Prior preparation and motivational characteristics mediate relations between gender and learning outcomes in introductory physics

Timothy J. Nokes-Malach[1], Z. Yasemin Kalender[2], Emily Marshman[2], Christian D. Schunn[1], and Chandralekha Singh[2]

[1]*Learning Research and Development Center,* [2]*Department of Physics and Astronomy*
University of Pittsburgh, Pittsburgh PA 15260

Abstract. Prior work has shown that women are underrepresented and underperform in physics compared to men. A number of factors have been identified as contributing to these gaps such as stereotypes, prior knowledge, and student motivation, but less work has tested the relations between them. We have collected motivational data and performance outcomes on students in introductory physics courses over the past several years. In the current work, we investigate how students' prior preparation and motivation relate to their learning outcomes and the gender gap. The results show that students' prior preparation and self-efficacy beliefs mediate the relation between gender and learning outcomes. Identifying factors that impact learning outcomes in physics courses is important for developing and implementing pedagogies and interventions to help all students learn and eliminate the gender gap.

I. INTRODUCTION & BACKGROUND

Much research has documented women's underrepresentation and underperformance in science, technology, engineering, and mathematics (STEM) fields [1]. Within STEM, physics has a particularly large gender gap between male and female students. We define the term *gender gap* as a measurement, a comparison typically between male and female students, that shows differences in educational or learning outcomes such as achievement, retention, and motivation. For example, even though women earn approximately 60% of all bachelor's degrees in the US only 20% of the physics undergraduate degrees are earned by women [2]. This shows that relative to ratio of women earning bachelor's degrees they are numerically underrepresented in earning physics degrees. Although efforts have been made to increase diversity in STEM and to identify the factors underlying these concerning patterns, much work remains to understand the interrelations of these factors.

Prior work has examined several factors related to the underrepresentation and underperformance of women in physics including socio-cultural reasons [3], teaching practices [4], student differences such as prior preparation [5] and motivation [6], among others [see 1, 3 for reviews]. For example, gender stereotypes can negatively affect performance outcomes by causing anxiety and rumination about one's anticipated performance [7]. These stereotypes can discourage women from majoring in STEM fields for fear of confirming the negative stereotype and can decrease their sense of belonging [8]. Women have also expressed concerns about poor teaching practices and non-welcoming, "chilly" classroom environments in which, for example, instructors are more likely to call on males than females [9]. A third factor, likely to be influenced by the previous two, is prior preparation. Research has shown that enrollment and proficiency in high school math and science courses positively predicts students' choice of a STEM major and persistence in the major in college [5]. However, high school girls are less likely than boys to take physical science and engineering courses [10]. Student motivation has also been proposed as a contributing factor to these gender gaps and prior work has examined differences between males and female students' self-efficacy, interest, feelings of belonging, and beliefs of physics intelligence, among others [1, 6].

Given the numerous factors proposed to contribute to the gender gaps in physics learning, performance, and retention, an important question arises as to how these different factors relate to one another. A few theoretical frameworks have been proposed to help researchers organize these factors and their relations with one another [1, 11]. Both Degol and Wang [11] and Eddy and Brownell [1] propose that some of these factors are more proximal (e.g., motivation) and others more distal (e.g., socio-cultural) in their impact on student performance and engagement. However, there is relatively little empirical work testing the relations between multiple factors. In the current work, we test a model that examines the relations between students' gender, prior preparation (e.g., scores on Advanced Placement (AP) science courses and Math SAT), self-efficacy beliefs, and learning outcomes in an introductory physics course.

Given the wide array of motivational factors that have been investigated, in our model we focused on students' self-efficacy beliefs, because past research has documented a particularly large gender gap in those beliefs in physics [6, 12]. Self-efficacy is one's belief about their ability to successfully accomplish a particular task or succeed in a specific domain [13]. Self-efficacy beliefs have been shown to predict course enrollment, increased persistence, self-regulated learning strategies, and academic achievement (see 14 for a review). Furthermore, some research has shown that for both sixth grade science students and college students solving math problems, their self-efficacy beliefs about the domain predicted their learning and performance outcomes,

even when controlling for their prior knowledge [15, 16]. This shows that these beliefs can have an influence on performance above and beyond prior knowledge. We test these relations between prior knowledge, self-efficacy beliefs, and performance with college students learning physics.

We predicted the relation between gender and learning outcomes for an introductory physics course would be mediated by more proximal factors such as students' prior preparation and self-efficacy beliefs. Figure 1 shows the hypothesized path analysis model that maps out the predicted directed dependencies between each of the measured variables. Specifically, we predicted that gender would be related to students' SAT math scores and their likelihood of taking AP courses. These relations are likely driven by what Eddy and Brownell call "ultimate causes" such as gender stereotypes and cultural biases [1]. We also expected AP courses and SAT math score would be correlated to students' self-efficacy such that taking more advanced science courses and having higher SAT scores would be related to higher self-efficacy. We also predicted that gender would have a correlation with students' self-efficacy in physics, with women exhibiting lower self-efficacy than men. We predicted that gender, SAT math, AP Physics, self-efficacy, CSEM pretest scores would each be correlated with CSEM post-test scores. Finally, we predicted that the measures of prior preparation and self-efficacy would explain learning on a conceptual test, mediating the relation between gender and learning outcomes.
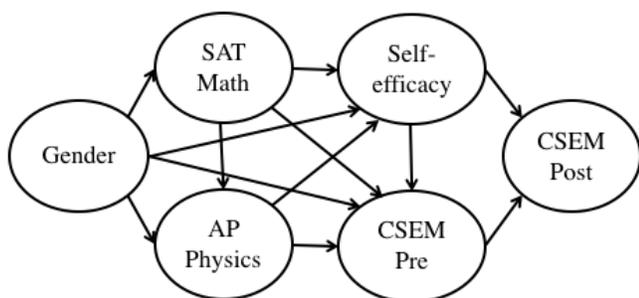

Figure 1. Hypothesized path model.

II. METHODOLOGY

To test our conceptual model, we conducted an analysis of institutional and survey data collected at a mid-size research University.

Participants were students in 8 sections of a calculus-based Introduction to Physics 2 course. A total of 542 students were included in the analysis who had complete data for all of the variables of interest (conceptual pre/posttest scores, self-efficacy, AP physics 1, and SAT Math). This was 64% of the total students enrolled in these classes. Students were primarily engineering and physics majors and the majority were 1st year students. Thirty-two percent of the sample were women.

The learning assessment was the *Conceptual Survey of Electricity and Magnetism* which consists of 32 multiple choice questions [17]. This assessment has been validated and used extensively as a conceptual measure of understanding of core physics phenomena and principles [17]. The pre and posttests were identical.

The motivation survey measured multiple physics-specific motivational constructs including self-efficacy, fascination, value, beliefs of intelligence, belonging, and identity, among others. In the current work, we focus on the physics self-efficacy construct. This measure includes items sampled from prior survey instruments and has been validated in prior research [6]. Six survey items were selected to represent a diverse set of contexts within which to assess student's sense of efficacy for physics content (see Table 1). Each item had a 4-point Likert scale. Previous work established internal coherence (Cronbach's $\alpha > .70$) and discriminability from the other motivation constructs vis-à-vis Principal Components Analysis [6]. A mean rating was calculated across the items.

Table I. The physics self-efficacy survey. One item is reverse coded and indicated with an (R).

Survey Item	Response options
I can complete the physics activities I get in a lab class	Rarely Half of the time Most of the time All of the time
If I went to a museum, I could figure out what is being shown about physics in:	None of it A few areas Most areas All areas
I am often able to help my classmates with physics in the laboratory or in recitation.	No! no yes Yes!
I get a sinking feeling when I think of trying to tackle difficult physics problems. (R)	
If I wanted to, I could be good at doing physics research.	
If I study, I will do well on a physics test.	

We obtained de-identified institutional data regarding students' sex, which was coded as 1 for male and 0 for female. In the current work, we do not focus on non-binary students due to their small sample size (n < 1%).

We also received information about whether students had taken an AP Physics 1 course. AP courses are elective high school classes that can count for college credit. There were two types of AP physics classes that students could have taken: Algebra and Calculus. We did not differentiate between these two types because they showed the same pattern of results when analyzed separately. AP exams were scored on a 1-5 basis, with higher numbers denoting higher performance. We assigned students who did not take an AP course a score of 1 because we found that those who did not take AP physics had the same average final grade in the Introduction to Physics 1 course as those who received a 1.

Students' math SAT scores were also obtained and ranged from 400 to 800. These scores are normed each year to generate a national mean of approximately 500.

Students completed a conceptual assessment and motivation survey at the beginning and at the end of the semester. These surveys were typically administered in paper booklets in the first and last recitations. Students first completed the motivation questionnaire (10-15 minutes) followed by the conceptual assessment (35-40 minutes). Instructors were encouraged to give a small amount of extra credit for completing the surveys. Institutional data and survey data were merged through an honest broker and de-identified for the research team.

We first report descriptive statistics of each variable by gender and correlations across key variables. Then we tested our path model to address our research questions. Path analysis techniques are used to examine the direct and indirect effects between the variables. We began by testing all directed dependencies in our hypothesized model and then removed non-significant paths that improved the fit of the model.

III. RESULTS

All students showed learning gains from pretest to posttest on CSEM assessment, see Table II (t= 17.4, $p < .05$, $d = .68$). In terms of gender differences, male students had higher scores than female students on math SAT, AP physics, self-efficacy, and the pre/post CSEM assessments (all F's > 4.4, p's $< .05$). The effect size differences between male and female students were small to medium for math SAT, AP physics, and the CSEM tests. The largest difference was for self-efficacy beliefs in which the mean for female students was at the middle of the scale ($M = 2.6$) and the mean for male students was 0.72 of a standard deviation above ($M = 2.9$).

Table III presents the correlations of the variables in the path analysis. All variables were significantly related to one another (p's $< .05$). Gender was significantly correlated with the CSEM posttest scores. Next, we examine whether this relation is mediated by the predictor variables of prior preparation and self-efficacy beliefs.

The strengths of association between variables within a path model was estimated using R (lavaan) and the maximum likelihood estimation method. The fit indices suggest an acceptable fit of the model to the data, χ^2 (2) = 1.97, p > .05, (CFI = 1 > .96, RMSEA = .00 < .06, AIC = 7,950, BIC = 8,027). In Figure 2, we present the standardized solution for the final path model showing all significant relations between gender, AP physics, math SAT, self-efficacy, and CSEM pre and posttest. The three significant predictors of CSEM posttest were CSEM pretest ($\boldsymbol{\beta}$=.34), self-efficacy ($\boldsymbol{\beta}$=.21), and SAT math scores ($\boldsymbol{\beta}$=.10). When all of the predictor variables are entered in the model there is no direct relation between gender and CSEM post scores (in

contrast to the initial correlation). That is, the relation between gender and conceptual posttest scores is mediated by the intervening variables of math SAT, AP physics, self-efficacy and CSEM pretest scores. We can quantify and compare the indirect effect flowing through different paths. The largest indirect path is through CSEM pretest ($\boldsymbol{\beta}$=.25), followed by self-efficacy ($\boldsymbol{\beta}$=.15), which is followed by SAT math ($\boldsymbol{\beta}$=.01).

Table II. Descriptive statistics including mean (*M*), standard deviation (*SD*), and Cohen's d effect size (*d*) for the key variables.

Variable	Males (n=368)		Females (n=174)		
	M	*SD*	*M*	*SD*	*d*
Math SAT	725	52	709	60	-0.27
AP Physics	2.1	1.5	1.8	1.3	-0.21
Self-efficacy	2.9	0.4	2.6	0.5	-0.72
CSEM pre	.42	.14	.37	.15	-0.33
CSEM post	.56	.19	.48	.19	-0.42

Table III. Zero-order correlations of variables in the path analysis.

Variables	Gender	Math SAT	AP Physics	Self-efficacy	Pre-CSEM	Post-CSEM
Gender	--					
Math SAT	.14	--				
AP Physics	.09	.33	--			
Self-efficacy	.34	.28	.34	--		
Pre-CSEM	.18	.36	.39	.43	--	
Post-CSEM	.20	.31	.30	.42	.50	--

IV. DISCUSSION

We observed small gender differences in student performance on measures of math SAT, AP physics, and on the conceptual pretest. These differences have been observed previously [3, 5, 11] and are likely driven by the combination of social expectations and stereotypes, teaching practices, and motivations at play during development and schooling. We also observed a large difference between males and females' self-efficacy beliefs in physics at the beginning of their second semester physics class. This is also consistent with prior observations and measurement of these competency beliefs [12]. We then examined the relations between measures of prior preparation and students' self-efficacy on learning outcomes. As predicted, we found that the relation between gender and learning outcomes was meditated by measures of prior preparation (math SAT, AP physics, CSEM pretest) and self-efficacy. As one might expect, prior knowledge as measured by CSEM pretest had a large effect on post CSEM scores. It is also not surprising that prior experience in mathematics and physics predicts scores on the initial CSEM pretest as we expect that practice in developing physics knowledge and mathematics skills will help learning new physics content.
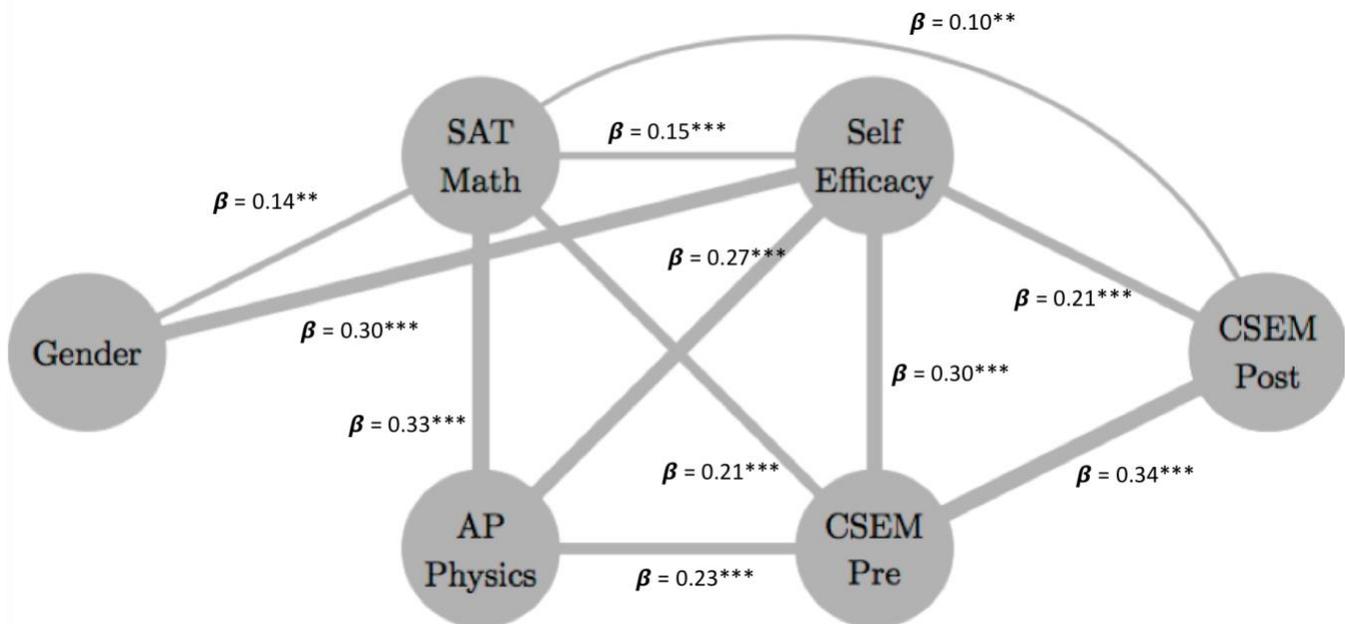
Figure 2. Standardized solution to the final path model. Line thicknesses are directly scaled by β. **p < .01, ***p < .001

This work also provides further support for the relations between the role of students' self-efficacy in predicting learning outcomes [15, 16]. That is, above and beyond prior knowledge, students' beliefs about their competence in a domain can have a large impact on what they learn. Students' prior performance and experiences (Math SAT and AP physics) shapes students' perceptions of self-efficacy but importantly, in addition to those sources of information, there is still a large relation between gender and self-efficacy, suggesting that other factors like socio-cultural expectations and stereotypes are still playing a role. The importance of self-efficacy in explaining the learning outcomes further highlights the importance of addressing the self-efficacy gap between men and women in these classes [12].

This work takes the first steps towards testing the relations of prior performance, experiences, and motivations in how they relate to gender differences in learning outcomes. These results suggest that that by measuring multiple factors hypothesized to affect learning outcomes we can begin to assess which variables have impact on others and which have the largest effects. Using this information, we can begin to develop interventions and educational innovations targeting these variables and interrelations to eliminate the gender gap.

[1] S. Eddy and S. Brownell, Phys. Rev. Phys. Educ. Res. **12**, 020106 (2016).

[2] J. Blue, A. Traxler & X. Cid, Phys. Today, **31**, 3 (2018).

[3] S. Cheryan, S. Ziegler, A. Montoya, and L. Jian, Psych. Bull. **143**, 1 (2017).

[4] E. Seymour and N. Hewitt, Westview Press (1997).

[5] E. Wang, Am. Educ. Res. J. **50**, 5 (2013).

[6] E. Marshman, Z. Y. Kalender, C. Schunn, T. Nokes-Malach, and C. Singh, Can. J. Phys. **96**, 4 (2018).

[7] C. Pennington, D. Heim, A. Levy, and D. Larkin, PLOS One **11**, 1 (106).

[8] C. Dweck, Carnegie Corporation of New York (2007).

[9] L. Fritschner, J. Higher Educ. **71**, 342 (2000).

[10] C. Nord et al., The nation's report card: America's high school graduates. US gov. printing office (2011).

[11] M. Wang and J. Degol, Dev. Rev. **33**, 304 (2013).

[12] J. Nissen and J. Shemwell, Phys. Rev. Phys. Educ. Res. **12**, 020105 (2016).

[13] A. Bandura, New York: Freeman (1997).

[14] B. Zimmerman, Contemp. Educ. Psych. **25**, 82 (2000).

[15] F. Pajares and D. Miller, J. Educ. Psych. **86**, 2 (1994).

[16] K. Boden, E. Kuo, T. Nokes-Malach, T. Wallace, and M. Menekse, AIP Conf. Proc., 2017, edited Ding, Traxler, and Cao, p. 60.

[17] C. Maloney, T. O'Kuma, C. Hieggelke, and A. Van Heuvelen, Am. J. Phys. Supplement **69**, 7 (2001).