European J of Physics Education Volume 14 Issue 2 1309-7202 Muehsler

Differences in Physics Self-Efficacy Among Personal and Course-wide Variables

Hans Eric Muehsler

Department of Educational Technology, Research and Assessment
Northern Illinois University
DeKalb, IL, USA

hmuehsler1@niu.edu

(Received 25.03.2023, Accepted 20.07.2023)

Abstract

Lab activities are a vital part of physics courses, providing students with hands-on experiences of classroom topics. At the same time, building self-efficacy is vital to student learning and retention; however, females traditionally have lower self-efficacy in STEM fields. This research explores the role of gender and the changes in interest in physics and the relation to changes in self-efficacy, as well as the interactions with math ability, the number of labs, and hours spent in the lab. Data from 260 institutions containing responses from over 11,000 women and 14,000 men was used. The self-efficacy construct was created from the E-CLASS survey. A two-level multi-level model was used to analyze the data. Level one variables included gender and change in physics interest. Level two variables included number of labs, hours spent in the lab, and math level; the dependent variable was change in self-efficacy. Change in interest positively correlated with change in self-efficacy, with women having a more favorable change. While the number of labs per semester was not a significant predictor of change in self-efficacy, females could tolerate more labs for no change in self-efficacy than men. Further, the gender effect (greater favorable change in self-efficacy for women) reverses with men showing a greater favorable change in self-efficacy than women after approximately 3.5 hours in the lab. Finally, math level had a positive correlation with change in interest and change in self-efficacy, with algebra-based labs demonstrating a more pronounced effect.

Keywords*:* physics education research, self-efficacy, gender differences, multi-level modeling, interest in physics, physics lab

INTRODUCTION

Lab activities are an integral part of the physics experience. These activities provide students with experience regarding the phenomena learned in class, development of ideas, and learning how to predict outcomes (AAPT, 1992). Developing these skills are vital to success in physics (AAPT, 1992). Building students' self-confidence, and specifically their self-efficacy, is vital to student learning and retention in the field (Espinosa, Miller, Araujo, & Mazur, 2019). Females often have lower self-efficacy regarding their ability to succeed in a STEM field. This lower self-efficacy results in fewer females in the STEM field (Ivie, White, & Chu, 2016) who are already an underrepresented group in college physics (National Center for Education Statistics (NCES), 2017) and in the profession (National Science Foundation (NSF), 2017).

Research has found that self-efficacy plays a large role in career choice (Hazari, Sonnert, Sadler, & Shanahan, 2010; Lent, Brown, & Hackett, 1994; Wang, 2013). Further, self-efficacy predicts STEM enrollment (Pajares, 2004; Zeldin, Britner, & Pajares, 2008) and persistence toward pursuing that career (Betz & Hackett, 1986; Zimmerman, 2000). In fact, after controlling for other student aspects like interest, aptitude, and achievement, self-efficacy has greater predictive power compared to most other variables (Brown, Lent, & Larkin, 1989; Lent, Lopez, & Bieschke, 1993).

Given the psychological effects self-efficacy may have on women, it is important to study this effect in STEM fields to reduce gender underrepresentation. This gender gap in STEM must close, and one way to close the gap is through studying how to improve women's self-efficacy in STEM (Espinosa et al., 2019). This paper presents the results of a study addressing how self-efficacy is related to gender, the change in interest in physics, the number of labs per semester, the number of hours per week in the lab, and the math level required for the course.

Literature Review

To address the issue of self-efficacy, Bandura's theory of human agency was employed. Agency is a person's ability to monitor, control, and regulate their own learning (Code, 2020); people make choices and engage those choices to make differences in their lives (Martin, 2004). Self-efficacy is part of this theory. Self-efficacy is one's belief in their ability to complete a task successfully in a given field (Dou et al., 2016; Kalender, Marshman, Nokes-Malach, Schunn, & Singh, 2018; Nissen & Shemwell, 2016). This characteristic is fluid and dynamic, varying across fields and potentially varying at different times within a field, based on experience (Nissen & Shemwell, 2016). Further, successful completion is based on actions the individual took to complete the task (Dou et al., 2016). Higher levels of success often result in increased self-efficacy (Nissen & Shemwell, 2016).

Self-efficacy also varies across gender; research has shown that women often leave STEM fields due to low self-efficacy, not due to lack of success (Goodman et al., 2002).

In general, self-efficacy diminishes over time in physics even when instructors use research-based teaching techniques that have demonstrated benefits to learning, especially for women (Brewe, Kramer, & O'Brien, 2009; Kost, Pollock, & Finkelstein, 2009). Further, the negative changes in self-efficacy and in attitude toward physics are larger for women than men (Kost-Smith, 2011; Kost-Smith, Pollock, & Finkelstein, 2010; Nissen & Shemwell, 2016); this finding was true across other STEM domains (Kalender et al., 2018). However, according to Nissen and Shemwell (2016), no other field had lower levels of self-efficacy as in a physics course. Women start at lower levels of self-efficacy, conceptual knowledge, attitude, and interest in physics than men; this disparity remains throughout the course (Kalender et al., 2018; Lindstrøm & Sharma, 2011; Nissen & Shemwell, 2016). This difference in self-efficacy was found across different methods of delivery of instruction (Cavallo, Potter, & Rozman, 2004; Lindstrøm & Sharma, 2011; Nissen & Shemwell, 2016; Sawtelle et al., 2010). Further, Nissen and Shemwell (2016) found women required more extrinsic motivation and felt greater stress than men in physics courses.

The surveyed literature did not demonstrate a relationship between gender, self-efficacy, and academic performance. Kalender et al. (2018) and Marshman et al. (2018) found large gender gaps in self-efficacy across all performance levels. Marshman et al. (2018) commented that this gender gap in self-efficacy may be due to students' self-perceptions than actual rated performance. To illustrate, Marshman et al. (2018) found female students in the middle performance group had the same self-efficacy as males in the low performance group on the Force Concept Inventory (FCI) test. While the gap exists, in general, students who performed at higher levels had higher self-efficacy scores (Kalender et al., 2018).

Researchers have attempted to determine the cause for the gender difference in self-efficacy. For instance, Nissen and Shemwell (2016) believed teaching method played a larger role in gender differences in self-efficacy than content. However, Nissen and Shemwell (2016) and Espinosa et al. (2019) also stated the content played a role in the self-efficacy difference. They found women consistently had lower and large differences in self-efficacy than men in physics compared to other STEM fields. Dou et al. (2016) also commented on this finding, but attributed the drop in male levels of self-efficacy in lower performing groups to a correction in self-confidence. Kalender et al. (2018) posited a feedback loop may be at play in these relationships; that is, students who performed well develop higher degrees of self-efficacy, which in turn led to even higher performance. Conversely, students who performed lower developed lower degrees of self-efficacy, which led to even poorer performance. The gender difference also may have been due to differences in student agency. For instance, women's self-efficacy was impacted through vicarious experience and social persuasion, whereas males were more influenced by mastery experiences (Anderson & Betz, 2001; Zeldin et al., 2008; Zeldin & Pajares, 2000).

Research gap

Much research has been performed regarding a gender gap in self-efficacy. However, little research has been performed regarding changes in interest in physics and its relation to changes in self-efficacy, much less changes in self-efficacy. Further, little research was found that addressed differences in required math abilities or the number of labs and hours spent in the lab. It was hypothesized that more time in the lab would increase self-efficacy, as students explore physics phenomena, in accordance with AAPT's statement about the importance of labs (AAPT, 1992), but only to a point. It was also hypothesized that calculus-based labs had higher levels of self-efficacy due to potentially increased interest in STEM fields.

Research Questions

To address these points, a series of research questions were designed to use with this data set:

RQ 1: Is there a relationship between change in interest in physics and gender (IVs) and the change in self-efficacy (DV)?

RQ 2: Does the math level predict the relationships between gender and change in interest in physics with the change in self-efficacy?

RQ 3: Do the hours in the lab predict the relationships between gender and change in interest in physics with the change in self-efficacy?

RQ 4: Do the number of labs predict the relationships between gender and change in interest in physics with the change in self-efficacy?

METHOD

The data for this study were described by Aiken (2021). Aiken and Lewandowski (2021) collected and posted a 70,000-response data set regarding student responses to prompts concerning student experiences participating in college-level physics labs. The anonymized sample included 133 institutions; the sample further consisted of 599 unique courses and 204 instructors. The data were collected between 2016 and 2019 (Aiken & Lewandowski, 2021). Participants responded to a survey at the start and end of the course. Overall, the data set consisted of 35,380 responses at the start of the course and 28,282 responses at the end of the course (Aiken & Lewandowski, 2021). Participating students were sampled in introductory and "beyond the first course" lab sections. Much of the student data consisted of responses to the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) survey. This survey provided 30 Likert-like items concerning students' epistemologies and expectations and asked students to posit an expert's responses to the items. Students also provided demographic information (Aiken & Lewandowski, 2021). The other part of the survey consisted of instructors' responses to items related to the course such as the required math level, pedagogical approaches, and frequency of laboratory experiences (Aiken & Lewandowski, 2021). The data were freely available and part of an effort to address recent calls to provide free access and open sharing of data (e.g. (American Institute of Physics (AIP), 2003; S.1701 - 115th Congress (2017-2018), 2017; Shunter, 2022).

The self-efficacy construct was created from items in the E-CLASS survey. An exploratory factor analysis (EFA) with a varimax rotation was performed on the self-efficacy variable. The Kaiser-Meyer-Olkin measure of sampling adequacy was .867, which is considered acceptable (Tabachnick & Fidell, 2001). Further, Bartlett's test of sphericity was found to be statistically significant ($p<.001$). Both of these conditions indicated a factor analysis was appropriate for the data. A six-factor solution was obtained, based on the scree plot and the eigenvalues greater than one. The component of interest consisted of six items; sample items included: "If I try hard enough, I can succeed at doing physics experiments," "When I approach a new piece of lab equipment, I feel confident I can learn how to use it well enough for my purposes," and "Nearly all students are capable of doing a physics experiment if they work at it." Response numerical values were added to form a composite value of self-efficacy for the pre and post-survey. The change in self-efficacy was calculated as the final composite values minus the initial composite values.

For this analysis, there were 11,338 women and 14,043 men. Further, 7,783 people stated their interest in physics increased, 12,986 people stated their interest remained the same, and 5,352 people stated their interest in physics decreased by the end of the semester. The number of included colleges was 260, with approximately 220 calculus-based lab courses and 141 algebra-based courses. The number of hours spent in the lab per week varied from one-half to five.

The method of analysis was a two-level multi-level model (MLM) using a maximum likelihood estimation. The dependent variable was the change in self-efficacy over the semester. The level one variables were gender and the change in physics interest. Gender was a discrete, nominal variable and was grand-mean centered, allowing for both genders to be considered and compared. Change in physics interest was considered a continuous, interval level variable, and was also grand-mean centered. This centering allowed for comparisons for the average change in physics interest to be zero. The level two variables were the number of labs per semester, the hours spent in the lab. Both the number of labs per semester and the number of hours spent in the lab were considered continuous, ratio-level data. The number of labs per semester and number of hours in the lab were screened for potential outliers. Values of greater than 20 hours per week or 20 labs per semester were removed from the final data set. Math level was a categorical variable and recoded as "Algebra-based" (0) or "Calculus-based" (1). The cluster variable was the class or course, represented by a unique identifier code; there were 260 clusters.

RESULTS

First, the amount of clustering was determined to ensure multilevel modeling was appropriate. The intraclass correlation coefficient was .025 with a mean cluster size of 101. This resulted in a design effect of 3.51. Because the design effect was larger than the recognized threshold of two (Hox & Maas, 2002; Muthen & Satorra, 1995), the clustering effect could not be ignored and was accounted for in this paper. Further, because of the interest in the level two predictors, clustering of the cases was used.

Research Question 1

The first research question was to determine if gender and change in interest in physics predicted change in self-efficacy in physics. This research question utilized a level 1 predictors model, allowing the slopes for gender and change in physics interest to change:

Level 1: $$\Delta Self-efficacy = \beta_{0j} + \beta_{1j} Gender_{ij} + \beta_{2j} \Delta Interest_{ij} + r_{ij}$$

Level 2: $$\beta_{0j} = \gamma_{00} + \mu_{0j}$$
$$\beta_{1j} = \gamma_{10} + \mu_{10}$$
$$\beta_{2j} = \gamma_{20} + \mu_{20}.$$

The fixed effects for gender and change in physics interest were both significant predictors of the change in self-efficacy. As the change in interest moved to more positive numbers, the change in self-efficacy became increasingly positive. Women had nearly the same change in the change in self-efficacy as men in this model; however, women had a greater mean favorable change in self-efficacy than men. The random effects were also found to be significant, indicating the slopes significantly varied, albeit slightly, among colleges. The results are shown in Table 1 and Figure 1.

The change in physics interest accounted for 3% of the variability in the change in self-efficacy, whereas gender accounted for less than 0.1% of the variability in the change in self-efficacy. A small, positive correlation was found between the slopes of the change in physics interest and gender; men were found to have more extreme loss in interest than women (*r*=.063).

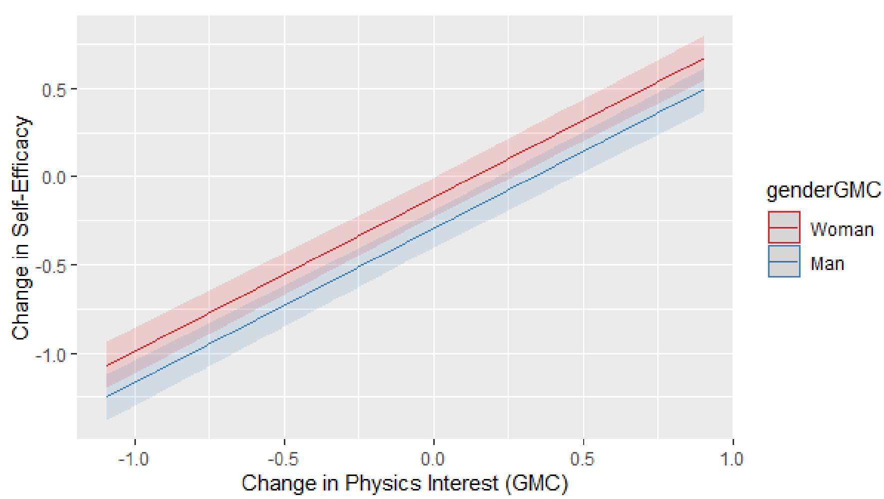

Figure 1. *Change in self-efficacy vs. change in interest by gender.*

Table 1. *Random Slopes And Intercepts For Change In Interest*

	Random Slopes and Intercepts for Level 1 Predictors		
Predictors	Estimates	Std. Error	*p*
Intercept	-0.12	0.06	.058
Change in Physics Interest (GMC)	0.95	0.06	< .001
Gender (GMC)	-0.16	0.06	.005
Random Effects			
σ^2	14.02		
τ_{00} Response ID	0.42		
τ_{11} ResponseID.ChgIntGMC	0.26		
τ_{11} ResponseID.genderGMC	0.10		
ρ_{01}	0.42		
	-0.83		
ICC	0.02		
$N_{ResponseID}$	260		
Observations	25019		
Marginal R^2	0.031		
Deviance	137390.788		

Research Question 2

Next, the effects the math level had on the change in self-efficacy were considered. The modeled equations were:

Level 1: $\Delta Self - Efficacy = \beta_{0j} + \beta_{1j} Gender_{ij} + \beta_{2j} \Delta Interest + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01} Math + \mu_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11} Math + \mu_{1j}$

$\beta_{2j} = \gamma_{20} + \gamma_{21} Math + \mu_{2j}$.

The interactions between gender and change in physics interest with math level also were considered. Change in physics interest and its interaction with math level were found to be significant predictors of the change in physics self-efficacy; see Figure 2 and Table 2. Significant variation in the intercepts also was found. Math level was a significant predictor of the relationship between a favorable change in interest and favorable change in self-efficacy.

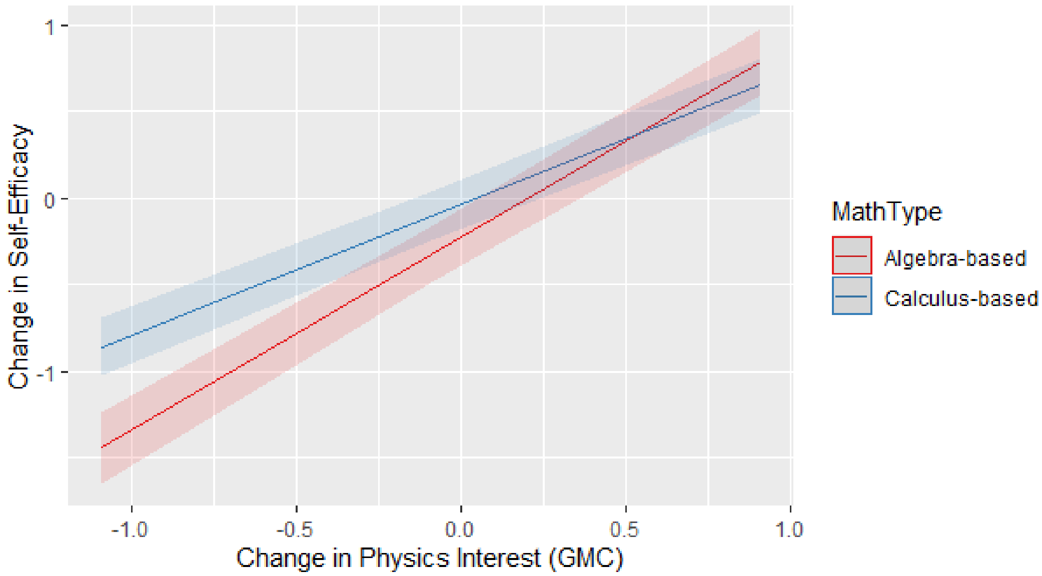

Figure 2. *Change in self-efficacy vs change in interest by math type.*

Table 2. *Intercepts as predictors for math level and level 1 predictors.*

Predictors	Estimates	Std. Error	*p*
Intercept	-0.23	0.08	.006
GenderGMC	-0.09	.09	.282
MathType	0.19	0.11	0.080
ChangeInterestGMC	-1.11	0.06	< .001
GenderGMC:MathType	-0.14	0.11	.186
MathType:ChangeInterestGMC	0.36	0.07	< .001
Random Effects			
σ^2	14.12		
τ_{00} Response ID	0.27		
Marginal R^2	.027		
Deviance	137448.606		

In both types of math, as a favorable change in interest increased, the favorable change in self-efficacy also increased. The effect was larger for algebra-based courses, however, than calculus-based courses.

Research Question 3

In this research question, the fixed and random effects of the number of labs had on the change in self-efficacy were considered. The modeled equations were:

Level 1: $Self-efficacy = \beta_{0j} + \beta_{1j} Gender_{ij} + \beta_{2j} \Delta Interest_{ij} + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01} NumLabs + \mu_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11} NumLabs + \mu_{1j}$

$\beta_{2j} = \gamma_{20} + \gamma_{21} NumLabs + \mu_{1j}$.

The number of labs per semester was not a significant predictor of change in physics self-efficacy nor did it significantly interact with gender or change in physics interest. A graph of the change in physics self-efficacy versus the number of labs per semester by gender is presented below; see Figure 3. The graph revealed that for men, the number of labs to exhibit no change in physics self-efficacy was four per semester; for females, the number of labs per semester was eight for no change in physics self-efficacy.

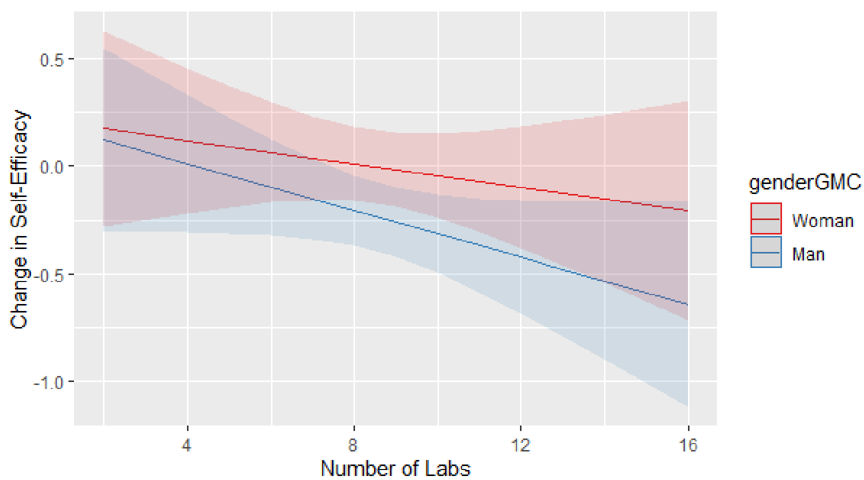

Figure 3. *Change In Self-Efficacy Vs Number Of Labs By Gender*

Research Question 4

In this question, the number of hours spent in the lab as a predictor for the change in physics self-efficacy were considered. The number of hours spent in the physics lab was not a significant predictor of the change in physics self-efficacy. The modeled equations were:

Level 1: $Self-efficacy = \beta_{0j} + \beta_{1j} Gender_{ij} + \beta_{2j} \Delta Interest_{ij} + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01} HrsLab + \mu_{0j}$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} HrsLab + \mu_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21} HrsLab + \mu_{1j}\,.$$

When gender, change in physics interest and number of hours spent in the physics lab were considered simultaneously, only gender and change and physics interest were considered significant predictors of the change in physics self-efficacy.

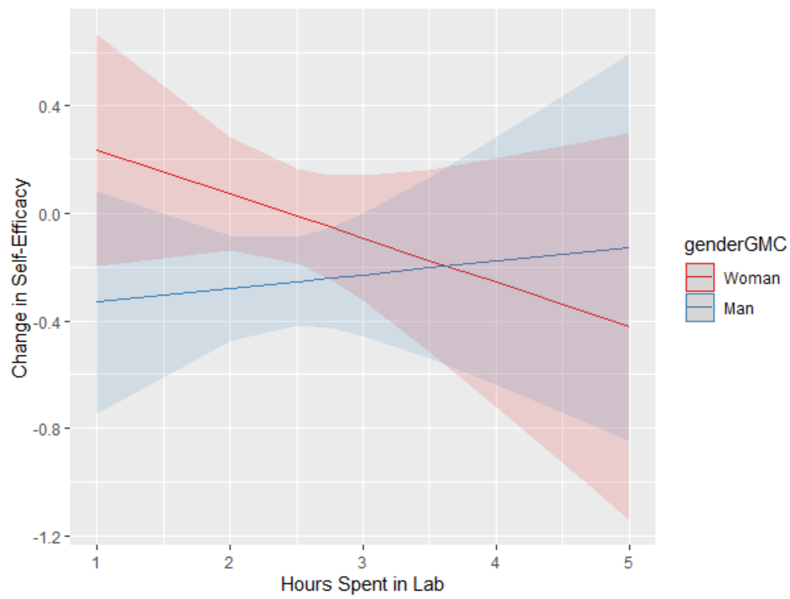

Figure 4. *Change In Self-Efficacy Vs. Hours In Lab By Gender*

An interaction was found between hours spent in the physics lab and gender; see Figure 4. This model and plot are consistent with the findings from research question 1 in which women showed a greater change in favorable self-efficacy than men, but up to a point; the effect will vary depending on the number of hours spent in the lab. When students spend more hours in lab, the gender effect (greater favorable change in self-efficacy for women than men) reverses; that is, men showed a greater favorable change in self-efficacy than women after approximately 3.5 hours in the lab. Further, at lower number of hours spent in the physics lab, women had a positive change in physics self-efficacy up to about 2.5 hours spent in the lab; after this point, the change in physics self-efficacy became negative.

DISCUSSION

This analysis presented a number of discrepant findings with the reviewed literature. The main difference was women's change in physics self-efficacy was generally superior to that of men's, in contrast to Ivie et al.'s (2016) findings. This finding depended on the number of hours spent in the lab, however. Further, this research did not support the findings of Kost-Smith et al. (2010), Kost-Smith (2011), nor Nissen and Shemwell (2016) in which they found the negative changes in self-efficacy and attitude were larger for women than men; this research found women's change in physics self-efficacy was generally superior to men's change in physics self-efficacy, but depended on the number of hours spent in the lab. Women showed a more favorable change in physics self-efficacy than men, except when they participated in more labs, contrasting with Espinosa et al.'s (2019) findings. In fact, men were found to have a more extreme loss in physics interest compared to women. This finding could be due to the loss of overconfidence Dou et al. (2016) mentioned. While the slopes for each college varied significantly, the differences were slight indicating these findings were fairly consistent across universities.

The interactions between gender and level 2 variables provided a surprise. When considering the hours spent in the physics lab, males demonstrated less of an extreme negative change in physics self-efficacy. The opposite was true for females. Women had a positive change in physics self-efficacy up until approximately 2.5 hours spent in the physics lab. A similar result was found with the number of labs per semester compared to the change in physics self-efficacy. At approximately four labs per semester, men exhibited no change in physics self-efficacy; after this point, their change in self-efficacy became negative. Likewise, women tolerated approximately 8 labs per semester with no change in physics self-efficacy. In general, women exhibited less change in physics self-efficacy with shorter hours in the lab, but tolerated more labs per semester with no change to a positive change in physics self-efficacy. In contrast, males tolerated fewer labs per semester but longer ones, with no loss in physics self-efficacy.

Finally, students in lab courses with differing math levels exhibited differences in the relationship between change in self-efficacy and change in interest. Students in calculus-based lab classes did not lose self-efficacy as much as algebra-based lab classes for a given amount of change in physics interest. Students in algebra-based physics lab courses generally had more extreme negative changes in self-efficacy compared to students in calculus-based lab courses. For instance, for no change in physics interest, students in algebra-based lab courses had a loss in self-efficacy, whereas students in calculus-based courses saw nearly no change in self-efficacy. However, for moderate increases in physics interest, students in algebra-based courses surpassed students in calculus-based classes in gain in self-efficacy.

These findings may have possible implications for designing the frequency and duration of lab experiences for students. If students' self-efficacy generally diminishes in physics, which in turn may explain the lack of retention of female students in physics, it would be beneficial to institutions to examine current lab practices. Lab activities are important; they help reinforce concepts from lecture, teach skills, and provide valuable experience (AAPT, 1992; Espinosa et al., 2019). However, being cognizant of the effects labs can have on a students' sense of self-efficacy may increase student retention, understanding of physics concepts, and provide a pleasant experience for all involved. Further, especially in algebra-based lab courses, stimulating student interest by even a moderate amount may be key to increasing student self-efficacy. It is recommended, in light of these findings, that universities investigate the optimal number of labs and time spent in lab to maximize both the learning and the affective experience of students. Examining and implementing these changes in co-ed courses may require a compromise because the tolerated number of hours in lab and the number of labs per week differ by a wide margin for males and females.

Another suggestion for future research includes examining the effects of varying the teaching method. As Brewe et al. (2009) and Kost et al. (2009) suggested, self-efficacy diminishes over time even though different pedagogies are implemented. It would be beneficial for universities to examine pedagogical techniques to determine which, if any, minimize the decline in self-efficacy and ideally improve this affective condition.

A few weaknesses should be noted as well. First, the self-efficacy measure was part of a larger instrument. These results were self-reported. Further, it was not clear when exactly the surveys were administered. For instance, if the surveys were administered after a particularly difficult lab or after a quiz or test, self-efficacy results may drop. It would also be beneficial to include other demographic information in the study such as major to determine the extent those factors account for changes in self-efficacy.

REFERENCES

AAPT (1992). *Role of labs in high school physics,* from American Association of Physics Teachers: https://www.aapt.org/resources/policy/roleoflabs.cfm#:~:text=Laboratory%20activities%20in%20high%20school,predictive%20power%20of%20their%20reasoning.

Aiken, J. (2021). *A new data set for physics education research,* from https://towardsdatascience.com/a-new-data-set-for-physics-education-research-25631745919.

Aiken, J. M., & Lewandowski, H. J. (2021). Data sharing model for physics education research using the 70 000 response Colorado Learning Attitudes about Science Survey for Experimental Physics dataset. *Physical Review Physics Education Research, 17*(2).

American Institute of Physics (AIP) (Ed.). 2003. *Conference Proceedings:* AIP.

Anderson, S. L., & Betz, N. E. (2001). Sources of social self-efficacy expectations: Their measurement and relation to career Development. *Journal of Vocational Behavior, 58*(1), 98–117.

Betz, N. E., & Hackett, G. (1986). Applications of self-efficacy theory to understanding career choice behavior. *Journal of Social and Clinical Psychology, 4*(3), 279–289.

Brewe, E., Kramer, L., & O'Brien, G. (2009). Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS. *Physical Review Special Topics - Physics Education Research, 5*(1), 121.

Brown, S. D., Lent, R. W., & Larkin, K. C. (1989). Self-efficacy as a moderator of scholastic aptitude-academic performance relationships. *Journal of Vocational Behavior, 35*(1), 64–75.

Cavallo, A. M., Potter, W. H., & Rozman, M. (2004). 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*(6), 288–300.

Code, J. (2020). Agency for Learning: Intention, Motivation, Self-Efficacy and Self-Regulation. *Frontiers in Education, 5*.

Dou, R., Brewe, E., Zwolak, J. P., Potvin, G., Williams, E. A., & Kramer, L. H. (2016). Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens. *Physical Review Physics Education Research, 12*(2), 1.

Espinosa, T., Miller, K., Araujo, I., & Mazur, E. (2019). Reducing the gender gap in students' physics self-efficacy in a team- and project-based introductory physics class. *Physical Review Physics Education Research, 15*(1), 1.

Goodman, I. F., Cunningham, C. M., Lachapelle, C., Thompson, M., Bittinger, K., Brennan, R. T., & Delci, M. (2002). *Final report of the women's experiences in college engineering (WECE) project.* Cambridge, MA: Goodman Research Group, Inc, from https://files.eric.ed.gov/fulltext/ED507395.pdf.

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching, 2*(6).

Hox, J. J., & Maas, C. J. M. (2002). Sample sizes for multilevel modeling. In J. Blasius, J. Hox, E. de Leeuw, & P. Schmidt (Eds.), *Social Science Methodology in the New Millennium* (2nd ed.). Opladen, RG: Leske + Budrich Verlag.

Ivie, R., White, S., & Chu, R. Y. (2016). Women's and men's career choices in astronomy and astrophysics. *Physical Review Physics Education Research, 12*(2).

Kalender, Z. Y., Marshman, E., Nokes-Malach, T., Schunn, C., & Singh, C. (2018). Large gender differences in physics self-efficacy at equal performance levels: A warning sign. In *Proceedings of the 2018 Physics Education Research Conference* .

Kost, L. E., Pollock, S. J., & Finkelstein, N. D. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics - Physics Education Research*, *5*(1), 288.

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2010). Gender disparities in second-semester college physics: The incremental effects of a "smog of bias". *Physical Review Special Topics - Physics Education Research*, *6*(2), 225.

Kost-Smith, L. E. (2011). *Characterizing, modeling, and addressing gender disparities in introductory college physics.* Dissertation, University of Colorado, Boulder.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*(1), 79–122.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1993). Predicting mathematics-related choice and success behaviors: Test of an expanded social cognitive model. *Journal of Vocational Behavior, 42*(2), 223–236.

Lindstrøm, C., & Sharma, M. (2011). Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter? *Int. Journal of Innovation in Science and. Mathematics*, *19*(1).

Marshman, E. M., Kalender, Z. Y., Nokes-Malach, T., Schunn, C., & Singh, C. (2018). Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm? *Physical Review Physics Education Research*, *14*(2), 82.

Martin, J. (2004). Self-regulated learning, social cognitive theory, and agency. *Educational Psychologist*, *39*(2), 135–145.

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267.

National Center for Education Statistics (NCES) (2017). *Digest of education statistics: 2016 Tables and figures: Table 318.30: Bachelor's, master's, and doctor's degrees conferred by postsecondary institutions, by sex of student and discipline division: 2014-15.*

National Science Foundation (NSF) (2017). *Women, minorities, and persons with disabilities in science and engineering.* Washington, D. C.: National Science Foundation.

Nissen, J. M., & Shemwell, J. T. (2016). Gender, experience, and self-efficacy in introductory physics. *Physical Review Physics Education Research*, *12*(2), 1–16.

Pajares, F. (2004). Gender differences in mathematics self-efficacy beliefs. In *Gender Differences in Mathematics* (1st ed., pp. 294–315). Cambridge University Press.

S.1701 - 115th Congress (2017-2018) (2017). Fair access to science and technology research act of 2017.

Sawtelle, V., Brewe, E., Kramer, L. H., Singh, C., Sabella, M., & Rebello, S. (2010). Positive impacts of modeling instruction on self-efficacy. In : *AIP Conference Proceedings, PERC Conference Proceedings* (pp. 289–292). AIP.

Shunter (2022). *Dissemination and sharing of research results | NSF - National Science Foundation.* Retrieved April 06, 2022.446Z, from https://www.nsf.gov/bfa/dias/policy/dmp.jsp.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.

Wang, X. (2013). Why students choose STEM majors. *American Educational Research Journal, 50*(5), 1081–1121.

Zeldin, A. L., Britner, S. L., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 45*(9), 1036–1058.

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37*(1), 215–246.

Zimmerman (2000). Self-efficacy: An essential motive to learn. *Contemporary educational psychology, 25*(1), 82–91.