference in self-efficacy in introductory algebra-based physics courses in which women make up a majority of students in the class. In particular, we investigated the gender gap in self-efficacy at matched performance levels. It is also important to examine whether physics self-efficacy is a reflection of actual performance difference or whether the self-efficacy gender gap is due to pervasive societal stereotypes and biases female students must contend with throughout their lifetime. Below, we first delineate the research questions, then describe the methodology, present and discuss results, and finally conclude with instructional implications and future directions. The following research questions were answered by analyzing data from a validated survey administered to students in introductory algebra-based physics courses mainly for bioscience majors in which women outnumber men at a large public research university in the U.S.:

II. RESEARCH QUESTIONS

- RQ1** Are there gender differences in self-efficacy throughout a two-semester physics course sequence controlling for student performance?
- RQ2** How do women's and men's self-efficacy change from the beginning to the end of each physics course (physics 1 and physics 2)?
- RQ3** If there is a self-efficacy gender gap at the beginning of the first physics course, does the gap get bridged by the end of a year-long physics course sequence in which women outnumber men?
- RQ4** If there is a self-efficacy gender gap, is it a reflection of the performance differences by gender or is there a gender gap beyond what is predicted by course performance due to pervasive stereotypes and biases female students must contend with throughout their lifetime?

III. METHODOLOGY

A. Participants who were administered the survey

The data were collected using a validated motivational survey by us and administered at a large public research

university in the U.S. to students in a two-semester introductory algebra-based physics sequence in which women outnumber men. The survey was administered on scantrons in the recitation at the beginning (pre, during the first recitation) and end (post, in the last two weeks of classes) of physics 1 and physics 2. These two courses are generally taken in consecutive semesters, and in a normal sequence. We analyzed data across two consecutive academic years from courses that are traditionally taught. The classes are typically taken as a requirement by students on the pre-medical or pre-health track (mainly bioscience majors) primarily in their junior or senior year of undergraduate studies. Both of these are mandatory physics courses for all bioscience majors.

In this study, we analyzed 474 matched students (students who took the survey at both time points in physics 1 and physics 2). The university provided demographic information such as age, gender, and ethnicity or race information using an honest broker process by which the research team received the information without knowledge of the identities of the participants. From the university data, the matched participants were 34% male and 66% female students. Thus, female students significantly outnumbered male students in these introductory physics courses. We note that the data provided by the university only include binary options of male and female. We recognize that gender is not a binary construct. However, we are limited to the binary data here (less than 1% of the students did not provide this information and thus were not included in this study).

B. Instrument validity

This study focused on students' responses to the physics self-efficacy survey items (see Table I) [60–62] in the motivational survey, which included other items on other motivational constructs such as physics perceived recognition, interest, identity, etc. The items in the study were designed on a Likert scale of 1 (low endorsement) to 4 (high endorsement) [63]. A lower score was indicative of a negative endorsement of self-efficacy while a higher score was related to a positive endorsement. The items for different motivational beliefs including self-efficacy were adapted from previously validated surveys [21,60,62,64–66]. The questions on self-efficacy covered students' belief in their ability to succeed in the physics course [26,27]. Self-efficacy question Q1 was taken from the Peer Instruction self-efficacy instrument [62]. The other questions are adapted from Godwin *et al.*'s performance or competence questions with minor changes based upon individual interviews during the validation of the survey at our institution [66]. Specifically, Q2, Q3, and Q4 (see Table I) are the same or slightly modified questions from the survey used by Godwin *et al.* [66]. The items on the survey were used in our prior work and revalidated in our own context [21] using one-on-one interviews with students to ensure that students interpreted

TABLE I. Each self-efficacy survey item along with factor loadings from the confirmatory factor analysis for all students ($N = 474$) in physics 1 and physics 2. Cronbach's alpha for physics 1 is 0.78 and Cronbach's alpha for physics 2 is 0.82. The factor loading of each self-efficacy item to the self-efficacy construct (factor) is statistically significant to $p < 0.001$.

Survey question	Physics 1	Physics 2
Q1. I am able to help my classmates with physics in the laboratory or in recitation.	0.581	0.672
Q2. I understand concepts I have studied in physics.	0.769	0.760
Q3. If I study, I will do well on a physics test.	0.730	0.762
Q4. If I encounter a setback in a physics exam, I can overcome it	0.771	0.703

the survey items correctly, exploratory factor analysis (EFA), confirmatory factor analysis (CFA) [21], and Cronbach's alpha [67]. Interviews with students suggest that they interpreted the items correctly. Cronbach's alpha was used to measure the internal consistency of the items. Cronbach's alpha is 0.78 for physics 1 and 0.82 for physics 2, which is considered reasonable [67]. To ensure that items measured self-efficacy coherently and separately from other motivational constructs in the survey at the end of physics 1 and physics 2, an EFA was conducted. Additionally, we conducted a CFA with five motivational constructs in the broader survey. Since this study focuses only on the self-efficacy construct, the survey items and factor loadings for each self-efficacy item from the CFA are given in Table I.

In addition, we preformed item response theory (IRT) in order to check the response option distances for the survey constructs [68,69]. The parametric grades response model using STATA was used to test the measurement precision of our response scale. The parametric grades item response model calculates the location parameter for each response and calculates the difference between the locations. The response scale discrimination values were 1.54 and 2.08. The numerical values for the location differences need to be approximately similar, as they are for the response scale so we can use the means for these data. In addition, we estimated IRT-based domain scores. The correlation coefficient between the Likert mean scale values for self-efficacy and the IRT-based scores were calculated to be 0.98. Since the IRT-based domain score is so highly correlated with the mean score, using mean values for the scores is acceptable when analyzing the data and we can use parametric tests on Likert-scale data [70,71]. In particular, because the psychological distance between adjacent response items and across items was approximately similar and complex factor scores derived from IRT or CFA are so highly correlated with the mean scores, it is reasonable to use mean scores [68,72].

TABLE II. The distribution of students' final grade (which includes women and men) and percentage of female students in each grade bin in physics 1 and physics 2.

Course	Total	D	C	B	A
Physics 1	$N = 474$ 66% female	$N = 1$ 100% female	$N = 52$ 71% female	$N = 195$ 73% female	$N = 226$ 58% female
Physics 2	$N = 474$ 66% female	$N = 16$ 63% female	$N = 94$ 72% female	$N = 168$ 70% female	$N = 196$ 60% female

C. Analysis

We obtained students' final grades in each course from the university and linked those to students' gender and self-efficacy via the honest broker process discussed earlier. For analysis, we grouped students into bins by grade ranges (i) C− and below (considered insufficient to move on to the next course for this mandatory course for bioscience majors); (ii) C, C+; (iii) B−, B, B+ (iv) A−, A. For convenience, we labeled these bins D (e.g., this refers to C− and below), C, B, and A, respectively. Table II shows the distribution of students' final grades and the percentages of female students in each grade bin.

We first compared female and male students' mean self-efficacy scores by grade bins for statistical significance using t tests and for calculating the effect size using Cohen's d [73]. In this analysis, we focused on self-efficacy differences of students who passed the course (C or above) since there were few students in the lowest grade bin. Next, to determine whether there are gender differences in self-efficacy controlling for grade (see Table III), we performed a linear regression in which the dependent variable was students' average post-self-efficacy score in physics 1 or physics 2 and the independent variables were students' gender and grade at the end of the course [74].

IV. RESULTS AND DISCUSSION

A. Gender differences in self-efficacy by grade

Regarding RQ1 for physics 1, we find women have lower average self-efficacy scores than men in most of the

TABLE III. Self-efficacy scores in physics 1 by grade bin for female and male students. N = number of students, Pre = mean pre-self-efficacy score, Post = mean post-self-efficacy score. Cohen's d for the gender differences in pre- and postscores for each grade bin is shown (all statistically significant gender differences have been bolded).

Physics 1 self-efficacy by grade									
	C			B			A		
	Pre	Post	N	Pre	Post	N	Pre	Post	N
Men	2.78	2.65	15	3.07	2.92	52	3.09	3.10	95
Women	2.79	2.43	37	2.76	2.58	143	2.90	2.96	131
Cohen's d	−0.01	0.36		0.65	0.60		0.40	0.30	
p value	0.966	0.252		<0.001	<0.001		0.003	0.027	

grade bins (A, B, C) at the beginning and end of the course (see Table III). Table III shows that the differences are statistically significant for students in the A and B bins at the beginning and end of the course. However, in regard to RQ2 for physics 1, the average self-efficacy scores of both men and women decreased for those students who received B's or C's in the physics 1 course. The differences between men and women range from a large effect size (greater than 0.50) to a medium effect size (greater than 0.30). Furthermore, Table III shows that women receiving A's have about the same self-efficacy as men who receive between a C and a B at the beginning of the course in physics 1. However, by the end of the course, women with A's have roughly the same self-efficacy as men with B's (Table III). Table III also shows that women and men's average self-efficacy scores were stable (did not significantly change over the course of the semester) for the students who received A's.

Regarding RQ1 for physics 2, the gender gap in self-efficacy controlling for grade continued. Table IV shows that women had lower self-efficacy scores than men in all grade bins at the beginning and at the end of physics 2, but the self-efficacy gender difference was not statistically significant in the group of students who received a C grade. However, by the end of the course in physics 2, the only statistically significant difference occurred for women and men in the A range with a medium to large effect size (Table IV). This finding differs from a similar study in the calculus-based introductory physics courses

TABLE IV. Self-efficacy scores in physics 2 by grade bin for female and male students. N = number of students, Pre = mean pre-self-efficacy score, Post = mean post-self-efficacy score. Cohen's d for the gender differences in pre- and postscores for each grade bin is shown (all statistically significant gender differences have been bolded).

Physics 2 self-efficacy by grade									
	C			B			A		
	Pre	Post	N	Pre	Post	N	Pre	Post	N
Men	2.88	2.64	26	3.00	2.74	51	3.10	3.10	79
Women	2.69	2.50	68	2.73	2.68	117	2.89	2.88	117
Cohen's d	0.40	0.27		0.56	0.10		0.51	0.50	
P value	0.085	0.244		<0.001	0.561		<0.001	<0.001	

(in which women are severely underrepresented) in which the physics self-efficacy gender gap increased from the beginning to the end of the course [52].

In regard to RQ2 for physics 2, at the beginning of the course, women who receive A's have the same self-efficacy as men who receive C's (Table IV). However, by the end of the course, women with A's have the same self-efficacy as men with grades between a B and an A (see Table IV). Again, women's and men's average self-efficacy scores did not significantly change over the course of the semester for the students who received A's in the physics 2 course. Also, the average self-efficacy scores of both men and women decreased for those students who received B's or C's in the course.

Regarding RQ3, Tables III and IV show that the self-efficacy gender gap did not get bridged by the end of a year-long course sequence. Data from both physics 1 and physics 2 show that there is some protective effect in students' self-efficacy for men and women who receive A's in both courses. However, the gender gap is maintained even among students who received an A grade and the self-efficacy scores decreased for students in the B and C bins. These findings suggest that there may be a need to make the learning environment equitable and inclusive. It is important to note that women had lower self-efficacy scores than men in many of the grade bins (as evidenced by both the pre- and post-test scores) even though they outnumber men in the course. Prior interviews with students suggest that some women in physics classes would get test grades comparable to their peers in the class, but they still felt that they were not good at physics [75]. We believe that this disparity may be due to the deep-rooted societal stereotypes and biases about who can excel in physics that women must contend with from an early age, starting long before they step into the college physics classroom [59].

Lastly, in regard to RQ4, we investigated the fraction of the self-efficacy gender gap that was related to students' course performance versus students' biased perceptions. First, we found the raw post-self-efficacy differences between women and men to be 0.28 ($p < 0.001$) in physics 1 and 0.20 ($p < 0.001$) in physics 2 in favor of men. Then we controlled for students' grade in each course with multiple linear regression [74] with post-self-efficacy as the dependent variable and gender and grades as the independent variables, including an interaction effect between grade and gender [see Eq. (1)]. The interaction effect (β_3) between gender and grade was not a significant predictor for self-efficacy in either physics 1 ($p = 0.221$) or physics 2 ($p = 0.299$) and thus was omitted in future analyses.

$$\begin{aligned} \text{post} - \text{self} - \text{efficacy} = & \beta_1(\text{gender}) + \beta_2(\text{grade}) \\ & + \beta_3(\text{grade} \times \text{gender}) + \text{const.} \end{aligned} \quad (1)$$

In physics 1, a significant overall regression correlation was found [$F(2, 471) = 47.96$, $p < 0.001$] for post-self-efficacy scores with main effects from both gender ($p < 0.001$) and final grade ($p < 0.001$) [73]. The regression equation is shown in Eq. (2). In physics 2, a significant overall regression correlation was found [$F(2, 471) = 34.78$, $p < 0.001$] for post-self-efficacy scores with main effects from gender ($p = 0.003$) and final grade ($p < 0.001$). The regression equation for this case is shown in Eq. (3). Gender was coded 0 for female students, or 1 for male students and grade was reported continuously from 1 to 4. Both gender and grade were significant predictors of students' average self-efficacy in both physics 1 and physics 2 even though women make up a numerical majority in the course.

$$\begin{aligned} \text{post} - \text{self} - \text{efficacy in physics 1} \\ = 1.67 + 0.21(\text{gender}) + 0.32(\text{grade}), \end{aligned} \quad (2)$$

$$\begin{aligned} \text{post} - \text{self} - \text{efficacy in physics 2} \\ = 2.00 + 0.15(\text{gender}) + 0.23(\text{grade}). \end{aligned} \quad (3)$$

Compared to the raw post-self-efficacy differences, the mean values of the gender gap in self-efficacy (from the multiple linear regression) decreased slightly to 0.21 ($p < 0.001$) in physics 1 [see Eq. (2)] and 0.15 ($p = 0.003$) in physics 2 [see Eq. (3)]. Therefore, while a small portion of the gap in self-efficacy can be attributed to performance in the course, the self-efficacy gender gap mainly comes from women's biased perceptions of how well they can do in physics: 75% in both physics 1 and physics 2 (see Fig. 1). Thus, similar to earlier investigation in calculus-based courses in which women are severely underrepresented [52], even in these algebra-based courses where women make up the majority of the students, biased beliefs about who can excel in physics that are pervasive in our society still affects their self-efficacy.

FIG. 1. Relative contribution of performance and biased perception to the post-self-efficacy gender gap in physics 1 and physics 2 for bioscience majors in which women outnumber men.

IV. IMPLICATIONS AND FUTURE DIRECTIONS

The physics self-efficacy gender gap we found at equal grade levels may seem unexpected since women are not underrepresented in these introductory algebra-based physics courses mainly taken by the bioscience majors (for whom both courses are mandatory). However, we hypothesize that the societal stereotypes and biases about who can excel in physics that women from a young age are bombarded with are at the heart of the gender gap in physics self-efficacy in these courses in which women outnumber men. Not only are female students exposed to gendered views of physics in informal situations growing up, e.g., from shows such as *The Big Bang Theory* or from their family and friends, they are also treated differently and given differential advice by K-12 educators and guidance counselors [76].

Our finding is important and this issue needs to be addressed since self-efficacy has been shown to affect students' persistence, learning and long-term outcomes [49]. Students with higher self-efficacy are more likely to engage with challenging problems without anxiety and self-efficacy plays an important role in predicting student performance in courses in that domain [49]. Low self-efficacy can create a negative feedback loop when students receive feedback via graded assignments in their classes. In particular, a student's lower self-efficacy can cause anxiety and deteriorate their performance, which can then weaken their self-efficacy further.

However, there appears to be some benefit to having a large cohort of women in these introductory algebra-based physics courses for bioscience majors. We find that the percentage of the gender gap attributed to biased perception does not change significantly from physics 1 to physics 2 (Fig. 1). However, the percentage of the gender gap in introductory calculus-based physics courses (in which women are severely underrepresented) attributed to biased perception started out lower in physics 1 and gets worse in physics 2 [52]. Thus, we observed a somewhat protective effect in women's self-efficacy when they make up about two-thirds of the students in these courses. Nevertheless, since women still have lower self-efficacy than men in these courses controlling for grade, there is urgent need to create an equitable and inclusive learning environment to bridge the self-efficacy gender gap.

In particular, equitable and inclusive pedagogical strategies are necessary to bridge the gender gap in self-efficacy and

raise the self-efficacy of all students. Active learning pedagogies could decrease women's self-efficacy if not implemented with equity and inclusion as central constructs. In one study in engineering, women working in mixed gender cooperative groups tended to be undermined by their male classmates and played a less active role in those groups [77]. Additionally, in a study in physics, students in courses that utilized active engagement strategies performed better than students in lecture-based courses, however, the gender gap in performance on conceptual surveys persisted and even became worse in some cases [78]. In relation to self-efficacy, some interactive engagement physics courses have been shown to decrease women's self-efficacy [32] whereas others have been shown to improve women's self-efficacy [79]. Therefore, it is important for the instructor or TA to ensure that all students feel they belong and have high self-efficacy and can contribute equally. One strategy physics instructors can use to encourage equal contribution in a group is to assign each student a role that rotates throughout the course. However, this approach alone is unlikely to be effective in bridging the self-efficacy gender gap if the learning environment is otherwise not equitable and inclusive.

A field-tested strategy to make the learning environment equitable and inclusive is through classroom interventions [9,80,81]. In particular, social-psychological classroom interventions, e.g., sense of belonging and mindset interventions, have been shown to eliminate gender performance gaps [8,81] and to have lasting effects beyond the time they are implemented [80]. At the same time, these interventions can help students develop positive feelings of being recognized by their peers, TAs, and instructors. We have implemented a short ecological belonging intervention in the calculus-based physics classes that eliminated the gender gap in physics performance [81]. These types of interventions could be adapted and their effectiveness investigated for the algebra-based physics courses discussed here.

ACKNOWLEDGMENTS

This work was supported by Grant NSF DUE-152457. APC charges for this article were fully paid by the University Library System, University of Pittsburgh. We thank all students who participated in this research and Dr. Robert Devaty for his constructive feedback on this manuscript.

- [1] Science & Engineering Degree Attainment: 2004-2014, (National Student Clearinghouse Research Center) <https://nscresearchcenter.org/snapshotreport-degreeattainment15/>.
- [2] Percent of Physics Bachelors and PhDs Earned by Women, Classes of 1975 through 2016. American Institute of Physics, <https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016>.
- [3] E. Seymour, Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, *Sci. Educ.* **86**, 79 (2002).
- [4] J. Blue, A. L. Traxler, and X. C. Cid, Gender matters, *Phys. Today* **71**, 40 (2018).
- [5] S. G. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, *J. Eng. Educ.* **87**, 369 (1998).
- [6] B. L. Whitten, S. Foster, and M. Duncombe, What works for women in undergraduate physics?, *Phys. Today* **56**, 46 (2003).
- [7] L. Tsui, Effective strategies to increase diversity in STEM fields: A review of the research literature, *J. Negro Educ.* **76**, 555 (2007).
- [8] G. M. Walton, C. Logel, J. M. Peach, S. J. Spencer, and M. P. Zanna, Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering, *J. Educ. Psychol.* **107**, 468 (2015).
- [9] J. M. Harackiewicz, E. A. Canning, Y. Tippetts, S. J. Priniski, and J. S. Hyde, Closing achievement gaps with a utility-value intervention: Disentangling race and social class, *J. Pers. Soc. Psychol.* **111**, 745 (2016).
- [10] E. Seymour, N. M. Hewitt, and C. M. Friend, *Talking About Leaving: Why Undergraduates Leave the Sciences* (Westview Press, Boulder, CO, 1997), Vol. 12.
- [11] S. Tobias, *They’re Not Dumb, They’re Different: Stalking the Second Tier* (Research Corporation, Tucson, AZ, 1990).
- [12] S. J. Correll, Gender and the career choice process: The role of biased self-assessments, *Am. J. Sociology* **106**, 1691 (2001).
- [13] S. J. Correll, Constraints into preferences: Gender, status, and emerging career aspirations, *Am. Soc. Rev.* **69**, 93 (2004).
- [14] 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).
- [15] 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).
- [16] S. L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, *J. Res. Sci. Teach.* **45**, 955 (2008).
- [17] C. P. Benbow and L. L. Minor, Mathematically talented males and females and achievement in the high school sciences, *Am. Educ. Res. J.* **23**, 425 (1986).
- [18] J. E. Stets, P. S. Brenner, P. J. Burke, and R. T. Serpe, The science identity and entering a science occupation, *Soc. Sci. Res.* **64**, 1 (2017).
- [19] M. M. Chemers, E. L. Zurbriggen, M. Syed, B. K. Goza, and S. Bearman, The role of efficacy and identity in science career commitment among underrepresented minority students, *Journal Soc. Issues* **67**, 469 (2011).
- [20] K. A. Robinson, T. Perez, J. H. Carmel, and L. Linnenbrink-Garcia, Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit, *Contemp. Educ. Psychol.* **56**, 180 (2019).
- [21] 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).
- [22] 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).
- [23] J. Stewart, R. Henderson, L. Michaluk, J. Deshler, E. Fuller, and K. Rambo-Hernandez, Using the social cognitive theory framework to chart gender differences in the developmental trajectory of STEM self-efficacy in science and engineering students, *J. Sci. Educ. Technol.* **29**, 758 (2020).
- [24] A. Traxler, R. Henderson, J. Stewart, G. Stewart, A. Papak, and R. Lindell, Gender fairness within the force concept inventory, *Phys. Rev. Phys. Educ. Res.* **14**, 010103 (2018).
- [25] A. Madsen, S. B. McKagan, and E. C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, *Phys. Rev. Phys. Educ. Res.* **9**, 020121 (2013).
- [26] A. Bandura, in *Encyclopedia of Psychology*, edited by R. J. Corsini (Wiley, New York, 1994), pp. 368–369.
- [27] A. Bandura, Self-efficacy: Toward a unifying theory of behavioral change, *Psychol. Rev.* **84**, 191 (1977).
- [28] D. H. Schunk and F. Pajares, in *Development of Achievement Motivation: A Volume in the Educational Psychology Series*, edited by A. Wigfield and J. S. Eccles (Academic Press, San Diego, 2002), pp. 15–31.
- [29] R. Dou, E. Brewé, G. Potvin, J. P. Zwolak, and Z. Hazari, Understanding the development of interest and self-efficacy in active-learning undergraduate physics courses, *Int. J. Sci. Educ.* **40**, 1587 (2018).
- [30] S. L. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, *J. Res. Sci. Teach.* **43**, 485 (2006).
- [31] H. Fencel and K. Scheel, Engaging students—an examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course, *J. Coll. Sci. Teach.* **35**, 20 (2005), https://my.nsta.org/resource/?id=10.2505/4/jcst05_035_01_20.
- [32] J. M. Nissen and J. T. Shemwell, Gender, experience, and self-efficacy in introductory physics, *Phys. Rev. Phys. Educ. Res.* **12**, 020105 (2016).
- [33] J. M. Bailey, D. Lombardi, J. R. Cordova, and G. M. Sinatra, Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy, *Phys. Rev. Phys. Educ. Res.* **13**, 020140 (2017).
- [34] V. Sawtelle, E. Brewé, and L. H. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, *J. Res. Sci. Teach.* **49**, 1096 (2012).

- [35] A. M. L. Cavallo, W. H. Potter, and M. Rozman, Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors, *School Sci. Math.* **104**, 288 (2004).
- [36] P. Vincent-Ruz and C. D. Schunn, The increasingly important role of science competency beliefs for science learning in girls, *J. Res. Sci. Teach.* **54**, 790 (2017).
- [37] P. Vincent-Ruz, K. Binning, C. D. Schunn, and J. Grabowski, The effect of math SAT on women's chemistry competency beliefs, *Chem. Educ. Res. Pract.* **19**, 342 (2018).
- [38] T. Bouffard-Bouchard, S. Parent, and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students, *Int. J. Behav. Dev.* **14**, 153 (1991).
- [39] K. Miller, J. Schell, A. Ho, B. Lukoff, and E. Mazur, Response switching and self-efficacy in peer instruction classrooms, *Phys. Rev. Phys. Educ. Res.* **11**, 010104 (2015).
- [40] S. Cheryan, S. A. Ziegler, A. K. Montoya, and L. Jiang, Why are some STEM fields more gender balanced than others?, *Psychol. Bull.* **143**, 1 (2017).
- [41] 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).
- [42] J. P. Gee, Chapter 3: Identity as an analytic lens for research in education, *Rev. Res. Educ.* **25**, 99 (2000).
- [43] K. L. Tonso, Student engineers and engineer identity: Campus engineer identities as figured world, *Cult. Stud. Sci. Educ.* **1**, 273 (2006).
- [44] P. Vincent-Ruz and C. D. Schunn, The nature of science identity and its role as the driver of student choices, *Int. J. STEM Educ.* **5**, 48 (2018).
- [45] R. Dou, H. Cian, and V. Espinosa-Suarez, Undergraduate STEM majors on and off the pre-med/health track: A STEM identity perspective, *CBE Life Sci. Educ.* **20**, 1 (2021).
- [46] 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).
- [47] 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).
- [48] 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).
- [49] B. J. Zimmerman, Self-efficacy: An essential motive to learn, *Contemp. Educ. Psychol.* **25**, 82 (2000).
- [50] E. A. Maloney, J. R. Sattizahn, and S. L. Beilock, Anxiety and cognition, *Wiley Interdisc. Rev.: Cog. Sci.* **5**, 403 (2014).
- [51] J. A. Raelin, M. B. Bailey, J. Hamann, L. K. Pendleton, R. Reisberg, and D. L. Whitman, The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention, *J. Eng. Educ.* **103**, 599 (2014).
- [52] 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).
- [53] C. Lindstrøm and M. D. Sharma, Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter?, *Int. J. Innov. Sci. Math Educ.* **19**, 1 (2011).
- [54] R. Henderson, V. Sawtelle, and J. M. Nissen, Gender & self-efficacy: A call to physics educators, *Phys. Teach.* **58**, 345 (2020).
- [55] S.-J. Leslie, A. Cimpian, M. Meyer, and E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines, *Science* **347**, 262 (2015).
- [56] S. Upson and L. F. Friedman, Where are all the female geniuses?, *Sci. Am. Mind* **23**, 63 (2012).
- [57] L. Bian, S.-J. Leslie, and A. Cimpian, Gender stereotypes about intellectual ability emerge early and influence children's interests, *Science* **355**, 389 (2017).
- [58] 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).
- [59] J. C. Blickenstaff, Women and science careers: Leaky pipeline or gender filter?, *Gender Educ.* **17**, 369 (2005).
- [60] 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).
- [61] Activation lab tools: Measures and data collection instruments, <http://www.activationlab.org/tools/>.
- [62] J. Schell and B. Lukoff, Peer Instruction Self-Efficacy Instrument [developed at Harvard University] (to be published) (2010).
- [63] R. Likert, A technique for the measurement of attitudes, *Arch. Psychol.* **22**, 55 (1932).
- [64] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, *Am. J. Phys.* **66**, 212 (1998).
- [65] PERTS academic mindsets assessment, <https://survey.perts.net/share/dlmooc>.
- [66] 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).
- [67] L. J. Cronbach, Coefficient alpha and the internal structure of tests, *Psychometrika* **16**, 297 (1951).
- [68] S. E. Embretson and S. P. Reise, *Item Response Theory for Psychologists* (Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 2000).
- [69] https://www.stata.com/meeting/australia15/abstracts/materials/oceania15_rosier.pdf.
- [70] G. Norman, Likert scales, levels of measurement and the "laws" of statistics, *Adv. Health Sci. Educ.* **15**, 625 (2010).
- [71] J. De Winter and D. Dodou, Five-point likert items: T test versus Mann-Whitney-Wilcoxon (addendum added october 2012), *Pract. Assess. Res. Eval.* **15**, 11 (2010).

- [72] F. Samejima, Estimation of latent ability using a response pattern of graded scores, *Psychometrika Monograph Supplement* **1968**, 17 (1969), <http://www.psychometrika.org/journal/online/MN17.pdf>.
- [73] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Routledge, London, 2013).
- [74] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Routledge, London, 2013).
- [75] 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).
- [76] D. M. Welsch and M. Winden, Student gender, counselor gender, and college advice, *Educ. Econ.* **27**, 112 (2019).
- [77] 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).
- [78] 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).
- [79] T. Espinosa, K. Miller, I. Araujo, and E. Mazur, Reducing the gender gap in students' physics self-efficacy in a team- and project-based introductory physics class, *Phys. Rev. Phys. Educ. Res.* **15**, 010132 (2019).
- [80] D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They're not magic, *Rev. Educ. Res.* **81**, 267 (2011).
- [81] K. R. Binning, N. Kaufmann, E. M. McGreevy, O. Fotuhi, S. Chen, E. Marshman, Z. Yasemin Kalender, L. Limeri, L. Betancur, and C. Singh, Changing social norms to foster the benefits of collaboration in diverse workgroups, *Psychol. Sci.* **31**, 1059 (2020).