Gender & Self-Efficacy: A Call to Physics Educators

Many students across the United States enter college with aspirations of becoming a successful career scientist within the field of Science, Technology, Engineering, and Mathematics (STEM). However, the number of bachelor's degrees awarded in physics has significantly lagged behind the number of bachelor's degrees awarded in the other STEM disciplines[1]. Of the physics bachelor's degrees awarded nationally in 2014, only 20% were conferred to women[1]. As a part of the conversation on recruitment, retention, and diversity in physics, researchers have focused on students' self-efficacy (SE), or one's personal beliefs in their capabilities to execute a specific task [2]. SE is highly correlated with performance and success[3], career aspirations[4] and student persistence, particularly in physics [5,6]. In addition, many studies have shown that men and women evaluate their science SE differently with women, on average, reporting a lower SE toward science.[7] This article will provide a robust literature review about the research reporting on the gender differences in science SE, specifically within the physics discipline. We will highlight common resources educators can use to measure students' SE in their own physics classrooms, the standard findings that SE decreases in introductory physics courses but not in other science courses, and within physics the decreases tend to be larger for female students.

Literature Review Methods

We reviewed 7 studies within physics[8–14] and 7 studies within other STEM disciplines[15–21] including chemistry, mathematics, and biology. The studies provided data for both our investigation of overall SE shifts in courses and gender differences in physics SE. The descriptive statistics for the studies on overall shifts in SE, which we used to construct Figure 1, are provided in Table 1. These studies included data from 15 physics courses and 14 STEM courses outside of physics (5 chemistry, 2 mathematics, and 7 biology); the specific details for each of these courses, including the type of course and pedagogy used in the classroom, can be found in the online appendix. Table 2 summarizes the descriptive statistics for the studies on gender differences in physics SE that we used to construct Figure 2; this figure only includes data from the 15 physics courses where differences in SE between men and women were reported.

Where possible, we reported the effect sizes as reported in the individual studies and in all other cases, we calculated the effect sizes by dividing the difference in the means by the standard deviation for the matched samples ($N_{pre} = N_{post}$) or the pooled standard deviation for the un-matched samples ($N_{pre} \neq N_{post}$). It is important to note that different scales were used in the various studies. Some studies reported the means on the same scale as the instrument (i.e., 1-5) while others calculated the mean of the sum of scores for the students (i.e., 33-165 is the range of scores for a 33-item survey with a 5-point Likert scale). In a few studies, the researchers reported descriptive statistics for the data disaggregated across student groups, instrument sub-constructs, or individual questions. In these cases, we aggregated the data using standard weighting practices for means and standard deviations.

The studies we include in this review are limited to those that include multiple measurements across time because, as Eddy and Brownell[22] points out, studies that do not follow students through time cannot inform the impact of degree programs or courses. Most of the studies we reviewed used surveys administered at the beginning and end of a course; however, there are additional studies that have used the experience sampling method to survey students self-efficacy over time and more proximal to the various experiences that students have in the classroom[23]--see the online appendix for more information.

Table 1. Descriptive statistics from studies of self-efficacy in physics, chemistry, mathematics and biology courses. The population column indicates if the course was intended for “majors” or “nonmajors” in that discipline. Studies that used a matched sample are indicated with “*” next to the author name. All effect sizes were either reported from the individual study or calculated from the descriptive statistics provided.

Article	Population	Scale	Pretest			Post-test			Change
			Mean	Sd	N	Mean	Sd	N	*(d)*
Physics									
Kost-Smith*	Majors	N/A	-	-	329	-	-	329	-0.47
Kost-Smith*	Majors	N/A	-	-	484	-	-	484	-0.38
Kost-Smith*	Majors	N/A	-	-	352	-	-	352	-0.23
Kost-Smith*	Majors	N/A	-	-	357	-	-	357	-0.16
Nissen et al.	Majors	1-5	3.43	0.55	117	3.36	0.65	117	-0.11
Dou et al.	Majors	33-165	135.3	13.8	147	129.0	17.2	147	-0.40
Cavallo et al.	Nonmajors	3-15	11.93	2.00	152	11.94	2.19	240	0.00
Lindstrøm et al.	Nonmajors	5-25	18.69	3.03	90	17.76	2.65	118	-0.33
Lindstrøm et al.	Majors	5-25	18.81	2.99	191	17.92	2.67	153	-0.31
Lindstrøm et al.	Nonmajors	5-25	17.46	2.77	64	17.62	3.08	74	0.05
Lindstrøm et al.	Majors	5-25	18.64	2.29	58	19.10	2.43	71	0.19
Sawtelle et al.	Majors	1-5	3.84	0.76	70	3.86	0.76	70	0.02
Sawtelle et al.	Majors	1-5	3.57	0.61	175	3.30	0.61	175	-0.43
Marshman et al.	Majors	1-4	2.79	0.43	798	2.65	0.47	595	-0.30
Marshman et al.	Majors	1-4	2.86	0.40	691	2.79	0.43	319	-0.16
Chemistry									
Ferrel et al.*	-	1-5	3.29	0.60	294	3.77	0.68	294	0.50
Ferrel et al.*	-	1-5	2.87	0.73	175	3.69	0.64	175	1.19
Dalgety et al.	Majors	1-7	4.34	1.33	126	4.4	1.36	109	0.04
Dalgety et al.	Majors	1-7	4.4	1.36	109	4.75	1.32	84	0.26
Villafane et al.	Nonmajors	1-5	2.93	0.83	297	3.39	0.81	229	0.56
Mathematics									
Brewer*	Nonmajors	N/A	-	-	72	-	-	72	0.51
Brewer*	Nonmajors	N/A	-	-	54	-	-	54	0.51
Biology									
Lawson et al.	Nonmajors	1-5	2.81	0.72	436	3.81	0.57	436	1.55
Ainscough et al.	Majors	23-115	65.7	12.6	614	75.9	12.4	614	0.81
Roster*	Combined	N/A	-	-	58	-	-	58	0.28
Roster*	Combined	N/A	-	-	46	-	-	46	0.63
Roster*	Combined	N/A	-	-	109	-	-	109	0.38
Roster*	Combined	N/A	-	-	76	-	-	76	0.39
Roster*	Combined	N/A	-	-	33	-	-	33	0.44

Table 2. Descriptive statistics from studies reporting gender differences in self-efficacy within the physics classroom. Studies that used a matched sample are indicated with “*”. All effect sizes were either reported from the individual study or calculated from the descriptive statistics provided.

	Women						Men							
	Pretest			Post-test			Pretest			Post-test			*d*	
Author	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Pre	Post
Kost-Smith*	-	-	-	-	-	-	-	-	-	-	-	-	0.14	0.13
Kost-Smith*	-	-	-	-	-	-	-	-	-	-	-	-	0.42	0.44
Kost-Smith*	-	-	-	-	-	-	-	-	-	-	-	-	0.38	0.58
Kost-Smith*	-	-	-	-	-	-	-	-	-	-	-	-	0.13	0.18
Nissen*	3.29	0.54	27	3.13	0.46	27	3.47	0.55	90	3.43	0.69	90	0.34	0.47
Dou*	134.8	12.5	61	127.5	17.9	61	135.7	14.8	86	130.2	16.7	86	0.06	0.16
Cavallo	11.4	1.96	76	11.3	2.24	120	12.4	2.03	76	12.6	2.13	120	0.51	0.58
Lindstrøm	17.8	3.12	51	16.9	2.52	66	19.9	2.91	39	18.9	2.8	52	0.71	0.75
Lindstrøm	18.1	3.54	66	16.8	3.04	59	19.2	2.65	125	18.6	2.41	94	0.38	0.66
Lindstrøm	17.2	2.64	38	16.9	3.13	43	17.9	2.96	26	18.6	3.01	31	0.25	0.54
Lindstrøm	18.2	2.67	30	18.5	2.41	36	19.1	1.8	28	19.7	2.45	35	0.42	0.52
Sawtelle*	3.73	0.61	40	3.81	0.61	40	3.99	0.92	30	3.93	0.92	30	0.34	0.15
Sawtelle*	3.45	0.61	65	3.16	0.61	65	3.63	0.62	110	3.39	0.62	110	0.24	0.30
Marshman	2.56	0.45	279	2.38	0.49	201	2.91	0.42	519	2.79	0.46	394	0.79	0.89
Marshman	2.70	0.42	224	2.52	0.45	98	2.93	0.39	467	2.91	0.43	221	0.57	0.90

Gender Differences in Self-Efficacy

The field of science education has studied gender differences in science SE, largely focused on "science" without an emphasis on disciplinary SE. In this section, we present recently published studies addressing the gender differences in SE with a focus on those within the physics classroom. While we recognize that this is not exhaustive, our intention is to focus on articles that included a thorough literature review from where the reader can learn more. Table 3 presents three main points that are important for studying gender differences in SE along with a few key studies that can provide more information to the reader. An online appendix provides additional references for interested readers.

Table 3: Summary points from surveying the literature. Key references to review these points are presented in this table. For additional references, see the appendix.

Summary Point from Literature	Key References
There are instruments available for teachers to study gender differences in self-efficacy within their physics classroom.	1. L. Kost-Smith, "Characterizing, Modeling, and Addressing Gender Disparities in Introductory College Physics," Dissertation 1–341 (2011). 2. 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, (2010). 3. H. Fencl 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, 20–25 (2005).
Students' self-efficacy tends to decrease in physics courses; however, this is not the case in other STEM courses.	1. J.M. Nissen, "Gender differences in self-efficacy states in high school physics," *Phys. Rev. Phys. Educ. Res.* 15, 13102 (2019). 2. A.M.L. Cavallo, M. Rozman, and W.H. 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," *Sch. Sci. Math*. 104, 288–300 (2004). 3. V. Sawtelle, E. Brewe, and L.H. Kramer, in *2010 Phys. Educ. Res. Conf.*, edited by C. Singh, M. Sabella, and S. Rebello (AIP, 2010), pp. 289–292.
Within the physics classroom, the gender difference in self-efficacy increases.	1. J.M. Nissen and J.T. Shemwell, "Gender, experience, and self-efficacy in introductory physics," *Phys. Rev. Phys. Educ. Res.* 12, 1–16 (2016). 2. 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, (2010). 3. 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).

1. **There are instruments available for teachers to study gender differences in self-efficacy within their physics classroom.**

Within the SE literature, many surveys have been used to study gender differences within a classroom. [24,25] Here, we discuss three instruments available to instructors to measure SE in their physics classroom. The first survey comes from Lindstrøm and Sharma[12] who developed and validated a five-item survey called the Physics Self-Efficacy Questionnaire (PSEQ). The PSEQ is a short survey intended on measuring a student's general SE in a physics classroom. This survey was designed and used within a first-year physics course and was given to students at the beginning and end of the course semester. For a more nuanced instrument to measure SE, Fencl and Scheel[26] developed a survey to analyze the four sources of self-efficacy within the physics context (mastery experiences, vicarious learning experiences, social persuasion experiences, and physiological state[2]). This second survey, the Sources of Self-Efficacy in Science Courses-Physics (SOSESC-P) is longer, consisting of 33 items and has been used to study SE in the introductory, calculus-based physics classroom.[5,10] Finally, Kost-Smith[8] developed a physics SE and identity survey with a total of 46 items on the pretest and 48 items on the posttest. The survey includes 15 items from the SOSESC-P and an additional 21 items to measure students' SE with respect to the various activities that they engage in within the physics classroom (performance, using math, getting help, and working with others); Nissen and Shemwell[9] used these 21 items in their study.

The PSEQ and the SOSESC-P are available on the PhysPort website[27] and the SOSESC-P is also available for administering online with the Learning Assistant Alliance's LASSO platform[28]. The survey developed by Kost-Smith can be found in Appendix A of her thesis[8]. These are only a few of the commonly used surveys to measure SE in the physics classroom; for instructors seeking to develop their own instrument, Lent[29] provides a guide for constructing SE scales.

2. Students' self-efficacy tends to decrease in physics courses; however, this is not the case in other STEM courses.

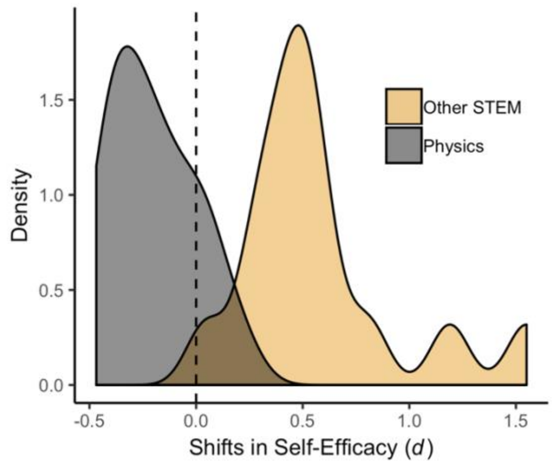

Figure 1. The distribution of shifts in self-efficacy from pre- to post-instruction for physics courses and for other STEM courses. The figure describes the shifts using Cohen's *d* [30], which standardizes the shifts across the various instruments used. Density plots show the probability distribution of the data—the probability that *d* has a value within an interval is the area within this interval and below the curve of density function. The total area under the curve is equal to one.[31] Self-efficacy decreases in most physics courses and increases in most other STEM courses. The data used for these plots comes from published work and are presented in Table 2. The citations can be found in the online appendix.

Researchers in science education have reviewed gender differences in SE across various STEM disciplines.[22,32] While these articles provide a broad review of gender differences across STEM disciplines, they provide few references for gender differences in physics, nor do they review shifts in students' SE from pre- to post-instruction. To provide context for gender differences in self-efficacy in physics, we reviewed the literature on overall shifts in SE across STEM disciplines. Figure 1 summarizes the shifts in SE in physics and other STEM disciplines. In general, students' SE increases from pre- to post-instruction in most STEM disciplines: introductory chemistry courses, introductory mathematics courses, and introductory biology courses. However, in introductory physics courses, students' SE decreases or does not change.[22,23] We are not aware of any studies within the fields of engineering or computer science that measured SE with pretests and posttests.

3. Within the physics classroom, the gender difference in self-efficacy increases.

Most of the studies in chemistry, biology, and mathematics reviewed above reported overall changes in SE and did not explore gender differences in SE. In those samples, it may be likely that the increase for all students represents an increase for women because women likely made up approximately half of the students. In contrast, many studies investigating SE in the physics classroom have focused on gender differences. As Fig. 2 shows, gender differences increased in all but two of the physics courses. In general, female students experienced a greater negative shift in physics SE than did their male peers[24]. While female students' physics self-efficacy never increased, male students' self-efficacy slightly increased in some physics courses. All of these studies reported gender as a binary construct; future work can add to understanding the role of gender in physics self-efficacy by using an inclusive measure of gender[25].

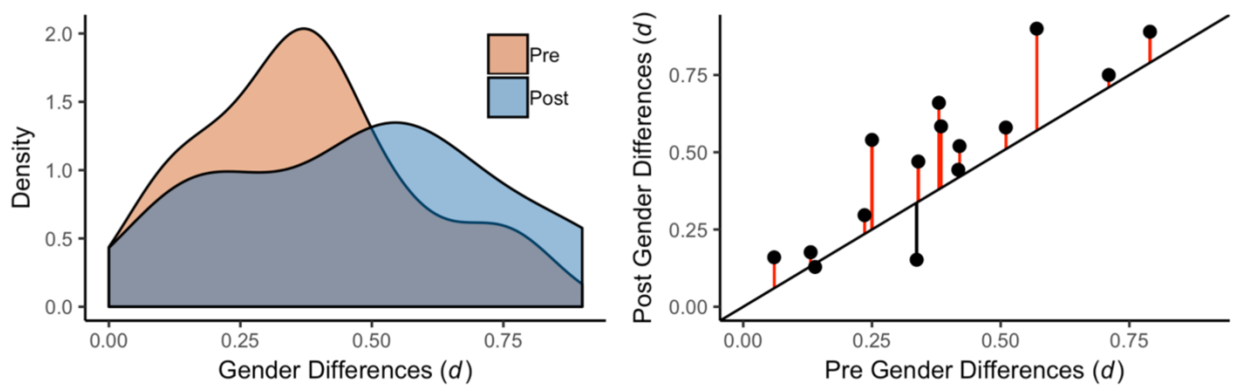

Figure 2. Gender differences in self-efficacy before and after instruction. The red lines indicate an increase in the SE gender differences from pretest to post-test while the black lines indicate a decrease. The distribution shows that in all courses gender differences favoring men existed and these inequities tended to increase from pre- to post-instruction. The figure describes the shifts using Cohen's d [30], which standardizes the shifts across the various instruments used. Density plots show the probability distribution of the data—the probability that *d* has a value within an interval is the area within this interval and below the curve of density function. The total area under the curve is equal to one[31]. The data used for these plots comes from published work and are presented in Table 3. The citations can be found in the online appendix.

A Call to Secondary Instructors and Two-Year College Faculty

Across decades, researchers have consistently found that introductory physics classrooms negatively impact SE and have a larger impact on female students than on male students. It is also likely that these negative impacts on SE are in part responsible for the lower rate in persistence of women in the field of physics. In order to change physics classrooms, we now need to turn our attention to classrooms and activities that *positively* impact SE. We would love to write a second half to this article that outlines that positive changes a teacher could make in their classroom to impact self-efficacy, but at this point we can only hypothesize as there is little work that targets classroom interventions for self-efficacy.

Fencl and Scheel's[26] work specifically looked at classroom teaching strategies in physics and their relationship to overall self-efficacy. They found strategies such as "Question and answer", "Electronic applications," and "Conceptual problem assignments" all significantly correlated with self-efficacy scores. However, looking at the details of how self-efficacy improves is a much needed and little studied effort. For instance, there has been some work to suggest that women rely on different kinds of experiences when evaluating their self-efficacy[33–35], and we have little knowledge about how particular classroom practices effect those experiences.

With this TPT article, we make a call to readers to consider measuring SE in your classrooms. We note that the majority of these data has been collected in large research universities across the United States. We believe that there are interesting activities and creative classrooms in the broader physics community that are telling a more positive story about SE. Specifically, we imagine that the teachers of secondary physics classrooms and those in our two-year colleges have a different story to tell about SE. We would like to invite these physics instructors to share those instructional designs that positively impact SE. In collecting these data, we ask the reader to consider the three measurement tools we have pointed to in this article alongside the practice of taking multiple measurements or collecting longitudinal SE data. We also note that results that show non-negative shifts, or similar shifts for female and male students are important to share. We look forward to reading about the positive impacts physics instructors are making on SE that tell counter-stories to those we have reviewed in this article.

Acknowledgements

The work that was presented in this publication was supported by DUE 1928596.

References

[1] P.J. Mulvey and S. Nicholson, *Physics Bachelor's Degrees: Results from the 2014 Survey of*

Enrollments and Degrees: Focus On. (Statistical Research Center of the American Institute of Physics, DC, 2015).

[2] A. Bandura, *Self-Efficacy:The Exercise of Control* (W.H. Freeman and Company, New York, NY, 1997).

[3] M. Richardson, C. Abraham, and R. Bond, "Psychological correlates of university students' academic performance: A systematic review and meta-analysis," Psychol. Bull. 138, 353–387 (2012).

[4] R.W. Lent, S.D. Brown, H.-B. Sheu, J. Schmidt, B.R. Brenner, C.S. Gloster, G. Wilkins, L.C. Schmidt, H. Lyons, and D. Treistman, "Social Cognitive Predictors of Academic Interests and Goals in Engineering: Utility for Women and Students at Historically Black Universities.," J. Couns. Psychol. 52, 84–92 (2005).

[5] V. Sawtelle, E. Brewe, and L.H. Kramer, "Exploring the relationship between self-efficacy and retention in introductory physics," J. Res. Sci. Teach. 49, 1096–1121 (2012).

[6] L.M. Larson, T.-F. Wu, J.D. Werbel, V.S. Bonitz, K.M. Pesch, and S. Surapaneni, "Predicting Graduation," J. Career Assess. 23, 399–409 (2014).

[7] E.L. Usher and F. Pajares, "Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions," Rev. Educ. Res. 78, 751–796 (2008).

[8] L. Kost-Smith, "Characterizing, Modeling, and Addressing Gender Disparities in Introductory College Physics," Ph.D. Thesis, Univesrity of Colorado, Boulder, 2011.

[9] J.M. Nissen and J.T. Shemwell, "Gender, experience, and self-efficacy in introductory physics," Phys. Rev. Phys. Educ. Res. 12, 1–16 (2016).

[10]] 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, 20124 (2016).

[11] A.M.L. Cavallo, M. Rozman, and W.H. 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," Sch. Sci. Math. 104, 288–300 (2004).

[12] 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, (2010).

[13] V. Sawtelle, E. Brewe, and L.H. Kramer, "Positive impacts of modeling instruction on self-efficacy," AIP Conf. Proc. 1289, 289–292 (2010).

[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] B. Ferrell and J. Barbera, "Analysis of students' self-efficacy, interest, and effort beliefs in general chemistry," Chem. Educ. Res. Pract. 16, 318–337 (2015).

[16] J. Dalgety and R.K. Coll, "The influence of first-year chemistry students' learning experiences on their educational choices ," Assess. Eval. High. Educ. 31, 303–328 (2006).

[17] S.M. Villafañe, C.A. Garcia, and J.E. Lewis, "Exploring diverse students' trends in chemistry self-efficacy throughout a semester of college-level preparatory chemistry," Chem. Educ. Res. Pract. 15, 114–127 (2014).

[18] D. Brewer, "The Effects of Online Homework on Achievement and Self-efficacy of College Algebra Students," All Grad. Theses Diss. (2009).

[19] A.E. Lawson, D.L. Banks, and M. Logvin, "Self-efficacy, reasoning ability, and achievement in college biology," J. Res. Sci. Teach. 44, 706–724 (2007).

[20] L. Ainscough, E. Foulis, K. Colthorpe, K. Zimbardi, M. Robertson-Dean, P. Chunduri, and L. Lluka, "Changes in Biology Self-Efficacy during a First-Year University Course," CBE—Life Sci. Educ. 15, ar19 (2016).

[21] N.O. Roster, The Effects of Inquiry-Based Teaching on Attitudes, Self-Efficacy, and Science

Reasoning Abilities of Students in Introductory Biology Courses at a Rural, Open-Enrollment Community College, PhD. Thesis, Oklahoma State University, 2006.

[22] S.L. Eddy and S.E. Brownell, "Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines," Phys. Rev. Phys. Educ. Res. 12, 1–20 (2016).

[23] J.M. Nissen, "Gender differences in self-efficacy states in high school physics," Phys. Rev. Phys. Educ. Res. 15, 13102 (2019).

[24] R.W. Marx, P.R. Pintrich, and R.A. Boyle, "Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change," Rev. Educ. Res. 63, 167–199 (1993).

[25] T.G. Duncan and W.J. McKeachie, "The Making of the Motivated Strategies for Learning Questionnaire," Educ. Psychol. 40, 117–128 (2005).

[26] H. Fencl 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, 20–25 (2005).

[27]PhysPort Assessments: Physics Self-Efficacy Questionnairehttps://www.physport.org/assessments/assessment.cfm?A=PSEQ.

[28]LASSO: Survey of Self-Efficacy in Science Courses - Physics (SOSESC - P) - LA Resourceshttps://sites.google.com/a/colorado.edu/la-resources/research/assessment/disciplinary-assessments/sosesc-p.

[29] R.W. Lent and S.D. Brown, "On conceptualizing and assessing social cognitive constructs in career research: A measurement guide," J. Career Assess. 14, 12–35 (2006).

[30] D. Lakens, "Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs," Front. Psychol. 4, (2013).

[31] J.L. Devore and K.N. Berk, *Modern Mathematical Statistics with Applications* (Springer,

New York, 2012).

[32] S. Cheryan, S.A. Ziegler, A.K. Montoya, and L. Jiang, "Why are some STEM fields more gender balanced than others?," Psychol. Bull. 143, 1–35 (2017).

[33] S.L. Anderson and N.E. Betz, "Sources of Social Self-Efficacy Expectations: Their Measurement and Relation to Career Development," J. Vocat. Behav. 58, 98–117 (2001).

[34] R.W. Lent, F.G. Lopez, S.D. Brown, and P.A. Gore, "Latent structure of the sources of mathematics self-efficacy," 49, 292–308 (1996).

[35] 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, 215–246 (2000).