

Reducing the gender gap in students' physics self-efficacy in a team- and project-based introductory physics class

Tobias Espinosa,¹ Kelly Miller,² Ives Araujo,³ and Eric Mazur²

¹*Mathematics Statistics and Physics Institute, Universidade Federal do Rio Grande, Santo Antônio da Patrulha-RS, 125 Barão do Cahy St., 95500-000, Brazil*

²*Department of Physics and Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

³*Physics Institute, Universidade Federal do Rio Grande do Sul, 9500 Bento Gonçalves Avenue, Porto Alegre-RS, 91501-970, Brazil*



(Received 18 July 2018; published 28 May 2019)

Self-efficacy represents an individual's belief that he or she can perform a particular task in a given domain. It is a strong predictor for performance and persistence in STEM education. Research shows that there is a large and persistent gender gap in student self-efficacy in STEM academic disciplines. In some cases, active teaching strategies have been shown to positively affect physics self-efficacy but it is unclear how this impact differs between men and women. In this study, we investigate the impact of a physics class taught with active teaching strategies on students' self-reported physics self-efficacy and how this impact varies across gender lines. We measured the change in physics self-efficacy over four different dimensions; conceptual understanding, problem-solving, lab and hands-on activities, and collaborative work. We report three main findings. First, the initial physics self-efficacy gender gap disappeared by the end of the semester. Second, female students' self-efficacy improved significantly, whereas there was no significant change in the self-efficacy of the male students. Third, the gender gaps in the conceptual understanding and problem-solving dimensions in particular, were significantly reduced. This study represents an initial step towards understanding the influence of active teaching strategies can have on reducing the self-efficacy gender gap.

DOI: [10.1103/PhysRevPhysEducRes.15.010132](https://doi.org/10.1103/PhysRevPhysEducRes.15.010132)

I. INTRODUCTION

In 1977 Bandura introduced the concept of self-efficacy— one's belief in their ability to complete a specific task (or a set of tasks) in a given domain [1]. Self-efficacy is a strong predictor for performance in science courses and has been shown to be related to students' persistence, resilience, and career choices in STEM [2–6]. Self-efficacy also influences a number of factors that are relevant to learning in an active environment, such as perseverance and self-regulation¹ [7,8].

Bandura theorized that an individual develops his or her own self-efficacy for a specific task through social and personal experiences [1]. These experiences fall into four categories: mastery experiences, vicarious learning experiences, social persuasion experiences, and an individual's physiological and affective state [7,9]. Mastery

experiences are recurrent episodes of success or failure experienced by the student through active participation. Vicarious experiences are observations of others (specifically peers or role models) performing a task. Social persuasion includes the verbal and nonverbal judgment of others (peers and teachers). Finally, a student's physiological and affective state refers to their mood, stress, anxiety, etc. as it emerges during performance of a task. Students draw on all four of these categories as sources of information in building their self-efficacy [9–12].

It is important that students have accurate judgments about their own capabilities. Overconfident students, who believe they can perform activities beyond their reach, can suffer unnecessary failures that can have damaging effects on their self-efficacy. Underestimating their own capabilities can also have negative consequences. Students who underestimate their self-efficacy often restrict their activities and create internal obstacles to their own performance [13].

It is desirable that all students have physics self-efficacy levels consistent with their real abilities. However, research has shown that women start introductory physics courses with lower physics self-efficacy than men, even when both achieve similar performances [14]. This gender gap increases over the course of the semester of a traditionally taught course [15–19]. This gap in self-efficacy is

¹Bandura defines self-regulation as the “exercise of influence over one's own motivation, thought processes, emotional states, and patterns of behavior” [7], p. 2].

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.

problematic given that females are underrepresented in science and technology-related courses. In the United States, for example, the percentage of degrees in the physical sciences earned by women is only 38% at the bachelor's level and 34% at the doctorate level [20]. The percentage of women at all levels of academic appointments in the physical sciences is even lower (20%) [21]. This gender gap is not unique to the United States; it is prevalent, to varying degrees, throughout the world [21]. Given that self-efficacy is a strong predictor of success and perseverance in physics and other STEM-related fields [11], closing the self-efficacy gender gap is important in addressing the gender gap in STEM. However, increasing women's self-efficacy has limitations in addressing the larger issues of the underrepresentation of women in STEM. There are broader cultural issues in STEM, e.g., sexism and gender microaggressions, that also have to be worked on [22].

For women, the development of self-efficacy is influenced more by vicarious experiences and social persuasions than for men [10,23–25]. Mastery experiences, on the other hand, are the most significant source of self-efficacy for male students [24,25]. Another gender difference is shown to occur in students' physiological and affective states. For example, research on test anxiety has shown female students report significantly higher levels of test anxiety in math and physics than male students [26,27].

Many studies have shown how teaching strategies can help reduce the gender gap in science achievement [28–39]. Some studies have shown that female students, in particular, benefit from the use of active teaching strategies [35,36,40]. Female students tend to perform better when they have the opportunity to verbally articulate their thoughts [34]. Teaching environments that focus on activities that enhance cooperation and communication between students and between students and instructors have been shown to decrease the gender achievement gap [30,33]. Pedagogies that focus on decreasing competitiveness, increasing understanding, and providing frequent and diverse assessment opportunities have been shown to help narrow the gender gap [28–30,37–39].

Students' physics self-efficacy decreases in traditional physics courses [16,41]; in actively taught physics courses it either decreases [41] or remains essentially unchanged [16]. Furthermore, some studies have shown that actively taught physics courses are detrimental to the self-efficacy of female students [19].

The purpose of this paper is to investigate physics self-efficacy for male and female students in a team- and project-based introductory physics course.² We measured students' physics self-efficacy across four different dimensions

(conceptual physics understanding, problem solving, collaborative work, lab and hands-on activities) at the beginning and the end of two semesters of this course. We find that while there is no significant change in the self-efficacy of male students, the female students' self-efficacy improves significantly, closing the initial gender gap by the end of both semesters. We also observe significant gender differences in the change in self-efficacy in each of the four dimensions.

II. METHODS

We administered the physics self-efficacy surveys from Appendix in two semesters of Applied Physics 50a (AP50a), a team- and project-based introductory physics course for engineering students at Harvard University. We collected data during two consecutive implementations of the same course (Fall 2016: $N = 65$, 37 female, 28 male; Fall 2017: $N = 39$, 25 female, 14 male). The population was 48%–50% premedical students and 50%–52% engineering students, and the students were evenly distributed among sophomore, junior, and senior years.

A. Pedagogy

The pedagogy in AP50a draws on features from both project-based learning [42] and team-based learning [43]. During project-based learning students work for an extended period of time on an inquiry-driven project, often inspired by a real-world problem. By researching and problem solving, students gain knowledge and skills in specific content areas. All of the learning goals for AP50a are addressed through three projects per semester that students work on in class, as part of a team of 4–5 students. Students are not allowed to form their own teams. The instructor creates the teams to ensure that they are well balanced and diverse with respect to a number of student characteristics (incoming physics knowledge, gender, college major, year in college, prior experience with building). Teams are constructed so that there is an equal balance of male and female students. Students are not part of the same team for the entire semester, they are reorganized for each of the three projects during the semester. Students work in these teams on all aspects of the course, including assessments, which have both an individual and team component.

The in-class structure of AP50a consists of a blend of five different types of activities, each of which is designed to help students acquire skills and learn content that is essential to the projects. Each class consists of anywhere from 1 to 3 of these activities with the more highly scaffolded activities (for example, Peer Instruction) at the beginning of the class. As there are no lectures in class, the content delivery part of the course takes place via an online preclass reading assignment that the students are required to complete before each class.

²This study is not inclusive of nonbinary gender identities. We are using the language of the registered data that we have, which is consistent with that used in most physics educational research.

B. In-class activities

Peer Instruction: over the course of each semester, the instructor conducts 8 Peer Instruction sessions, each of which is between 1.5 and 2 h long. During each session students answer 8–12 ConcepTests on difficult concepts selected from the preclass reading assignment. ConcepTests are short conceptual questions that focus on a single topic [44]. Students answer individually initially and, after discussing each question with their team, they answer again. These sessions occur at the beginning of the class as they allow the instructor to probe students' understanding of the reading and address difficult concepts.

Tutorials.—worksheets that are designed to address common misconceptions about the course content. Depending on the semester, we use 6–8 tutorials from the “Tutorials in Introductory Physics” [45] developed by the Physics Education Group at the University of Washington. During this activity, which lasts 1 h, students work with their team on the worksheet and this allows them to explore their thinking about the more difficult concepts in the material.

Estimation activity.—students use their physics knowledge and reasoning skills to estimate five quantities related to the content of the class. Students are given 30 min to think and work with their team to estimate the quantities to the nearest order of magnitude.

Experimental design activity.—Hands-on, lablike activities or online simulations (typically PhET) [46] that help students develop experimental and analytical skills that are important for the projects.

Problem set reflection.—problem sets are comprised of 4–5 physics problems that students are given a week to work on at home. Students are instructed to give the problems their best effort (without consulting others) before coming to class and to bring their solutions to class.

C. Assessment

The emphasis in AP50a is on continuous, low-stakes, formative assessment—there are no high-stakes, summative assessments. In particular, traditional exams are replaced by *Readiness Assurance Activities* (RAA) which occur at the end of each of the 5 learning units and are designed to help ensure students master the material in each unit. During an RAA, students first work individually to solve a set of complex physics problems. They are free to consult textbooks or the internet (the problems on the RAA are not available on the internet), but are not allowed to discuss the problems with others. After submitting their individual responses via an online system, students work together to solve the same set of problems, submitting their responses as a team. As the team submits responses, the system provides immediate feedback. If the response is incorrect, teams can resubmit responses for reduced credit up to 3 times before the system reveals the correct answer with a detailed explanation. The overall score each student receives is the average of that student's individual score and the team score. This form

of assessment provides a low-stakes testing environment in which students learn collaboratively and receive immediate feedback.

D. Self-efficacy surveys

For the purpose of this study we developed the Physics Self-Efficacy Survey (PSES, Appendix). The PSES is adapted, mostly, from the Source of Self-Efficacy in Science Courses (SOSESC) survey [47,48]. We administered the PSES twice during each semester of the study, once as a presemester survey at the start of the semester and again as a postsemester survey at the end of the semester. The survey asks students to rank, on a ten-point scale, their degree of confidence in their ability to perform specific physics-related tasks. On this scale, a 1 represents “highly certain cannot do,” 5 is “moderately certain can do,” and 10 means “highly certain can do.” According to Bandura [49], self-efficacy scales that use only a few points (e.g., a Likert five-point scale), as well as negative numbers, should be avoided because they are less sensitive and less reliable. Individuals usually avoid the extreme positions, therefore including too few points loses differentiating information. Bandura emphasizes that a ten-point survey scale is stronger than a five-point scale.

We designed the PSES to measure student self-efficacy across the conceptual physics understanding (CPU), problem solving (PS), collaborative work (CW), and lab/hands-on activities (LHA) dimensions. The PSES consists of 20 items, 5 items for each of the four dimensions. In Appendix, we have indicated which questions on the PSES survey pertain to each of the four dimensions of physics self-efficacy. In the lab and hands-on activities section, for example, students are asked to rate their belief in their ability to design physics experiments using materials and use equipment during hands-on activities. In the collaborative work part of the survey, students are asked to rate their belief in their ability to work together with classmates to complete a complex task and encourage classmates to participate in discussions. Other self-efficacy instruments used in the literature (the SOSESC, for example) focus on the conceptual understanding and problem-solving dimensions and neglect to measure students' self-efficacy in the collaborative work and lab and hands-on dimensions. Given that this class emphasizes teamwork and hands-on activities, we were particularly interested in measuring students' self-efficacy in both collaborative work and lab and hands-on activities.

E. Validation of the physics self-efficacy survey

We created one self-efficacy survey based on existing, validated instruments. We established content validity by asking four physics education research experts to review and provide feedback on all survey items. We verified internal consistency by calculating Cronbach's alpha for the overall survey and for each of the four domains measured

TABLE I. Previous items gathered from published physics self-efficacy surveys and the new ones.

Validated instrument	Dimension	Item
College Chemistry Self-efficacy Scale (CCSS), Uzuntiryaki and İpa Aydın (2009)	CPU	11, 12
	LHA	8, 9, 14, 20
Sources of Self-Efficacy in Science Courses—Physics (SOSESC-P), Fencil and Scheel (2004)	PS	5*, 7, 13
	CW	4
New items	CPU	1, 2, 3
	PS	16, 19
	CW	6, 15, 17, 18
	LHA	3

by the PSES survey. In his *Guide for Constructing Self-Efficacy Scales* [49], (p. 307), Bandura recommends that researchers build their own self-efficacy instrument, because “there is no all-purpose measure of perceived self-efficacy. [...] Scales of perceived self-efficacy must be tailored to the particular domain of functioning that is the object of interest.”

Table I summarizes the PSES items that were adapted from other, validated self-efficacy instruments as well as the new items created to fit the specific AP50 context.

The PSES survey was designed with the specific features of AP50 in mind. The activities, teamwork, and projects in AP50 can be grouped into the four dimensions described. Table II shows how the relevant aspects of AP50 pertain to each of the four dimensions and the corresponding PSES survey items.

To establish content validity, four experts were asked to review both instruments (PSES and SPSES). The experts consisted of four physics professors and researchers in physics education, three of whom were experts in both self-efficacy and active learning environments, and one who was an experienced survey designer. This meets the standard consensus in the literature regarding the number of experts considered sufficient for content validity [50,51]. The four experts were asked to review the first version of the surveys. More specifically to (a) associate each item with the instruments’ dimension and objective, (b) comment on the appropriateness of the objectives, (c) criticize the items, (d) comment on both the format and phrasing of the statements, and (e) suggest improvements. Feedback from the experts was then used to revise the first version of the surveys and many of the original items were modified or removed to create the final version of each survey. A limitation to our validation procedure is that we cannot be sure that students are interpreting the

TABLE II. Mapping of survey items onto the AP50 features, organized into the four dimensions of physics self-efficacy.

Dimension	AP50 features	Survey items
CPU	<ul style="list-style-type: none"> • Preclass reading • Peer Instruction • Tutorials • Project-based learning approach 	<ul style="list-style-type: none"> • Understand physical concepts • Relate different physics concepts with each other • Relate physics concepts with daily life applications • Interpret the physical meaning of an equation
PS	<ul style="list-style-type: none"> • Peer Instruction • Problem set reflection • Estimation activity • Project-based learning approach 	<ul style="list-style-type: none"> • Interpret graphs explaining physical phenomenon • Answer conceptual physics questions (e.g., learning catalytics) in class individually • Solve qualitative physics problems • Handle mathematical calculations while solving physics problems • Evaluate the plausibility of results of physics problems
CW	<ul style="list-style-type: none"> • Team-based learning approach • Project-based learning approach 	<ul style="list-style-type: none"> • Apply physical equations in order to solve physics problems • Work together with classmates to complete a complex task (e.g., a physics project) • Be flexible in the face of conflicts and disagreements in group activities • Communicate physics in a way that the classmates understand • Encourage the classmates to participate in discussions in group activities • Listen the opinion of the classmates, even when the student thinks he/she is right in a discussion
LHA	<ul style="list-style-type: none"> • Experimental design activity • Project-based learning approach 	<ul style="list-style-type: none"> • Design physics experiments using materials in hands-on activities (i.e., in class or in lab) • Interpret data while conducting physics experiments • Write reports summarizing physics experiments • Collect data while conducting physics experiments • Use the equipment during hands-on activities (e.g., in class or in lab)

TABLE III. Cronbach's alpha for the entire PSES survey and its four domains of physics self-efficacy.

Dimension	Statements	α
Entire survey	20 items	0.924
CPU	5 items (1, 2, 10, 11, 12)	0.875
PS	5 items (5, 7, 19, 16, 13)	0.834
CW	5 items (18, 17, 15, 4, 6)	0.760
LHA	5 items (20, 14, 8, 9, 3)	0.805

survey questions in exactly the same way as the experts. In future implementations of this survey, we intend to conduct student interviews using think-aloud protocol [52] to strengthen the content validity of the instrument.

The Cronbach's alpha coefficients were computed for the entire PSES survey and for each of the four domains measured by this survey. The survey was completed by 175 students. The results are presented in Table III. As it shows, the Cronbach's alphas were adequate [53].

III. RESULTS

Figure 1 shows the physics self-efficacy score at the beginning of the semester (pre) and the end of the semester (post) for male and female students for 2016 and 2017. To calculate this physics self-efficacy score, we averaged students' responses for the 20 items on the PSES survey.

The data from the two years are nearly identical. We performed two-tailed t tests to compare average student self-efficacy at the beginning of the semester to that at the end of the semester for both male and female students. At the beginning of the semester, female students' physics self-efficacy is 13% lower than that of the male students ($p < 0.001$). At the end of each semester, the physics self-efficacy gender gap disappears; female students' overall physics self-efficacy is not statistically different from male students' physics self-efficacy ($p = 0.19$). Figure 1 shows that female students' post-semester physics self-efficacy is significantly higher than their presemester physics

self-efficacy, while there is no significant change in the physics self-efficacy of male students.

Figure 2 and Table IV show the change in physics self-efficacy, overall and for each of the four dimensions separately. Data from both semesters (2016 and 2017) have been combined. We performed two-tailed t tests to compare the change in self-efficacy (overall and in each of the four dimensions) for male students to the change in self-efficacy for female students.

Given that the classes consist of 48%–50% premedical students, an alternative hypothesis could be that these results are more a “pre-med” vs “engineers” effect than a gendered one. We also performed a two-tailed t -test to compare students' average change of self-efficacy. There is no statistically significant difference between the overall change in physics self-efficacy of premedical and engineering students ($p = 0.57$). In addition, despite having prior formal instruction in physics, which is important to the development of the personal confidence, women consistently report lower physics self-efficacy [16].

The first set of bars in Fig. 2 reflects the changes seen in Fig. 1. Female students' physics self-efficacy improves, on average, by one point on the ten-point scale ($p < 0.001$), which corresponds to an effect size of 0.45. Male students' overall self-efficacy improves, on average, one-third of a point, but this change is not statistically significant ($p = 0.13$). Female students' self-efficacy also shows a statistically significant change in each of the four dimensions over the course of the semester. The only dimension that shows a statistically significant change for male students is the lab and hands-on activities dimension, which improves by almost 1 point ($p < 0.01$). The changes in both the conceptual physics understanding and problem-solving dimensions for female students are statistically significantly higher than the corresponding changes for male students ($p = 0.003$).

Table IV also shows the statistical significance of the differences in self-efficacy between male and female students. We performed two types of two-tailed t tests;

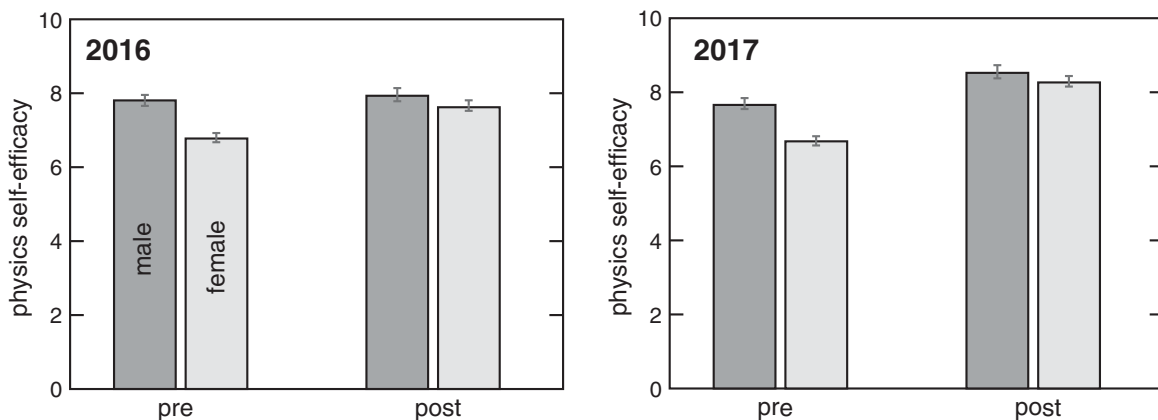


FIG. 1. Physics self-efficacy scores obtained using the PSES survey in Appendix at the beginning (pre) and end (post) of the semester. Dark gray: male students; Light gray: female students. $N = 65$ for 2016 and $N = 39$ for 2017.

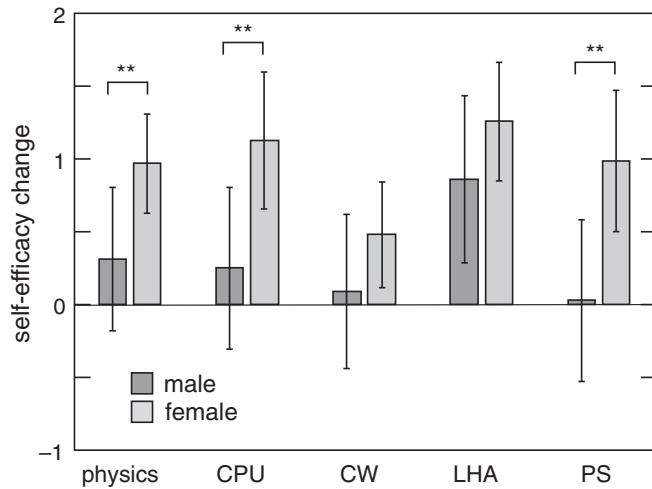


FIG. 2. Self-efficacy change (overall and in each of the four dimensions) over the course of the semester for male (dark gray) and female (light gray) students. Data from both semesters (2016 and 2017) have been combined. Error bars represent the standard error of the mean; error bars that cross the axis indicate a change that is not statistically different from zero. $**p < 0.01$ statistical significance of difference between male and female students. Abbreviations for the sources of self-efficacy are as shown in the text.

to compare the pre- and postsemester self-efficacy (overall and in each of the four dimensions) and also to compare self-efficacy for both male students and for female students.

TABLE IV. Average self-efficacy at the beginning (pre) and end (post) of the semester overall and in each of the four dimensions for male and female students showing statistical significance of the changes and differences in each item (st. err. = standard error). Data averaged for 2016 and 2017.

Dimension	Male (st. err.)	Female (st. err.)	Difference
Physics (pre)	7.86 (0.19)	6.93 (0.12)	0.92 ^a
Physics (post)	8.17 (0.16)	7.90 (0.12)	0.26
Change	0.31	0.97 ^a	
CPU (pre)	7.77 (0.21)	6.29 (0.16)	1.48 ^a
CPU (post)	8.02 (0.19)	7.42 (0.17)	0.60 ^b
Change	0.25 (0.28)	1.13 ^a	
PS (pre)	7.87 (0.22)	6.41 (0.17)	1.46 ^a
PS (post)	7.90 (0.18)	7.4 (0.18)	0.5
Change	0.03	0.99 ^a	
CW (pre)	8.25 (0.21)	7.97 (0.15)	0.28
CW (post)	8.34 (0.17)	8.45 (0.11)	0.11
Change	0.09	0.48 ^b	
LHA (pre)	7.53 (0.24)	7.06 (0.17)	0.47
LHA (post)	8.40 (0.17)	8.33 (0.12)	0.07
Change	0.86 ^a	1.26 ^a	

^a $p < 0.001$.

^b $p < 0.01$.

At the beginning of the semester, there is a statistically significant gender gap in the conceptual physics understanding and problem-solving dimensions, with the physics self-efficacy of males higher than that of females. At the end of the semester, the difference between male and female students in the problem-solving dimension is no longer statistically significant. The difference in the conceptual physics understanding dimension decreases significantly but does not disappear completely.

IV. DISCUSSION

In agreement with previous studies, we find that male students initially have significantly higher physics self-efficacy than female students [15]. We also find that over the course of team- and project-based physics class, this gender gap disappears. More specifically, physics self-efficacy improves significantly across all four self-efficacy dimensions for female students and improves only in the lab and hands-on activities dimension for male students. This disappearance of a gender gap in physics self-efficacy is a significant improvement over the generally detrimental effect instruction has on physics self-efficacy in traditionally taught physics courses [16,41], and in so-called interactive engagement courses [19].

The fact that women's self-efficacy improves significantly more than men's self-efficacy in this class is of particular interest. This may be due to the fact that there were more female students than male students in both semesters of the class and perhaps this helped decrease the overall stereotype threat [14,54]. However, other physics self-efficacy studies with similar gender distribution (more than 50% of women) pointed out that women still report lower physics self-efficacy [18].

Another explanation for women's self-efficacy improving significantly more than men's self-efficacy could be the specific course activities and structure. The emphasis on team work, group discussion, and projects creates ample opportunity for social persuasion experiences, which increases women's self-efficacy relative to that of men [10,23–25]. Furthermore, as women have also been shown to have higher test-anxiety than men [26,27], AP50a—with its frequent, low-stakes assessment and an emphasis on teamwork and collaboration during assessment—may alleviate physiological and affective stressors. Specifically, frequent testing itself may increase students' test-taking self-efficacy, and therefore lower test anxiety and stress [55]. Taken together, the conditions of AP50a provide some promising information about the critical role of project-based courses with low-stakes assessments in increasing self-efficacy.

V. CONCLUSION

We have shown that, over the course of a team- and project-based introductory physics course, female students'

physics self-efficacy improves to the same level as that of male students. Given that female students start introductory physics courses with lower physics self-efficacy than male students, and given that this gender gap normally increases in physics courses, our findings have important implications for improving physics self-efficacy for female students. Results support the hypothesis that introducing team- and project-based learning provide students with social persuasion experiences; and deemphasizing high-stakes testing improves physiological and affective states. Further research is necessary to clarify this point.

Furthermore, education research also has shown that pedagogy affects self-efficacy, but there is no consensus on what causes gender differences in self-efficacy [56]. It is important to investigate conditions that inadvertently increase or decrease students' academic self-efficacy and offer strategies for avoiding negative conditions and enhancing positive ones.

ACKNOWLEDGMENTS

Several people contributed to the work described in this paper. T. E. conceived of the basic idea for this work. T. E., I. A., and K. M. designed and carried out the experiments, and analyzed the results. E. M. supervised the research and the development of the manuscript. K. M. wrote the first draft of the manuscript; all authors subsequently took part in the revision process and approved the final copy of the manuscript. Julie Schell, Erik Knall, and Isa Gallegos provided feedback on the manuscript throughout its development. Julie Schell and Brian Lukoff developed an earlier version of the self-efficacy survey. Tobias Espinosa developed the surveys used in this research. The research described in this paper was supported by National Science Foundation - Department of Undergraduate Education under Contract No. NSF DUE-1504664. Ives Araujo partially supported by a CNPq grant (Brazil).

APPENDIX: PSES SURVEY

The appendix contains the Physics Self-efficacy Survey (PSES) designed to measure students self-efficacy across four dimensions: Conceptual Physics Understanding (CPU), Problem Solving (PS), Collaborative Work (CW), and Lab and Hands-on Activities (LHA). All PSES items are tagged with its respective dimension. The tags were not shown to respondents.

Physics Self-efficacy Survey (PSES)

For each statement, rate your belief in your ability to do the following tasks by recording a number from 0 to 10. 0 = Certain cannot do; 5 = Moderately certain can do; 10 = Highly certain can do

[Dimension of physics self-efficacy that the question pertains to]

1. Understand physical concepts [Conceptual Physics Understanding]

2. Relate different physics concepts with each other [Conceptual Physics Understanding]

3. Design physics experiments using materials in hands-on activities (i.e., in class or in lab) [Lab and hands-on activities]

4. Communicate physics in a way that my classmates understand [Collaborative Work]

5. Answer conceptual physics questions (e.g., Learning Catalytics) in class by myself [Problem Solving]

6. Work together with my classmates to complete a complex task (e.g., a physics project) [Collaborative Work]

7. Solve qualitative physics problems [Problem Solving]

8. Collect data while conducting physics experiments [Lab/Hands-on Activities]

9. Write reports summarizing physics experiments [Lab/Hands-on Activities]

10. Relate physics concepts with daily life applications [Conceptual Physics Understanding]

11. Interpret the physical meaning of an equation [Conceptual Physics Understanding]

12. Interpret graphs explaining physical phenomenon [Conceptual Physics Understanding]

13. Handle mathematical calculations while solving physics problems [Problem Solving]

14. Use the equipment during hands-on activities (e.g., in class or in lab) [Lab/Hands-on Activities]

15. Be flexible in the face of conflicts and disagreements in group activities [Collaborative Work]

16. Evaluate the plausibility of results of physics problems [Problem Solving]

17. In group activities, encourage my classmates to participate in discussions [Collaborative Work]

18. In a discussion, listen the opinion of my classmates, even when I think I am right [Collaborative Work]

19. Apply physical equations in order to solve physics problems [Problem Solving]

20. Interpret data while conducting physics experiments [Lab/Hands-on Activities]

- [1] A. Bandura, Self-efficacy: Toward a unifying theory of behavioral change, *Psychol. Rev.* **84**, 191 (1977).
- [2] S. Andrew, Self-efficacy as a predictor of academic performance in science, *J. Adv. Nurs.* **27**, 596 (1998).
- [3] R. W. Lent, S. D. Brown, and K. C. Larkin, Self-efficacy in the prediction of academic performance and perceived career options, *J. Counsel. Psychol.* **33**, 265 (1986).
- [4] K. D. Multon, S. D. Brown, and R. W. Lent, Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation, *J. Counsel. Psychol.* **38**, 30 (1991).
- [5] J. Pietsch, R. Walker, and E. Chapman, The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school, *J. Educ. Psychol.* **95**, 589 (2003).
- [6] J. Dalgety and R. K. Coll, Exploring first-year science students' chemistry self-efficacy, *Int. J. Sci. Math. Educ.* **4**, 97 (2006).
- [7] A. Bandura, in *Encyclopedia of Human Behavior*, edited by V. S. Ramachandran (Academic Press, New York, 1994), pp. 71–81.
- [8] G. Trujillo and K. D. Tanner, Considering the role of affect in learning: monitoring students' self-efficacy, sense of belonging, and science identity, *CBE Life Sci. Educ.* **13**, 6 (2014).
- [9] A. Bandura, *Self-Efficacy: The Exercise of Control* (Longman, New York, 1997).
- [10] R. W. Lent, F. G. Lopez, S. D. Brown, and P. A. Gore, Latent structure of the sources of mathematics self-efficacy, *J. Vocat. Behav.* **49**, 292 (1996).
- [11] 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).
- [12] S. L. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, *J. Res. Sci. Teach.* **43**, 485 (2006).
- [13] A. Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory* (Prentice Hall, Englewood Cliffs, NJ, 1986).
- [14] E. M. 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).
- [15] L. E. Kost-Smith, Characterizing, Modeling, and Addressing Gender Disparities in Introductory College Physics, Ph.D. thesis, University of Colorado, 2011.
- [16] V. Sawtelle, E. Brewes, and L. H. Kramer, Positive impacts of modeling instruction on self-efficacy, *AIP Conf. Proc.* **1289**, 289 (2010).
- [17] A. Cavallo, M. Rozman, and W. Potter, 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).
- [18] C. Lindström and M. Sharma, Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter?, *Int. J. Innovation Sci. Math. Educ.* **19**, 1 (2011).
- [19] J. M. Nissen and J. T. Shemwell, Gender, experience, and self-efficacy in introductory physics, *Phys. Rev. Phys. Educ. Res.* **12**, 020105 (2016).
- [20] National Center for Education Statistics, Table 318.30: Bachelor's, Master's, and Doctor's Degrees Conferred by Postsecondary Institutions, By Sex of Student and Discipline Division: 2014-15," Digest of Education Statistics: 2016 Tables and Figures (2017).
- [21] National Science Foundation, *Women, Minorities, and Persons with Disabilities in Science and Engineering* (National Science Foundation, Washington, DC, 2017).
- [22] R. S. Barthelemy and M. McCormick, Gender discrimination in physics and astronomy, *Phys. Rev. Phys. Educ. Res.* **12**, 020119 (2016).
- [23] S. L. Anderson and N. E. Betz, Sources of social self-efficacy expectations: Their measurement and relation to career development, *J. Vocat. Behav.* **58**, 98 (2001).
- [24] 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).
- [25] A. L. Zeldin and F. Pajares, Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers, *Am. Educ. Res. J.* **37**, 215 (2000).
- [26] R. Hembree, Correlates, causes, effects, and treatment of test anxiety, *Rev. Educ. Res.* **58**, 47 (1988).
- [27] U. Zoller and D. Ben-Chaim, Interaction between examination type, anxiety state, and academic achievement in college science; an action-oriented research, *J. Res. Sci. Teach.* **26**, 65 (1989).
- [28] J. B. Kahle and J. Meece, Research on gender issues in the classroom, in *Handbook of Research on Science Teaching and Learning*, edited by D. L. Gabel (Macmillan, New York, 1994), pp. 542–557.
- [29] A. Zohar and D. Sela, Her physics, his physics: Gender issues in Israeli advanced placement physics classes, *Int. J. Sci. Educ.* **25**, 245 (2003).
- [30] P. Labudde, W. Herzog, M. P. Neuenschwander, E. Violi, and C. Gerber, Girls and physics: Teaching and learning strategies tested by classroom interventions in grade, *Int. J. Sci. Educ.* **22**, 143 (2000).
- [31] L. McCullough, A gender context for the Force Concept Inventory, *Proceedings of the Winter Meeting of the American Association of Physics Teachers*, San Diego, CA, 2001 (unpublished).
- [32] J. Taylor, Sexist bias in physics textbooks, *Phys. Educ.* **14**, 277 (1979).
- [33] L. Danzl-Tauer, The relationship between intervention, equity, and excellence in rural high school biology classrooms, Ph.D. dissertation, Purdue University, 1990; Results described in J. B. Kahle and J. Meece, Research on gender issues in the classroom, in *Handbook of Research on Science Teaching and Learning*, edited by D. L. Gabel (Macmillan, New York, 1994), pp. 542–557.
- [34] A. Kitchenham, Vive la difference: Gender, motivation, and achievement, *School Libraries in Canada* **22**, 34 (2002).
- [35] P. Laws, P. Rosborough, and F. Poodry, Women's responses to an activity-based introductory physics program, *Am. J. Phys.* **67**, S32 (1999).
- [36] M. Schneider, Encouragement of women physics majors at grinnell college: A case study, *Phys. Teach.* **39**, 280 (2001).

- [37] B. Smail, Organizing the curriculum to fit girls' interests, in *Science for All Girls?*, edited by A. Kelly (Open University Press, Buckingham and Philadelphia, 1984).
- [38] G. M. Hilderbrand, Redefining achievement, in *Equity in the Classroom: Towards Effective Pedagogy for Girls and Boys*, edited by P. F. Murphy and C. V. Gipps (Falmer and UNESCO, Washington, DC, 1996), pp. 149–171.
- [39] H. Stadler, R. Duit, and G. Benke, Do boys and girls understand physics differently?, *Phys. Educ.* **35**, 417 (2000).
- [40] M. Lorenzo, C. H. Crouch, and E. Mazur, Reducing the gender gap in the physics classroom, *Am. J. Phys.* **74**, 118 (2006).
- [41] R. Dou, E. Brewe, J. P. Zwolak, G. Potvin, E. A. Williams, and L. H. Kramer, Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens, *Phys. Rev. Phys. Educ. Res.* **12**, 020124 (2016).
- [42] P. C. Blumenfeld, E. Soloway, R. Marx, J. Krajcik, M. Guzdial, and A. Palincsar, Motivating project-based learning: Sustaining the doing, supporting the learning, *Educ. Psychol.* **26**, 369 (1991).
- [43] L. K. Michaelsen, A. B. Knight, and L. D. Fink, *Team-based Learning: A Transformative Use of Small Groups* (Greenwood Publishing Group, Westport, 2002).
- [44] E. Mazur, *Peer Instruction A User's Manual* (Prentice-Hall, Upper Saddle River, NJ, 1997).
- [45] L. C. McDermott and P. S. Shaffer, *Tutorials in Introductory Physics* (Pearson, Boston, 2010).
- [46] PhET Simulations, <https://phet.colorado.edu/en/simulations/category>.
- [47] H. Fencl and K. Scheel, Pedagogical approaches, contextual variables, and the development of student self-efficacy in undergraduate physics courses, *AIP Conf. Proc.* **720**, 173 (2004).
- [48] V. Sawtelle, *A Gender Study Investigating Physics Self-Efficacy* (Florida International University, Miami, FL, 2011).
- [49] A. Bandura, in *Self-Efficacy Beliefs Adolesc.*, edited by F. Pajares and T. Urdan (Information Age Publishing, Greenwich, 2006), pp. 307–337.
- [50] F. Yaghmaei, Content validity and its estimation, *J. Med. Educ.* **3**, 25 (2003).
- [51] D. M. Rubio, M. Berg-Weger, S. S. Tebri, E. S. Lee, and S. Rauch, Objectifying content validity: Conducting a content validity study in social work research, *Social Work Research* **27**, 94 (2003).
- [52] R. Jääskeläinen, Think-aloud protocol, in *Handbook of translation studies*, edited by Y. Gambier and L. van Doorslaer (John Benjamins publishing company, Amsterdam, 2010), pp. 371–374.
- [53] R. F. DeVellis, *Scale Development: Theory and Applications* (Sage, Newbury Park, CA, 2003).
- [54] C. Hill, C. Corbett, and A. Rose, Why so few? Women in science, technology, engineering, and mathematics (2010), <https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf>, <http://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2010.tb01066.x/epdf>.
- [55] J. A. Schell and A. C. Butler, Insights from the science of learning can inform evidence-based implementation of peer instruction, *Front. Educ.* **3**, 1 (2018).
- [56] J. Dalgety and R. K. Coll, Exploring first-year science students' chemistry self-efficacy, *Int. J. Sci. Math. Educ.* **4**, 97 (2006).