

Investigating the role of prior preparation and self-efficacy on female and male students' introductory physics course achievements

Z. Yasemin Kalender,¹ Emily Marshman,² Christian D. Schunn,³ Timothy J. Nokes-Malach,³ and Chandralekha Singh¹

¹*Department of Physics and Astronomy University of Pittsburgh, Pittsburgh, PA, 15260*

²*Department of Physics, Community College of Allegheny County, Pittsburgh, PA, 15212*

³*Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, 15260*

Research suggests that self-efficacy is one of the central factors predicting students' engagement, participation and retention in STEM fields. Physics is one of the STEM fields in which women are severely underrepresented. Prior research has found that there is a gender gap in conceptual assessments and sometimes even in the final exam favoring men. Women also report lower self-efficacy than men in physics. The origins of these gender disparities are complex, not well understood, and include systemic societal biases and stereotypes that disadvantage women from a very young age both in and out of classroom. Since self-efficacy can impact performance and vice versa, lower physics self-efficacy than men can disadvantage women in physics classes. We studied female and male students' self-efficacy and its relation to learning outcomes in calculus-based introductory physics courses in which women are severely underrepresented. In particular, we discuss an investigation examining students' self-efficacy scores across gender and investigate the extent to which self-efficacy mediates learning outcomes for male and female students, controlling for students' relevant prior academic preparation such as AP Physics or SAT scores. We found that gender differences in course grade were partially mediated by students' prior knowledge and gender effect became non-significant factor after we include students' pre self-efficacy scores. This study can be helpful in catalyzing design and structuring of the physics classroom environment and curriculum to improve the self-efficacy and learning of all students, particularly those from traditionally disadvantaged groups such as women.

I. INTRODUCTION

In the disciplines of science, technology, engineering, and mathematics (STEM), there have been efforts to enhance the participation and advancement of women, yet the historical pattern of overall unequal gender representation remains in many STEM disciplines. In particular, over the past decades, some STEM fields, such as biology and chemistry, have shown great improvement in the number of degrees earned by women. However, other STEM fields, e.g., physics, have seen little progress. For instance, the percentages of bachelor and Ph.D. degrees in physics earned by women in the US are approximately 20% [1].

Despite the ongoing efforts, physics education researchers have still not fully understood and explained the reasons for the underrepresentation and underperformance of women in physics. The origins of these gender disparities are complex and include systemic societal biases and stereotypes that disadvantage women from a very young age both in and out of the classroom. There has been some effort in enhancing the participation and achievement of female students in physics with a broad focus on pedagogy and content improvement [2]. Researchers in physics education have recently started to investigate students' motivational factors and how they can be related to male and female students' learning of and retention in physics. In this study, we focus on self-efficacy, defined by Albert Bandura as "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations" [3]. Since self-efficacy can impact performance and vice versa, lower physics self-efficacy than men can disadvantage women in physics classes. Here, we investigate the role of self-efficacy in explaining the gender performance gap in calculus-based physics courses, in which women are numerical minorities making up 30% of the class.

II. BACKGROUND

A. Gender Differences in Physics Performance, Prior Physics Experience, and Academic Preparation

Previous studies have documented gender differences in physics performance at various institutions [4–7]. Some researchers have suggested that the gender differences in college level physics performance stem from the differences between female and male students' high school experiences and preparation [8, 9].

In high school in the US, the Advanced Placement (AP) physics courses are learning opportunities that can benefit students as they prepare for college physics. The AP physics courses can help students develop a deeper understanding of physics principles and enhance their confidence and interest in pursuing physics-related careers. Due to various reasons including societal biases and stereotypes, female students enroll in AP physics at lower rates than their male counterparts [10]. Moreover, the percentage of female students is lower in the more advanced AP physics classes, such as AP Physics

C: Mechanics [10]. Relatedly, developing robust mathematical skills can help students in college-level physics courses [9]. For example, the number of mathematics courses taken in high school is a strong independent predictor of students' college achievements in introductory science courses [9]. Likewise, research suggests that high school math grades and SAT math scores can predict college physics course success [11, 12]. In one study, high school preparation in math was found to be the strongest predictor of students' physics grades in college [8]. Mathematics as a foundation to physics is particularly relevant because gender differences have also been found in math performance [13, 14]. Societal stereotypes and biases about who is good at math and the type of encouragement and support that men and women receive to excel in math throughout their lives from family and teachers can also severely disadvantage women from a very young age. Gender performance gap in pre-college math can further impact self-efficacy beliefs of women, increase their anxiety and impact their college science course performance [15]. How students do in the first-year college courses in addition to their beliefs about how they expect to do can be one of the factors that impact students' retention.

B. Self-efficacy and Academic Performance

In learning science and educational research, self-efficacy is one of the central factors pertaining to students' beliefs about their capability to perform well in a particular domain [3]. In addition to interest, self-efficacy has been found to shape and be shaped by students' sense of being recognized as well as their effort and engagement in class [16]. Another study also showed that some of the central motivational constructs, such as interest or perception of recognition can be related to self-efficacy and they can all contribute to students' retention [17]. Likewise, the higher the students' self-efficacy in a particular learning activity, the more perseverance and resilience they are likely to show when faced with adversity [18]. When encountering challenging activities, students with low self-efficacy become less interested, make less effort and time, and eventually disengage from the class [19]. These behaviors act as a barrier to learning and development.

There is also a strong link between students' self-efficacy and academic performance. Studies in middle and high school level have shown that self-efficacy can predict student performance in science courses when controlling for prior knowledge and academic skill differences [20, 21]. Relatedly, non-physical science majors' self-efficacy was also shown to be a predictor of conceptual understanding and course achievement in physics [22].

C. Self-efficacy, Gender, and Performance in Physics

Prior research shows a prevailing gender gap in students' self-efficacy levels in physics courses [22–26]. In one study, female students were found to feel less efficacious in physics learning than male students regardless of the type of instruction (i.e., evidence-based active-engagement versus tra-

ditional) [23]. Another study identified a large self-efficacy gender gap for equally performing female and male students for all achievement groups (low, medium, high) [24, 25]: women earning As in physics often had self-efficacy levels similar to men earning Cs in physics.

Building on the success of self-efficacy studies in predicting students' achievement and retention, here we focus on the impact of self-efficacy across gender on students' college level calculus-based introductory physics achievements. Physics is one of the pillar courses taken during the first year of college and it is fundamental to almost all STEM degrees. Positive experiences in freshmen physics courses are especially important since students often decide to stay or exit the major at the end of the freshman year [28].

The research presented here focuses on the introductory level calculus-based Physics 2 courses that students take in the second half of their first year. Our primary goal is to explore the mediational mechanism of self-efficacy in explaining gender differences in Physics 2 learning outcome (physics grade), while also integrating academic skills (SAT Math and AP Physics test scores) and initial physics knowledge into the structural equation model. We hypothesize that gender differences in physics grades are mediated with prior knowledge and self-efficacy. Moreover, we also explore the contributions of SAT Math and prior knowledge as measured by standardized conceptual test as additional possible mediators of gender differences in physics grades.

In our prior work, we investigated students' conceptual post-test scores as a learning outcome in this course and attempted to explain the gender gap in the conceptual test scores using students' self-efficacy, SAT, AP Physics test and conceptual pre-test scores [26]. Here, we take our research further by using the same participants [26] and conduct an analogous analysis using students' grades as a learning outcome. In contrast to conceptual tests which measure students' conceptual understanding of physics, students' grades mostly depend on multiple midterm and final exams which evaluate students' understanding primarily using quantitative problem solving. Therefore, in the research presented here, we aim to use the same framework that we used in our prior work [26] to understand the gender performance gap in students' Physics 2 grades.

III. METHODOLOGY

A. Participants' Demographic Information and Class Context

Participants were 642 students in calculus-based physics courses who intended to major in engineering or physical sciences at a large research university. Students in the sample were enrolled in eight sections of Physics 2 courses that were taught by 4 male instructors having varying teaching experience. The demographic data (i.e., gender) were obtained from the university data warehouse that also provided records of students' pre-college test scores (AP tests, SAT, etc.) and university grades. Both motivational survey and conceptual test were administered in the first week of classes during the

recitation sections. The motivational survey took approximately 8-10 minutes and conceptual test took 40 minutes to complete. After motivational and conceptual survey responses were collected, they were sent to an honest broker to be linked with students' demographic information from the university records. Completion of this process gave researchers access to students' survey results merged with their gender and ethnic/racial identities as a de-identified dataset. In terms of demographics, 31% of the students were reported by the university as female; less than 1% of the students had not given gender information and were therefore excluded from this analysis. We acknowledge that the construct of gender is a multi-dimensional social construct that can have more than binary options [27]; however, the gender data that we use in our analysis are limited by these binary options.

B. Measurement

We previously developed and validated a self-efficacy survey that was built from prior survey instruments [24, 29]. Our instrument was iteratively refined and validated with exploratory factor analysis (EFA) and individual student interviews [24, 29]. Self-efficacy questions evaluated students' belief in their ability to understand concepts in physics and their self-perceptions of how they perform certain physics-related activities in and out of the classroom.

Students' course grades were used as a measure of their learning outcome. The final course grade was largely determined by students' midterm and final exam scores. To correct for instructor variation in grading, grades were z-scored using the class mean μ and standard deviation σ to calculate $z = (Grade - \mu)/\sigma$, essentially converting each student's grade to units of standard deviation.

The Conceptual Survey of Electricity and Magnetism (CSEM) [30] was administered to measure students' conceptual understanding of introductory electricity and magnetism, in contrast to their ability to solve quantitative problems (which can sometimes be solved algorithmically without conceptual understanding of the underlying concepts). The test was given at the beginning (pre) to measure students' initial physics knowledge.

The university uses a wide variety of measures to determine admission to the university. Here, we use two of the pre-college academic scores. The first is the Scholastic Assessment Test (SAT) Math score, which ranges from 400–800. The second measure is the standardized final exam score obtained in an Advanced Placement Physics course.

C. Analysis

An initial examination compared female and male students' scores in predictors (Self-efficacy, AP Physics test score, SAT Math, and CSEM pre-test scores) and learning outcome (Physics 1 grade) for statistical significance using t-tests and for effect sizes using Cohen's d [31]. Further, we calculated the correlations between the key constructs for two reasons: highly correlated constructs (> 0.90) would signal

that they measure non-distinguishable dimensions whereas low correlations (< 0.20) would indicate that the interrelation between the constructs was so low as to not require a direct link in the model (or could be excluded as a variable if not connected to any other variable).

To test the hypothesized path between the variables, we used Structural Equation Modeling (SEM) [32] as a statistical tool by using R (lavaan package) with a maximum likelihood estimation method [33]. There are commonly used thresholds for deciding whether the fit is acceptable or not: CFI and TLI > 0.90 ; SRMR and RMSEA < 0.08 [34]. Here, we tested the proposed theoretical model and examined the resulting structural paths between constructs. In creating a final model, we began with the saturated model (i.e., included all possible pathways), and then dropped the connections of variables that were non-significant predictors to obtain a model that produced an acceptable fit to the data and contained only statistically significant paths.

IV. RESULTS

A. Correlations and *t*-test Results

Pairwise Pearson correlations are given in Table I. Focusing on the correlations among the predictors (self-efficacy, SAT Math, AP Physics, the CSEM Pre) given in Table I, there were medium-level correlations ranging from 0.28 to 0.43, showing that the predictors are not so correlated as to be impossible to separate in the regression analyses, but also sufficiently intercorrelated that simple Pearson correlations with outcomes can be artificially higher than the true direct relationships. Table I shows that the strongest correlation was between the students' initial self-efficacy and CSEM pre-test scores, which assessed prior knowledge of Physics 2 topics. Thus, these self-efficacy judgements had a basis in reality. But the $r = 0.43$ correlation represents less than 20% shared variance, so self-efficacy is not identical to performance measured by this test or necessarily free from biases based on stereotypes and social interactions. Furthermore, CSEM Pre was moderately correlated with AP Physics test scores, suggesting that prior experience with college-level physics is quite important.

The last row of Table I presents the correlation values between the course grade and the predictors. Course grades were roughly equally correlated with CSEM pre-test, Math SAT and self-efficacy. The CSEM pre-test was most closely correlated with students' Physics Grades.

TABLE I. Pearson inter-correlations within the predictors, and between the course grade and the predictors.

	AP Physics	SAT Math	CSEM Pre	Self-efficacy
SAT Math	0.33	–		
CSEM Pre	0.40	0.36	–	
Self-efficacy	0.34	0.28	0.43	–
Course Grade	0.20	0.34	0.37	0.30

Statistically significant gender differences in favor of male students were found on all of the variables (see Table II). A very large gender gap is observed in students' initial self-efficacy reports [31]. While men had a mean score of 2.93 in self-efficacy beliefs which corresponded to a positive confidence level, women had a neutral level of confidence (mean ~ 2.6) in physics at the beginning of the course, despite all students being physical science or engineering majors. Further, the gender differences in the objective performance measures were smaller (Cohen's d), with medium differences in CSEM pre, and small differences in AP Physics, Math SAT, and Physics grade. Thus, while there are pre-existing differences based on high school experiences, the largest gender difference appeared to be one of perceived, rather than actual, physics skills and knowledge.

TABLE II. Mean predictor and outcome values by gender, along with statistical significance (p -value) and effect sizes (Cohen's d) for the gender contrast.

Predictors and Outcome	Mean		p -value	Cohen's d
	Female	Male		
AP Physics (1-5)	1.83	2.11	< 0.05	0.20
SAT Math (400-800)	709	725	< 0.01	0.29
Self-efficacy (1-4)	2.57	2.93	< 0.001	0.76
CSEM Pre (0-1)	0.37	0.42	< 0.001	0.38
Course Grade (0-4)	2.58	2.72	< 0.05	0.20

B. SEM Path Model

Our model proved to fit the data well (see Figure 1), and provided a strong fit: CFI = 0.99, TLI = 0.99, RMSEA = 0.02, and SRMR = 0.01. We report the standardized regression coefficients (β values) between the variables. Statistically significant p -values are indicated by *** for $p < 0.001$ and ** for $p < 0.01$. In the final model, direct predictors of course grade were CSEM pre-test score, SAT math, and self-efficacy variables. CSEM pre-test had the strongest connection to Physics 2 grade ($\beta = 0.27^{***}$), followed by SAT math ($\beta = 0.13^{**}$) and self-efficacy ($\beta = 0.12^{**}$) (see Figure 1). Most importantly, there was no link between gender and students' course grade in the model, which indicates that the initial gender differences were mediated by differences in students' self-efficacy and prior knowledge.

C. Total Indirect Effect

Since gender is not directly connected to Physics 2 grade in the final path model, it is possible to examine the relative contribution of gender to grade via different mediators. Therefore, we calculated the total mediated effects between gender and course grade. Within the path models, the indirect effects of gender to the outcome variables were found by multiplying the coefficients of the particular predictor that connected gender and course grade. If the predictor had more than one path between gender and learning outcome, we summed each

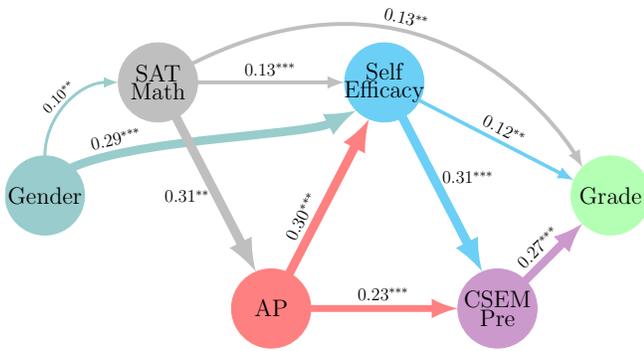


FIG. 1. Results of the structural equation modeling between gender and course grade through self-efficacy, SAT math, AP physics test score and the CSEM pre. The line thicknesses correspond to the magnitude of β values. All p -values are indicated by *** for $p < 0.001$ and ** for $p < 0.01$.

path’s contribution. For instance, one of the self-efficacy mediation paths between gender and grade flowed through SAT math so the calculation was (see Figure 1): (gender \rightarrow Sat Math) \times (SAT Math \rightarrow self-efficacy) \times (self-efficacy \rightarrow grade), i.e., $0.10 \times 0.13 \times 0.12 = 0.002$. The other self-efficacy indirect paths were also calculated by a similar procedure and then added to calculate the total indirect effect (see Table III).

The initial gender gap in students’ course grade was smaller, so it is not surprising that the mediated effects are all smaller for course grades. Overall, gender had the highest mediated effect through self-efficacy for students’ physics grade, followed by CSEM pre-test and SAT math (see Table III). In other words, the gender gap in students’ final course performances was predominantly mediated by the large gender differences in self-efficacy.

TABLE III. Direct contribution of predictors to Physics 2 Grade and the size of the total mediated contribution of the gender effect to Physics 2 Grade via each predictor. AP physics is not presented in this table since we did not find a direct connection to Grade in the model. All p -values are indicated by *** for $p < 0.001$, ** for $p < 0.01$ and * for $p < 0.05$.

Predictor	Physics 1 Course Grade	
	Direct Predictor to Outcome	Gender to Outcome via Predictor
CSEM Pre	0.27 ***	0.03 ***
Self-efficacy	0.12 **	0.04 **
SAT Math	0.13 **	0.01 *

V. CONCLUSION

We modeled the relationship between gender and physics grades as potentially driven by math performance, AP physics outcomes, initial prior knowledge of Physics 2 conceptual content, and physics self-efficacy. Each of those factors could have mediated the physics grade differences because there were statistically significant gender differences for all of the

variables (although varying effect sizes), with male students scoring higher. Consistent with prior work [13–15], SAT math showed a medium level gender gap in favor of male students ($d \sim 0.30$) but the gender performance gap was much smaller in AP physics test scores ($d \sim 0.20$).

The most striking result in our initial analysis was the large gender gap in students’ self-efficacy at the beginning of the course ($d \sim 0.8$). These self-efficacy belief differences across gender are well beyond differences in their actual physics performances, as we previously found [24, 25]. For example, we have found that even equally performing female and male students in calculus-based physics courses (both in Physics 1 and Physics 2) have large self-efficacy differences [24, 25]. Prior research has shown a relationship between self-efficacy and students’ achievements [4, 18, 22, 24] with self-efficacy predicting students’ learning after controlling for prior knowledge [24, 35]. Our path analyses went beyond that prior work to show that self-efficacy predicts performance even after controlling for prior physics knowledge and high-school-based academic differences.

Our most important finding is that the connection between gender and learning outcomes becomes non-significant in the final models of grade outcome measures. Further, the analysis of indirect effects revealed that the gendered patterns in course grade were mainly associated with students’ self-efficacy, with a modest contribution of prior physics preparation and a very small contribution from mathematics skill. We note, however, that mathematics skills and prior physics preparation appeared to be part of the causes of the large differences in self-efficacy. In particular, differences in prior mathematics and physics learning appear to play a small direct role in shaping later physics learning outcome differences by gender but plays an indirect role in shaping physics learning outcomes via undermining/supporting student self-efficacy, which then itself influences learning.

Failure to create classroom environment that supports women’s self-efficacy especially during their first-year college experiences not only has the potential for short-term negative outcomes, but is likely to lead to long-term effects, such as gendered patterns of retention in STEM domains. Supporting women’s achievement and self-efficacy necessitates promoting and supporting positive recognition and endorsement for their competence from mentors, academic advisors and course instructors as well as their families. Designing and implementing classroom interventions [36] can help to further promote higher self-efficacy, interest, and sense of belonging and improve performance and engagement in the physics classrooms for female students.

ACKNOWLEDGMENTS

This work was supported by grant DUE-1524575 from the NSF and grant No. 220020483 from the James S. McDonnell Foundation. We would also like to thank Dr. Robert Devaty for his contribution to this manuscript.

- [1] <https://www.aps.org/programs/education/statistics/womenstem.cfm>. Retrieved 6/6/2019.
- [2] P. V. Engelhardt, An introduction to classical test theory as applied to conceptual multiple-choice tests, in *Getting started in PER*, edited by C. Henderson and K. A. Harper (American Association of Physics Teachers, College Park, MD, 2009).
- [3] A. Bandura, Self-efficacy, in *Encyclopedia of Psychology*, edited by R. J. Corsini (Ed.) (New York: Wiley (1994), pp. 368-369.
- [4] V. Sawtelle, E. Brewster and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, *Journal of Research in Science Teaching* **49**, 1096 (2012).
- [5] P. M. Sadler and R. H. Tai, Success in introductory college physics: The role of high school preparation, *Science Education* **85**, 111 (2001).
- [6] L. E. Kost, S. J. Pollock and N. D. Finkelstein, Characterizing the gender gap in introductory physics, *Physical Review Physics Education Research* **5**, 010101 (2009).
- [7] 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?, *Physical Review Physics Education Research* **9**, 020121 (2013).
- [8] Z. Hazari, R. H. Tai and P. M. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, *Science Education* **91**, 847 (2007).
- [9] R. H. Tai and P. M. Sadler, Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA, *International Journal of Science Education* **23**, 1017 (2001).
- [10] American Institute of Physics, retrieved from <https://www.aip.org/statistics/reports/women-physics-and-astronomy-2019#files>
- [11] G. Hart and P. Cottle, Academic backgrounds and achievement in college physics, *The Physics Teacher* **31**, 470 (1993).
- [12] D. D. Long, G. W. McLaughlin and A. M. Bloom, The influence of physics laboratories on student performance in a lecture course, *American Journal of Physics* **54**, 122 (1986).
- [13] D. K. Leonard and J. Jiang, Gender bias and the college prediction of SATs: A cry for despair, *Research In Higher Education* **40**, 375 (1999).
- [14] H. Wainer and L. S. Steinberg, Sex differences in performance on the math section of the SAT: A bidirectional validity study, *Harvard Educational Review* **66**, 323 (1992).
- [15] P. Vincent-Ruz and C. Schunn, The increasingly important role of science competency beliefs for science learning in girls, *Journal of Research in Science Teaching* **54**, 790 (2017).
- [16] B. Zimmerman, Self-efficacy: An essential motive to learn, *Contemporary Educational Psychology* **25**, 82 (2000).
- [17] R. M. Lock, Z. Hazari, and G. Potvin, Physics Career Intentions: The Effect of Physics Identity, Math Identity, and Gender, *Physics Education Research Conference 2012*.
- [18] D. Schunk and F. Pajares, The development of academic self-efficacy, in *Development Of Achievement Motivation: A Volume in the Educational Psyc. Series*, edited by A. Wigfield and J. Eccles, (San Diego, CA: Academic Press 2002), p 15.
- [19] J. M. Bailey, D. Lombardi, J. R. Cordova and G. M. Sinatra, Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy, *Physical Review Special Topics - Physics Education Research* **13**, 020140 (2017).
- [20] S.L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, *J. of Res. in Science Teaching* **45**, 955 (2008).
- [21] K. Boden, E. Kuo, T. Nokes-Malach, T. Wallace, and M. Menekse, What is the role of motivation in procedural and conceptual physics learning? An examination of self-efficacy and achievement goals, *AIP Conference Proceedings*, edited by Ding, Traxler and Cao, (p. 60) (2017).
- [22] A. Cavallo, W. 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 Science and Mathematics* **104**, 288 (2004).
- [23] J. M. Nissen and J. T. Shemwell, Gender, experience, and self-efficacy in introductory physics, *Physical Review Special Topics - Physics Education Research* **12**, 1 (2016).
- [24] E. Marshman, Z. Y. Kalender, T. Nokes-Malach, C. Schunn and C. Singh, Female students with A's have similar self-efficacy as male students with C's in introductory courses: A cause for alarm?, *Physical Review Physics Education Research* **14**, 020123 (2018).
- [25] Z. Y. Kalender, E. Marshman, T. Nokes-Malach, C. Schunn and C. Singh, Large gender differences in physics self-efficacy at equal performance levels: A warning sign?, in *Physics Education Research Conference 2018*, (Washington, DC, August 1-2, 2018).
- [26] T. J. Nokes-Malach, Z. Y. Kalender, E. Marshman, C. D. Schunn, and C. Singh, Prior preparation and motivational characteristics mediate relations between gender and learning outcomes in introductory physics, *Physics Education Research Conference 2018* (Washington, DC, August 1-2, 2018).
- [27] A. L. Traxler, X.C. Cid, J. Blue, and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future, *Physical Review Physics Education Research* **12**, 020114 (2016).
- [28] R. Felder, G. Felder and E. Dietz, Longitudinal study of engineering student performance and retention. V. Comparisons with traditionally-taught students, *Journal of Engineering Education* **87**, 469 (1998).
- [29] E. 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, *Canadian Journal of Physics* **96**, 391 (2017).
- [30] C. Maloney, T. O'Kuma, C. Hieggelke and A. Van Heuvelen, Surveying students' conceptual knowledge of electricity and magnetism, *Am. J. Phys. Supplement* **69**, 12 (2001).
- [31] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers (1988).
- [32] R. B. Kline, *Methodology in the social sciences. Principles and practice of structural equation modeling* (4th ed.). New York, NY, US: Guilford Press, (2016).
- [33] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing (2013), <http://www.R-project.org/>
- [34] R. C. MacCallum, M. W. Browne and H. M. Sugawara, Power analysis and determination of sample size for covariance struc-

- ture modeling, *Psychological Methods* **1**, 230 (1996).
- [35] C. Lindstrøm and M. D. Sharma, Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter?, *International Journal of Innovation in Science and Mathematics Education* **19**, 1 (2011).
- [36] K. R. Binning, N. Kaufmann, E. McGreevy, O. Fotuhi, S. Chen, E. Marshman, Z. Y. Kalender, L. Limeri, L. Betancur and C. Singh, Changing social norms to foster the benefits of collaboration in diverse workgroups, unpublished manuscript, <http://d-scholarship.pitt.edu/id/eprint/36006>