INVESTIGATING FEMALE AND MALE STUDENTS' MOTIVATIONAL CHARACTERISTICS AND PERFORMANCE IN INTRODUCTORY PHYSICS

by

Z. Yasemin Kalender

B.S., Boğaziçi University, 2012

M.Sc., University of Pittsburgh, 2016

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Z. Yasemin Kalender

It was defended on

June 6, 2019

and approved by

Dr. Robert Devaty, Associate Professor, Department of Physics and Astronomy

Dr. Adam Leibovich, Professor, Department of Physics and Astronomy

Dr. Russell Clark, Senior Lecturer, Department of Physics and Astronomy

Dr. Larry Shuman, Professor, Department of Industrial Engineering

Thesis Advisor/Dissertation Director: Dr. Chandralekha Singh, Professor, Department of Physics and Astronomy

INVESTIGATING FEMALE AND MALE STUDENTS' MOTIVATIONAL CHARACTERISTICS AND PERFORMANCE IN INTRODUCTORY PHYSICS

Zeynep Yasemin Kalender, PhD

University of Pittsburgh, 2019

Physics is one of the science, technology, engineering, and mathematics (STEM) disciplines in which the participation of women and ethnic/racial minorities is unacceptably low. The representation gaps among certain groups of students in physics have motivated education researchers and practitioners to understand and address the underlying reasons of low diversity in physics. Although there has been an increasing effort in the physics community to increase the participation of women, only 20% of the all bachelors' degrees in physics across U.S. is currently earned by women. Lack of prior preparation, low encouragement from mentors, gender-based ability attributions, and stereotype threat are some of the proposed reasons that can undermine women's sense of belonging and self-efficacy and lead to have higher anxiety and self-doubt in their physics content knowledge and related skills. In this thesis, I focused on understanding the gender differences in students' motivational characteristics in introductory physics courses for STEM majors and how they vary across time by investigating the relation between gender, physics performance, and motivational characteristics. In particular, I examined how physics self-efficacy, interest, intelligence mindset views (e.g., whether intelligence is innate or can grow with effort) and getting recognition can predict students' physics course performance overall and on standardized conceptual physics tests after controlling for their prior academic preparation (e.g., SAT scores, AP exam, High School GPA) and their demographics (gender, ethnicity/race). Since physics is one of the pillar courses in STEM programs and first-year college experiences in these courses are significantly influential for students' retention in STEM, I also examined the physics identity of students as STEM majors in the introductory physics courses. Specifically, I have studied the female and male students' perception of how their peers and teaching assistant/instructors recognize them as someone who is good at physics. In all of the studies I have conducted, I have found large gender differences in the motivational constructs examined which created gendered-patterns in students' identity formation as STEM majors. The implication of these results for how to enhance women's participation and achievement in STEM disciplines in which they are severely underrepresented is discussed throughout the thesis.

Table of Contents

List of Tables

List of Figures

Preface

First, I would like to thank my research advisor, Dr. Chandralekha Singh. Dr. Singh has always been generous with her time for our research discussions and has been available whenever I needed her help and advice. Her expert guidance has helped me to become a more independent and confident researcher. Dr. Singh served as an extremely supportive mentor at this critical time in my life, always respecting new research suggestions. Thanks to her careful guidance, I always felt comfortable posing new models to test, which I recognize as an important step in becoming a successful researcher. I would also like to thank Dr. Singh for being a great female role model for me as a physicist, and as a successful researcher in our field.

Next, I would like to express my sincere appreciation to my co-advisors, Dr. Christian Schunn and Dr. Tim Nokes-Malach, for their never-ending support and constructive feedback. Their contributions to my Ph.D. training have been indispensable. They always provided positive endorsements of my capabilities, particularly when I struggled with learning statistical tools. Their uplifting energy and enthusiasm served to continually motivate me and keep me focused.

I would like to thank Dr. Robert Devaty for his valuable time and help with my thesis. His constructive feedback on my papers and thesis has greatly improved my writing skills and contributed a lot to my academic work.

I am also thankful to Dr. Emily Marshman, an excellent role model, supportive mentor, and amazing friend. Our lengthy discussions and anecdotal conversations over coffee about the issue of gender representation and the chilly environment in physics contributed enormously to shaping my research theories. Emily is a great researcher and much-loved physics instructor: it is very rare to have both qualities. I am so honored to be her mentee, colleague, and friend!

There is no way to express my feelings and gratitude to Maryam Kohram –the first friend that I made in the U.S.A. As two female physicists, living away from family, and sharing similar customs, we strengthened our commitment to our Ph.D. journeys despite difficulties and challenges. My friendship with Maryam was what kept me from missing my family, culture, and country—basically everything I left behind. It is hard to find a true friend who stands beside you and makes you feel strong and not alone. There is no doubt that without her, I could not have survived the past four years (Thanks M!).

I also want to thank another special friend, Wendy Bennett, for her earnest support during my Ph.D. life. She has always found a way to uplift my mood. During the last ten months when writing my thesis and searching for a job, I felt especially empowered by Wendy's contagious joy and positive energy. She is and will always be one of the strong women in my life who constantly motivate me to do more. I will always remember my surprise and joy when I found the book—*"Women Don't Ask: Negotiation and the Gender Divide"*, in my mailbox sent by Wendy, which I have read many times and cited in my thesis. I am beyond being grateful that she has embraced me as her Turkish daughter and become my family in the USA.

Very special and hearty thanking goes to my college friend since 2006—from the moment we sit on that same English preparation class of college until now and future: my beloved "dostum" Büşra İnce. She has, as always, wholeheartedly supported my achievements and endeavors during my Ph.D. chapter of my life from miles away and regardless of time difference.

Finally, I would like to thank my family members (too many to list here). My beloved parents play the most significant role in the last 24 years of my academic journey. They have stood behind my decisions even if it meant moving to a country beyond the ocean, for at least five years—a great sacrifice on their part. I am more than proud to be the only child of such wonderful parents. I also want to recognize my aunt, Rabia Inci Dikmen, who always told me that I could do anything that I desired to. She was always there to cheer me, hug me and support me during my stint as a professional tennis player or as a physicist. My only sadness is that today she is not here with me, to see me earn my Ph.D. Rest in peace 'tetecim'.

I dedicate this thesis to my family, friends, and all women in STEM fields who, through their amazing work, dedication and persistence, prove every day that that they have what it takes to be successful in this field.

1.0 Introduction

The representation of women in many science, technology, engineering, and mathematics (STEM) fields is unacceptably low. Although there have been several efforts to increase women's involvement in STEM majors and jobs, unequal gender participation among students majoring in a number of STEM fields persists with little change [1]. There is greater participation of women in biology, chemistry, industrial engineering, and chemical engineering, but physics, computer science, and electrical engineering have made little progress in increasing the representation of women [1,2].

In the US, the discipline of physics has had a low level of participation from women at all educational levels [1,2]. Even though the number of female students enrolling in regular high school physics classes has increased in the past several decades, female students are less likely to enroll in more advanced high-school physics courses (AP physics B and C) [2]. Similarly, the number of women majoring in physics has increased since the 1980s, but the percentage of all bachelor's degrees in physics earned by women is still only around 20% [2]. The underrepresentation of women in physics exhibits similar trends in Ph.D. degrees. This pervasive issue surrounding the lack of gender diversity in physics—beginning at the high school level and becoming increasingly salient in college, advanced degrees, and academic leadership positions—is of concern to the physics community. Despite existing efforts to understand low diversity in physics [2], the reasons for women's low enrollment and participation in physics are not fully understood. In recent years, the investigation of origins of the gender gap in physics and the effective approaches to creating a more inclusive and equitable learning environment have gained traction in both the physics education and broader science education research communities. The specific focus of this thesis is examining students' motivational characteristic differences by gender in the foundational college-level introductory calculus-based (and to a lesser extent algebra-based) physics courses. These introductory level physics courses are one of the sets of fundamental pillar courses in students' academic program, so investigating students' attitudes in these physics courses is a matter of great importance to identify the experiences of under-represented groups, such as women.

1.1 Factors that can impact the underrepresentation of women in STEM fields

Deciding to pursue a STEM field and continuing to remain in the field are generally affected by several interrelated factors, including students' prior preparation and skills, quality of teaching and type of teaching approach, sociocultural factors, and motivational factors [3-11]. Researchers have hypothesized that these factors affect women's choices to pursue STEM degrees that produce the underrepresentation of women in these fields [7,12-14]. In particular, enrollment and proficiency in math and science courses in high school have a positive impact on students' choice of an engineering major and their persistence in the major [11,15]. However, high school girls are significantly less likely than boys to enroll in engineering courses and majors [16].

Furthermore, the quality of teaching can also affect students' motivation to persist in engineering fields. Experiences in freshmen science courses are influential in students' decision to switch out of the major [17], and a majority of students who switched out of science-related majors said that they were concerned about the poor quality of teaching in their science courses [8]. However, teaching approaches such as the use of real-world and challenging tasks, cooperative learning groups, and evaluation practices that emphasize effort and improvement over normative ability can foster student motivation to persist in STEM fields [18-20]. Some research suggests that in-class active-learning activities and “flipped courses” can reduce the gender gap in performance and improve student attitudes and motivation [21,22]. However, there are also other studies indicating that the performance gap between female and male students increases in active-engagement physics courses [23].

In addition, sociocultural factors can also affect students’ career choices [10]. Stereotypes about engineering, e.g., that engineers are stereotypically male or that engineering is “thing-oriented” as opposed to “people-oriented” can deter female students from choosing to major in and persist in engineering fields [23]. In addition, since women are underrepresented in STEM disciplines, they may experience stereotype threat (i.e., the fear of confirming negative stereotypes about one’s group). Negative stereotypes about women’s abilities (e.g., that men perform better than women in STEM courses) can discourage them from majoring in STEM fields and reduces their sense of belonging in STEM [24-26].

1.2 Students’ motivational factors in STEM learning

Motivational factors, such as intelligence mindset, interest in science, and self-efficacy in science can also influence students’ motivation to major and persist in STEM fields. For example, when individuals believe that intelligence is an innate, unmodifiable trait, they become frustrated when confronted with challenging tasks, give up more easily, and attribute their struggles to a lack of talent [27,28]. Research also suggests that female students who view intelligence as a fixed trait and experience stereotype threat in STEM courses may have decreased motivation and interest in pursuing STEM careers [27,28]. Furthermore, interest in science can play a role in students’ career aspirations, and students leave engineering fields partly due to a decreased interest in engineering [29]. Research suggests that from early adolescence, girls express less interest in math or science careers than boys do [13,30].

Interest in a particular field is also influenced by one's self-efficacy [31]. Self-efficacy is the belief in one's capability to be successful in a particular task, course, or subject area [32-34] and it can impact one's interests [35]. Because self-efficacy can shape interest, it also influences engagement during learning [36]. Furthermore, students with high self-efficacy in a domain often enroll in more difficult courses in that domain than those with low self-efficacy because they perceive difficult tasks as challenges rather than threats [37]. Self-efficacy in STEM also predicts initial career goals and enrollment in STEM courses [27,28] as well as persistence toward long-term career goals [35,38].

Self-efficacy in particular can have large long-term effects because of its role in feedback loops. Obviously, as students complete short-term goals (e.g., complete a physics course requirement for an engineering major), they obtain feedback about their performance that shapes their self-efficacy. More surprisingly, the reverse effect also occurs: for a variety of reasons, self-efficacy beliefs constrain performance in science courses beyond the more normative effects of students' prior knowledge and skills [39]. For example, students with high self-efficacy are more likely to exhibit effective learning strategies such as self-monitoring [35,40 ,41] and tend to make more efficient use of problem-solving strategies and time-management [35]. In addition, students with high self-efficacy are less likely to reject correct hypotheses prematurely and tend to be better at solving conceptual problems than students with low self-efficacy but equal performance [42]. As a result, there are feedback loops in which higher initial self-efficacy produces higher performance which further strengthens self-efficacy; by contrast, lower initial self-efficacy produces lower performance which further weakens self-efficacy.

Unfortunately, several studies have shown that female students have significantly lower self-efficacy than male students in STEM-related domains [10,43], including mathematics, engineering [3,13,44], computer science [45,46], and physics [47-55]. Research has shown that female freshmen engineering students exhibit lower self-confidence about their background knowledge and skills, problems solving abilities, and overall engineering abilities, and have negative perceptions about how engineers contribute to society than did male freshman engineering students [3]. Furthermore, female students experience a decline in self-efficacy throughout their engineering education at college [44]. In our preliminary study, we investigated the self-efficacy of engineering students in different domains (mathematics, chemistry, physics and engineering). We find that physics self-efficacy shows a larger gender gap compared to other domains (see Figure 1-1). A strong sense of self-efficacy, especially for female students in engineering courses, can help them persist on an engineering track and become practicing engineers [56-60]. Given the central role of self-efficacy in career choice and career persistence, the self-efficacy gap likely contributes to the low representation of female students pursuing careers in engineering fields. Therefore, understanding the source and nature of the gap is important.

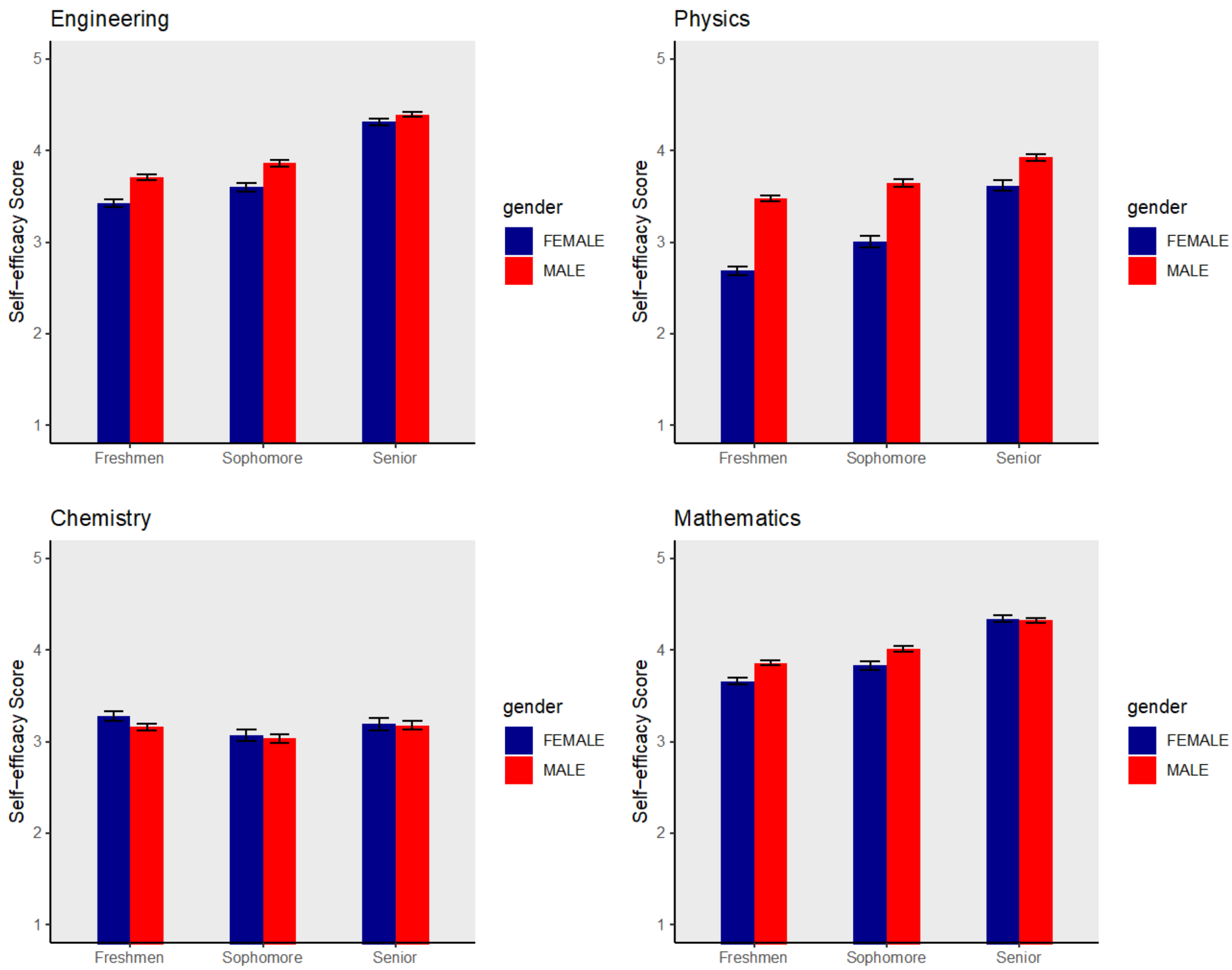

Figure 1-1 The mean self-efficacy scores of engineering students at the end of their freshman, sophomore, and senior years in each of the foundational subjects in engineering are plotted along with their standard error. Self-efficacy was measured on a Likert scale from 1 to 5.

One major focus of this thesis is investigating the impact of motivational factors on students' learning for both men and women in introductory physics courses and how their motivational characteristics change over the course of an introductory physics sequence. The results of this thesis can be a stepping stone for developing better learning tools and determining more effective ways to implement learning tools in order to increase student engagement and success in physics courses. One primary goal of this thesis is to determine the motivational factors for which there are gender differences and when those differences are present (e.g., at the beginning of a physics course sequence or later). Studies in this thesis can also inform future instruction and interventions. For example, if there are differences between male and female students' interest at the beginning of a physics course sequence, interventions may involve strategies for boosting female students' interest in physics early on in the course (e.g., by using examples that are more meaningful to them).

1.3 Investigating students' motivational characteristics in the introductory physics courses

This thesis consists of several studies conducted to understand the motivational characteristics of students in the introductory physics courses. The first study presents a framework for helping all students learn in science courses that takes into account four factors: 1) the characteristics of instruction and learning tools, 2) implementation of instruction and learning tools, 3) student characteristics, and 4) the students' environments. While there has been much research on factor 1 and 2 (characteristics of instruction and learning tools), there has been less focus on factor 3 (students' characteristics, and in particular, motivational factors). In this study, a longitudinal analysis of students' motivational characteristics in two-semester introductory physics courses was performed by administering pre- and post-surveys that evaluated students' self-efficacy, grit, fascination with physics, value associated with physics, intelligence mindset, and physics epistemology. I found that female students reported lower levels of self-efficacy, fascination and value associated with physics, and held a more "fixed" view of intelligence in the context of physics compared to male students. Female students' fascination and value associated with physics decreased significantly more than males' after an introductory physics course sequence. In addition, female students' view of physics intelligence became more "fixed" compared to male students' by the end of an introductory physics course sequence. Grit was the only factor for which female students reported averages that were equal to or higher than those of male students throughout introductory physics courses. The findings inform the framework and have implications for the development and implementation of effective pedagogies and learning tools to help all students learn.

The chapter three focuses on self-efficacy and how it can affect performance, career goals, and persistence. This study examines the self-efficacy of male and female students with similar performance in introductory physics courses and investigates whether gender gaps in self-efficacy are persistent across different instructors and course formats. Students filled out a self-efficacy in physics survey before Physics 1, before Physics 2, and at the end of Physics 2. Students' achievement was measured by their performance on research-based conceptual physics tests and course grades. The physics courses were taught by several instructors and varied in the type of pedagogy used, with some using a "flipped" format and others using a traditional, lecture-based format. I found that female students had lower self-efficacy than male students at all performance levels in both Physics 1 and Physics 2. The self-efficacy gaps continued to grow throughout the introductory physics course sequence, regardless of course format (i.e., traditional or flipped) and instructor. The findings suggest that female students' self-efficacy was negatively impacted by their experiences in introductory physics courses, and this result is persistent across various instructors and course formats. Female students' lower self-efficacy compared to similarly performing male students can result in detrimental short-term and long-term impacts.

The chapter four focuses on the female and male students' self-efficacy and its relation to learning outcomes in physics. In a longitudinal-study, we surveyed approximately 600 students in calculus-based Physics 2 courses to investigate students' attitudes in physics. I examined female and male students' self-efficacy scores and the extent to which self-efficacy mediates the relation between gender and learning outcomes. There were also several other control variables: AP physics test scores, SAT math, and conceptual pre-test results in physics. I found gender effects on learning outcomes by the model variables. In particular, initial self-efficacy differences show direct effect on outcomes even when controlling for students' prior physics knowledge and skill differences, and self-efficacy also has the strongest total effect in explaining the relationship between gender and learning outcomes.

The next study in chapter five investigates gender differences in students' physics identity in introductory physics courses. Exploring the components that influence these identities is critical to developing a better understanding of the under-representation of women in physics courses and physics-related majors. Using the physics identity framework, I investigated whether the relation between gender and physics identity was mediated by motivational factors, such as self-efficacy, interest, and perceived recognition by others. We surveyed approximately 500 students in introductory level calculus-based physics courses in which 30% of the students are women. Analysis revealed that the relation between gender and physics identity was mediated by students' self-reported motivation at the end of the semester. The model showed that perceived recognition by others played a major role in students' endorsement of physics identity with female students less likely to endorse statements that others perceived them as a "physics person." Individual interviews with the first-year students provided further evidence that unlike male students, female students' experiences in physics courses may reinforce their perception that they are not likely to be recognized by their instructors, teaching assistants, and peers as someone who is "good" at physics.

In a similar study in chapter six, I examined the gender differences in how students identify themselves as a physics person and how their perceived recognition from others, such as their TAs/instructors, peers or family members explains their physics identity. I tested a model that examined how different motivational constructs relevant for physics classrooms can explain students' physics identities by gender. I found that the perception of being recognized by the influential others such as the course instructor or teaching assistants was differentially related to female and male students' sense of belonging and self-efficacy in the physics classroom. These findings have implications for strategies to improve female students' sense of belonging and self-efficacy in the physics classroom so that they can develop a stronger physics identity.

Despite the existing studies on motivational characteristics of students in physics, views of intelligence mindset (fixed versus growth mindset) in the context of physics have been rarely examined. Considering that physics is one of the few STEM disciplines that has been commonly associated with being brilliant or "gifted" man to be successful, students might carry fixed mindset views in physics courses, which can be especially detrimental for minority groups' achievement. Therefore, in chapter seven, I also examined physics mindset views in the calculus-based physics courses and contrasted these views between female and male students. In my mindset framework, I divided students' mindset views into four group (My growth, My fixed ability, Others' growth and Others' fixed ability) to identify the specific mindset views that can largely explain students' course performance in physics. I found that underrepresentation and underperformance of women in physics courses is correlated with specific mindset view, which can further impact their achievement and persistence in the field. Physics mindset can be important aspect in students' course achievements and for designing certain interventions that can help improve learning of diverse groups of students in physics.

In the final chapter, I conducted a longitudinal analysis of students' motivational characteristics in introductory physics courses across three racial/ethnic groups: White, Asian, and Underrepresented Racial/ethnic Minorities (URM). In this study, the three racial/ethnic groups' self-efficacy and interest in physics are reported, and implications for instruction are discussed.

1.4 Chapter references

1. National Science Board Science and Engineering Indicators. National Science Foundation Report, retrieved from website: https://www.nsf.gov/statistics/2018/nsb20181/digest/sections/preface

2. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016

3. Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, Journal of Research in Science Teaching **47**, 978 (2010).

4. M. Besterfield-Sacre, M. Moreno, L. Shuman and C. Atman, Gender and ethnicity differences in freshmen engineering student attitudes: A cross-institutional study, J. Eng. Educ. **90**, 477 (2001).

5. C. Seron, S. Silbey, S. Cech and B. Rubineau, Persistence is cultural: Professional socialization and the reproduction of sex segregation, Work and Occupations **43**, 178 (2016).

6. 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).

7. A. W. Astin, *What Matters In College: Four Critical Years Revisited*. San Francisco, CA: Jossey-Bass (1993).

8. N.M. Hewitt and E. Seymour, A long, discouraging climb, ASEE Prism **1**, 24 (1992).

9. E. Seymour and N.M. Hewitt, *Talking About Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO: Westview Press (1997).

10. S. Tobias, *They're Not Dumb, They're Different: Stalking the Second Tier*, Tuscon, AZ: Research Corporation (1990).

11. S. Cheryan, S. Ziegler, A. Montoya and L. Jian, Why are some STEM fields more gender balanced than others? Psychological Bulletin (2016). doi: http://dx.doi.org/10.1037/bul0000052

12. M. Wang and J. Degol, Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields, Dev. Rev. **33**, 304 (2013).

13. S. Cheryan and V. Plaut, Explaining underrepresentation: A theory of precluded interest, Sex Roles **63**, 475 (2010).

14. C. Hill, C. Corbett and A. Rose, Why so few? Women in science, technology, engineering, and mathematics (2010). Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf.http://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2010.tb01066.x/epdf

15. R. Felder, G. Felder, M. Mauney, C. Hamrin and E. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Engineering Education **84**, 151 (1995).

16. E. Lichtenberger and C. George-Jackson, Predicting high school students' interest in majoring in a STEM field: Insight into high school students' postsecondary plans, Journal of Career and Technical Education **28**, 19 (2013).

17. C. Nord et al., The nation's report card: America's high school graduates (NCES 2011-462). Washington, CD: U.S. Government Printing Office (2011). Retrieved from https://files.eric.ed.gov/fulltext/ED518324.pdf

18. A. Strenta, R. Elliot, R. Adair, M. Matier and J. Scott, Choosing and leaving science in highly selective institutions, Res. High. Educ. **35**, 513 (1994).

19. E. Pascarella and P. Terenzini, *How College Affects Students: Findings and Insights from Twenty Years of Research*, San Francisco, CA: Jossey-Bass (1991).

20. A. Wigfield et al. Motivation to succeed. In Damon, W. and Eisenberg, N (Eds.) *Handbook of Child Psychology Vol.* ***3*** *Social Emotional, and Personality Development*, New York: Wiley (2006).

21. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

22. R. Beichner et al., The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project retrieved from http://www.per-central.org/per_reviews/volume1.cfm? (2007).

23. N. Karim, A. Maries and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics? Eur. J. Phys. **39**, 025701 (2018).

24. A. Diekman, E. Brown, J. Johnston and E. Clark, Seeking congruity between goals and roles: A new look at why women opt out of STEM careers, Psychological Science **21**, 1051 (2010).

25. J. Smith, C. Sansone and P. White, The stereotyped task engagement process: The role of interest and achievement motivation. J. Educ. Psy. **99**, 99 (2007).

26. S. Ahlqvist, B. London and L. Rosenthal, Unstable identity compatibility: How gender rejection sensitivity undermines the success of women in science, technology, engineering and mathematics fields, Psychological Science **24**, 1644 (2013).

27. C. Dweck, Is math a gift? Beliefs that put females at risk. In S. Ceci and W. Williams (Eds.), *Why Aren't More Women in Science? Top Researchers Debate the Evidence* (pp. 47-55). Washington, DC: APA Press (2007).

28. C. Dweck, *Mindsets and Math/Science Achievement*, New York: Carnegie Corporation of New York, Institute for Advanced Study, Commission on Mathematics and Science Education (2008).

29. I. Goodman, C. Cunningham and C. Lachapelle, The Women's Experience in College Engineering. Retrieved from https://files.eric.ed.gov/fulltext/ED507395.pdf (2002).

30. R. Lapan, A. Adams, S. Turner and J. Hinkelman, Seventh graders' vocational interest and efficacy expectation patterns, Journal of Career Development **26**, 215 (2000).

31. J. Eccles, Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices, Psychology of Women Quarterly **18**, 585 (1994).

32. A. Bandura, Self-efficacy. In R. J. Corsini (Ed.), Encyclopedia of Psychology (2nd ed., Vol. **3**, pp. 368-369). New York: Wiley (1994).

33. A. Bandura, *Self-efficacy: The Exercise of Control*. New York: Freeman (1997).

34. F. Pajares, Gender differences in mathematics self-efficacy beliefs. In A.M. Gallagher and J.C. Kaufman (Eds.), *Gender Differences in Mathematics: An Integrative Psychology Approach* (p. 294). New York: Cambridge University Press (2005).

35. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemporary Educational Psychology **25**, 82 (2000).

36. D. Schunk and F. Pajares, *The Development of Academic Self-efficacy*. In A. Wigfield and J. Eccles (Eds.), Development of Achievement Motivation: A Volume in the educational psychology series (p. 15). San Diego, CA: Academic Press (2002).

37. H. Watt, The role of motivation in gendered educational and occupational trajectories related to maths, Educational Research and Evaluation **12**, 305 (2006).

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

39. S. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, Journal of Research in Science Teaching **43**, 485 (2006); S.L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, Journal of Research in Science Teaching, **45**, 955 (2008).

40. A. Bandura, Social cognitive theory of self-regulation, Organizational Behavior and Human Decision Processes **50**, 248 (1991).

41. P. Pintrich, A motivational science perspective on the role of student motivation in learning and teaching contexts, Journal of Educational Psychology **95**, 667 (2003).

42. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students, International Journal of Behavioral Development **14**, 153 (1991).

43. P. Huang and S. Brainard, Identifying determinants of academic self-confidence among science, math, engineering, and technology students, Journal of Women and Minorities in Science and Engineering **7**, 315 (2001).

44. S. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, J. Eng. Educ. **87**, 369 (1998).

45. L. Irani, Understanding gender and confidence in CS course culture. Paper presented at the SIGCSE 2004, Norfolk VA. Retrieved from http://ai2-s2-pdfs.s3.amazonaws.com/f5c2/cd1620808de342808613cd0cd4a67dbd8931.pdf (2004).

46. M. Rosson, J. Carroll and H. Sinha, Orientation of undergraduates toward careers in the computer and information sciences: Gender, self-efficacy and social support, ACM Transactions on Computing Education **11**, 1 (2011).

47. 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).

48. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1069 (2012).

49. 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).

50. 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).

51. 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).

52. 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?, Phys. Rev. Phys. Educ. Res. **14**, 020123 (2018).

53. T. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, in the *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

54. 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, *Proceedings of the 2018 Physics Education Research Conference* (2019) [Washington, DC, August 1-2, 2018], edited by A. Traxler, Y. Cao, and S. Wolf.

55. 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?, *Proceedings of the 2018 Physics Education Research Conference* (2019) [Washington, DC, August 1-2, 2018], edited by A. Traxler, Y. Cao, and S. Wolf.

56. R. Marra, K. Rodgers, D. Shen and B. Bogue, Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy, J. Eng. Educ. **98**, 27 (2009).

57. J. Raelin, M. Bailey, J. Hamann, L. Pendleton, R. Reisberg and D. Whitman, The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention, J. Eng. Educ. **103**, 599 (2014).

58. O. Eris et al., Outcomes of a longitudinal administration of the persistence in engineering survey, J. Eng. Educ. **99**, 371 (2010).

59. K. Meyers, S. Silliman, N. Gedde and M. Ohland, A comparison of engineering students' reflections on their first-year experiences, J. Eng. Educ. **99**, 169 (2010).

60. B. M. Jones, Pursuing diversity at state-supported residential STEM schools, NCSSSMST Journal **16**, 30 (2010).

2.0 A longitudinal analysis of students' motivational characteristics in introductory physics courses: gender differences

2.1 Introduction

Women are underrepresented in many STEM courses [1], and this is especially true in the field of physics [1]. Furthermore, prior research suggests that women generally under-perform in introductory physics courses [2-10]. Some efforts have been made to increase the diversity in physics courses [11-23], yet, the reasons for the low percentages and the under-performance of women in physics are not yet fully understood. Some researchers have hypothesized that the reasons may include, e.g., their prior preparation, career goals, self-efficacy, sense of belonging, mindset, and epistemology [24-63]. Increasing the diversity in the field of physics hinges, in part, on taking student characteristics into account in instructional design and implementation to improve teaching and learning for all students in physics courses.

In the past few decades, physics education researchers have investigated the challenges students face in learning physics and developed research-based instructional tools to improve student understanding of physics and their problem-solving and reasoning skills [64-76]. Indeed, effective pedagogies and learning tools build on students' prior knowledge and skills and provide appropriate coaching and support to help all students learn [77-81]. However, there has been less focus on how other factors such as aspects of students' motivation, students' environments, and the implementation of instructional strategies affect the efficacy of instructional interventions. Here, we first describe a holistic framework for engaged learning that not only takes into account students' cognitive skills but also incorporates their motivational beliefs (e.g., their self-efficacy, intelligence mindsets, and epistemologies), their environments, and the implementation of the pedagogies and learning tools to help all students learn physics. We then focus on the baseline measures of motivation of introductory physics students (in particular, women) obtained from survey data. This baseline data on students' motivation can aid in understanding the interaction between instructional interventions and aspects of motivation.

2.2 Background and theoretical framework informing the investigation of students' motivational characteristics

Below, we describe the theoretical framework, Strategies for Engaged Learning Frameworks (SELF) [82] and how prior research informs the framework. The Strategies for Engaged Learning Framework (SELF) (see Figure 2-1) is a holistic framework which suggests that instructional design and learning tools, their implementation, student characteristics, and social and environmental factors collectively determine how effectively students engage with and learn from instruction in a particular course. The framework consists of four quadrants as shown in Figure 2-1 and all of them must be considered holistically to help a diverse group of students learn effectively. The horizontal dimension involves the characteristics of learning tools (e.g., how the tool provides efficiency and innovation in learning and builds on students' prior knowledge) and students (e.g., the students' prior preparation and motivational characteristics), which both are taken into account when developing effective learning tools. The vertical dimension involves internal and external characteristics of the learning tools and the students. This dimension focuses on how the characteristics of the learning tools and students as well as the environments in which the tools are implemented are important to consider when helping students engage with and learn from instruction. The internal characteristics of the tool pertain to the tool itself (e.g., whether it includes formative assessment) and the external characteristics of the tool pertain to how the tool is implemented (e.g., whether the tool is framed appropriately to get student buy in). The internal characteristics of the students pertain to, e.g., their prior preparation, motivation, goals, and epistemological beliefs. The external characteristics of the students pertain to social and environmental factors such as support from mentors and balance of coursework. The SELF framework can be used as a guide in thinking holistically about students, learning tools, and the environment to inform instruction. Below we describe how the SELF framework is informed by cognitive theories and prior empirical research.

	Learning Tool characteristics	**Student Characteristics**
Internal	Factor 1. Learning tool characteristics (internal) – pertaining to features embedded in the learning tools that help students learn □ Based on "cognitive apprenticeship model" to promote mastery of material for a variety of students □ Include material providing scaffolding support □ Involve efficiency and innovation in learning □ Incorporate elements of productive engagement and productive struggle □ Involve formative assessment	Factor 2. Student characteristics (internal) □ Prior knowledge and skills o Prior preparation o Cognitive / metacognitive skills □ Motivational and affective factors o Goals o Interest and value o Self-efficacy o Epistemology beliefs o Intelligence mindset o Grit □ Self-regulation
External Characteristics	Factor 3. Learning tool characteristics (external) – pertaining to how the tool is implemented in a particular course □ Embed features to frame the importance of learning from tools and to get student buy in □ Embed motivational features within tools conducive to effective learning □ Reinforce learning by coupling learning of different students via creation of learning communities □ Make explicit connection between in-class lessons and out of class assignments and assessments □ Incentivize students to engage with tools via grades and other motivational factors □ Support to help students manage their time better □ Support to improve students' self-efficacy and epistemological beliefs	Factor 4. Student characteristics (external) - pertaining to the student-environment interaction □ Collaboration skills □ Balance of coursework and/or work □ Family encouragement and support □ Support and mentoring from advisors and counselors □ Time-management o Minimizing unimportant activities that appear urgent (e.g., socializing) o Maximizing important activities that may not appear to be urgent (e.g., learning physics)

Figure 2-1 Strategies for Engaged Learning Framework (SELF).

2.2.1 Factors 1 and 2: Internal characteristics of learning tools and students

Factors 1 and 2 of the framework (the internal characteristics of the learning tool and students) are informed by several cognitive theories that point to the importance of knowing students' prior knowledge and difficulties in order to develop effective instructional tools. For example, Hammer proposed a "resource" model that suggests that students' prior knowledge, including their learning difficulties, should be used as a resource to help students learn better [78]. Similarly, the Piagetian model of learning emphasizes an "optimal mismatch" between what the student knows and is able to do and the instructional design [79]. In particular, this model focuses on the importance of knowing students' skill levels and reasoning difficulties and using this knowledge to design instruction to help them assimilate and accommodate new ideas and build a good knowledge structure. Vygotsky developed a theory which introduces the notion of the "zone of proximal development" (ZPD). The ZPD refers to the zone defined by the difference between what a student can do on his/her own and what a student can do with the help of an instructor who is familiar with his/her prior knowledge and skills [80]. Scaffolding is at the heart of this model and can be used to stretch students' learning beyond their current knowledge by carefully crafted instruction. These cognitive theories all point to the fact that one must determine the initial knowledge states of the students (i.e., students' prior knowledge and skills in Factor 2) in order to design effective instruction commensurate with students' current knowledge and skills (i.e., learning tool characteristics in Factor 1).

For example, instructional design that conforms to the field tested cognitive apprenticeship model [81] can help students learn effectively (see Factor 1). The cognitive apprenticeship model involves three major components: modeling, coaching and scaffolding, and weaning. This approach has also been found effective in helping students learn effective problem-solving heuristics and develop their cognitive skills. In this approach, "modeling" means that the instructor demonstrates and exemplifies the skills that students should learn. "Coaching and scaffolding" refer to providing students suitable practice, guidance, and feedback so that they learn the skills necessary for good performance. "Weaning" means gradually fading the support and feedback with a focus on helping students develop self-reliance.Much research in physics education has focused on these cognitive factors in developing effective instruction and assessment. For example, in physics, tutorials, peer instruction (clicker questions with peer discussion), collaborative group problem solving with context-rich physics problems, POGIL (process-oriented guided-inquiry learning) activities, etc. have been found effective in helping students learn [64-76]. Furthermore, several validated conceptual standardized surveys have been developed [83] to assess students' conceptual understanding of various physics topics.

However, Factors 1 and 2 are also informed by several "non-cognitive" aspects such as students' motivational characteristics, affective factors, and self-regulation. Several studies have focused on students' motivation (e.g., factors such as fascination and value associated with physics, self-efficacy, grit, intelligence mindset, and physics epistemology) [25-63]. For example, researchers have focused on students' interest in and value of science and their relationship to performance and STEM degree achievement [25-35]. Research suggests that interest in math and science is associated with the number of math and science courses taken in high school and career aspirations [28]. In addition, self-efficacy is another factor that can impact students' motivation and learning [36-46]. Self-efficacy is the belief in one's capability to be successful in a particular task, subject area, or course [36]. Students with high self-efficacy often exhibited effective learning behaviors during conceptual learning, such as self-monitoring and persistence, were less likely to reject correct hypotheses prematurely, and were better at solving conceptual problems than students with low self-efficacy but of equal ability [41]. Self-efficacy has been shown to positively predict performance in science courses and scores on the Force Concept Inventory [16]. Moreover, in Peer Instruction [69] environments in a physics course, students with low self-efficacy are more likely to switch to a wrong response after discussions with a peer than those with high self-efficacy [44]. Self-efficacy is also positively correlated with expected grades in physics [40], enrollment in STEM courses, and career choices in STEM [27].

Another factor that is associated with motivation is the construct of grit, introduced by Duckworth et al [47-52]. They defined grit as a trait-level perseverance and passion for long-term goals—a capacity to sustain both effort and interest in projects that take months or even longer to complete and adherence to goals even in the absence of positive feedback. Grit has been shown to predict achievement in challenging domains over and beyond measures of talent [47]. However, to our knowledge, there are no studies of students' grit within an introductory physics course sequence.

Dweck and colleagues defined another motivational construct that can affect motivation and learning, i.e, theories of intelligence (or intelligence mindset) [53-57]. Theories of intelligence involve views about the nature of intelligence—an "entity theory" in which intelligence is viewed as a fixed trait that one is born with or an "incremental theory" in which intelligence is viewed as malleable and can be shaped by the environment. Research has shown that children's implicit theories about the nature of intelligence impacts the goals they pursue, their response to difficulty, and how well they do in school [53]. Students who view intelligence as a fixed trait are concerned with demonstrating intelligence and prefer tasks in which they feel capable and smart. When faced with difficult tasks, students who view intelligence as a fixed trait tend to become debilitated and disengage. On the other hand, students who view intelligence as malleable are concerned with learning new concepts and improving competence. When faced with difficult tasks, these students appear to experience less anxiety, put forth more effort, and increase their engagement [53].

Students' epistemologies, or their beliefs about knowledge and learning, can also affect their motivation and learning of science [59-63]. Students' epistemologies in physics includes beliefs about the nature of physics (the belief that that physics knowledge is a set of disconnected facts and formulas or an organized system of interconnected concepts and principles), agency in learning physics (the belief that physics knowledge comes from authority as opposed to constructing their own knowledge), and study strategies (the belief that one should memorize formulas as opposed to understand them). Hammer's study on college students showed that students' ideas about knowledge and learning in physics affected how they solved homework problems during think-aloud interviews [58]. Physics education researchers have developed surveys that assess students' epistemologies in physics [59,84]. Research has found that students' responses to epistemological surveys stay approximately the same or become less "expert like" after taking traditionally taught introductory physics courses, e.g., students view physics knowledge as a collection of formulas as opposed to a set of interconnected concepts and view knowledge as coming from authority as opposed to developing their own understanding [59].

2.2.2 Factors 3 and 4: External characteristics of learning tools and students

We note that instructional tools and student characteristics (Factors 1 and 2) do not exist in a "vacuum." One must also take into account the interplay between instruction, students, and the environment. Factors 3 and 4 focus on the environment, i.e., how instructional design is implemented in a particular course and students' social environments, respectively. One example of Factor 3 involves "framing" of instruction to achieve student "buy-in," which is helpful in motivating students to engage with learning. Several studies have shown that providing rationales that, e.g., identify why it is worth the effort to engage and communicates why it is useful to students can help them engage constructively with a lesson [85,86]. Motivational researchers also posit that providing stimulating and interesting tasks that are personally meaningful, interesting, relevant, and/or useful to students can increase their interest and value in a subject, increasing their motivation and engagement in learning [87]. In physics, researchers have developed "context-rich" physics problems, i.e., problems that involve "real-world" applications of physics principles and are complex and ill-defined [88]. These types of problems can often increase students' interest and value associated with physics. Furthermore, instruction that fosters a community of learners can also encourage productive engagement in learning. Within this community of learners, students are encouraged to construct their own knowledge and are held accountable to others, which can, in part, encourage them to engage deeply with the content [89,90]. Factor 4 focuses on the student-environment interaction. Indeed, having supportive parents, teachers, and mentors can be beneficial in fostering students' motivation and engagement in the classroom [91]. Students' time management skills have also been shown to correlate with performance in college [92].

In sum, the four factors of the SELF Framework can be thought of holistically when designing instruction. We note that each of the factors can interact with other factors. For example, the way in which instructional tools are implemented (Factor 3) is informed by the characteristics of the learning tool (Factor 1), the characteristics of the students (Factor 2), and the way the students interact with their social environments (Factor 4). Furthermore, the student characteristics in Factor 2 can inform the learning tool characteristics (Factor 2), the implementation of learning tools (Factor 3), as well as how the student interacts with the environment (Factor 4). Here, we investigate baseline measures of, e.g., student characteristics in Factor 2 such that instructors, instructional designers, and researchers can develop effective learning tools and implement them appropriately to help all students

2.2.3 Gender differences in factor 2 of the SELF framework

In this study, we focus on baseline measures of student characteristics in Factor 2, and in particular, we focus on gender differences in motivational characteristics throughout an introductory physics course sequence. Prior research has already revealed several gender differences in regards to Factor 2 of the SELF framework. For example, on many physics conceptual tests, female students underperform male students [2-8,83]. Docktor and Heller found a consistent gap in the scores of men and women on the Force Concept Inventory in introductory undergraduate physics courses, with men scoring, on average, approximately 13% higher than women on the posttest [7]. A prior study has also showed that women had lower gains on the Force Concept Inventory than men, despite the fact that there were no differences in course grades [6]. Similarly, women score lower than men on other conceptual surveys, such as the Test of understanding Graphics in Kinematics and Determining and Interpreting Resistive Electric Circuit Concepts Tests [83].

Researchers have also discovered gender differences with regard to motivational characteristics (see Factor 2 of the SELF framework). For example, one of the reasons women choose not to pursue STEM degrees is because their career interests focus on altruistic efforts and helping others, and they did not view STEM fields as those in which they could help others [30]. However, one study has shown that when a STEM career was presented to females as beneficial to society, their interest in the field increased [30]. Furthermore, females demonstrate lower self-efficacy relative to males beginning in middle school and throughout high school and college [27].

Females' intelligence mindset has also been investigated. It was found that girls who attributed intelligence to effort and learning had math grades comparable to those of their male classmates and superior to girls who viewed intelligence as a fixed trait [54]. Research has also shown that individuals targeted by ability stereotypes (e.g., underrepresented students in STEM courses) tend to show similar characteristics of individuals who believe that intelligence is fixed—they tend to choose easier, success-assuring tasks when their abilities are subject to scrutiny or if their ethnicity or gender is made more salient [55], experience anxiety when the tasks are evaluative and challenging [56], and devalue ability domains in which they have performed poorly [57]. Additional research has also shown that girls who viewed math ability as a trait (had a fixed view of intelligence) and also experienced stereotype threat (that girls are not good at math) had decreased motivation and interest in pursuing math careers [53]. However, to our knowledge, there have been relatively few studies of students' theories of intelligence in physics courses [76]. On the Colorado Learning Attitudes about Science Survey (CLASS) [84], which is a survey of students' epistemological views about physics, women are generally less "expert-like" on statements involving real-world connections, personal interest, problem solving confidence, and problem-solving sophistication than males.

2.2.4 Longitudinal studies of factors related to students' motivation

Here we examine male and female students' motivational characteristics longitudinally, i.e., over the course an introductory physics sequence. We note that there have been relatively few longitudinal studies of students' motivation in physics [16,93-98]. In the study by Bates et al. [93], researchers used the Colorado Learning Attitudes about Science Survey (CLASS) to investigate how undergraduate physics students' attitudes changed over a period of three years. They found that “expert-like thinking” was unchanged over the duration of the undergraduate physics program. Gire et al. [96] also performed a similar study in which the CLASS was administered to students in undergraduate and graduate physics courses over a period of two years. They found that physics majors have relatively “expert-like” views when they begin their undergraduate study of physics and maintain these views throughout the undergraduate physics program. Cavallo et al. [16] investigated gender differences in learning constructs and shifts in a yearlong, inquiry-based physics course for life science majors. They found that males had significantly higher self-efficacy, performance goals, and physics understanding compared to females, and these differences persisted throughout the course. In addition, they found that self-efficacy significantly predicted physics understanding and course achievement for both men and women. Lindstrom and Sharma [97] investigated students’ self-efficacy in a first-year university physics course. They found that women consistently reported lower self-efficacy than males. Both gender and prior formal physics instruction impacted students’ self-efficacy.

However, none of these studies focused on students’ fascination with physics, value associated with physics, grit, or intelligence mindset—all of which may play a role in students’ motivation and learning and need to be taken into account holistically as posited by the SELF framework. Furthermore, to our knowledge, there have been no longitudinal studies in physics that focus on factors related to females’ motivation. Thus, we investigated student characteristics (factor 2), i.e., fascination with physics, value associated with physics, self-efficacy, grit, intelligence mindset, and physics epistemology. We report baseline data on these measures for both men and women in physics and how their motivational characteristics change over the course of an introductory physics sequence. The baseline data can be a stepping stone for developing better learning tools (factor 1) and determining more effective ways to implement learning tools (factor 3) in order to increase student engagement and success in physics courses. Providing a baseline can also help determine for which motivational factors there are gender differences and when those differences are present (e.g., at the beginning of a physics course sequence or later). This information can inform future instruction and interventions. For example, if there are differences between male and female students' interest at the beginning of a physics course sequence, one intervention may involve boosting female students' interest in physics early on in the course.

2.3 Methodology

To investigate students' fascination with physics, perceived value of physics, physics self-efficacy, grit, intelligence mindset, and physics epistemology, we created a 29-item questionnaire in which we adapted items from previously published surveys for each construct [48,53,59,84,99-101]. We administered the survey in introductory physics courses three times over a period of one academic year. Below, we describe the validation and administration of the survey as well as how the student demographics (such as gender and race) were linked to students' responses on the surveys.

2.3.1 Development of the survey

The survey on motivational factors was developed by four of the researchers iteratively. The researchers first decided which constructs to assess based upon a review of the literature on student motivation (see the background section). Then, initial versions of the survey were iterated which included a set of questions from other well-validated surveys in the literature on motivation [48,53,59,84,99-101]. Surveys that served as sources were selected based on prior empirical validation and appropriateness for university undergraduate students. Individual items were selected to cover the conceptual space of each construct without redundancy; surveys sometimes improve reliability by including nearly identical questions. Minimalistic scales for each construct were necessary because many constructs were included in the overall survey, and the researchers decided that the students must be able to complete the survey in fifteen minutes or less to encourage high completion rates and longitudinal use. The final version of the survey includes 29 Likert-scale items. All items were on a Likert scale of 1-4 except for grit, which was on a scale of 1-5. Table 2-1 shows a representative list of items in survey, the corresponding motivational or epistemological factor, and the original survey the item comes from. The final version of the survey is included in the appendix.

Table 2-1 For each factor, number of items, Cronbach alpha from each of the first two administrations (α_1 and α_2), example survey items, and source of items.

Factor	Example Survey Item	Scale	Original survey
Fascination with Physics 3 items α_1=0.63, α_2=0.67	□ I wonder about how nature works	Never Once a month Once a week Every day	see ref. [99]
	□ In general, I find physics	Very boring boring Interesting Very interesting	
Valuing Physics 5 items α_1=0.69, α_2=0.70	□ Knowing physics is important for being a good citizen … □ Physics makes the world a better place to live …	No! no yes Yes!	see ref. [99]
Physics Self-Efficacy 6 items α_1=0.75, α_2=0.74	□ I am often able to help my classmates with physics in the laboratory or in recitation … □ If I study, I will do well on a physics test …	No! no yes Yes!	see refs. [84,99-101]
Intelligence Mindset 4 items α_1=0.65, α_2=0.64	□ You have a certain amount of intelligence, and you can't really do much to change it …. □ Anyone can become good at solving physics problems through hard work …	Strongly disagree Disagree Agree strongly agree	see refs. [53,59]
Grit 4/3 items α_1=0.58, α_2=0.68	□ I often set a goal but later choose to pursue a different one …. □ I finish whatever I begin …. □ I have difficulty maintaining my focus on projects that take more than a few months to complete.	Not like me at all Not much like me Somewhat like me Mostly like me Very much like me	see ref. [48]
Physics Epistemology 8 items α_1=0.64, α_2=0.70	□ I do not expect to understand physics equations in an intuitive sense; they must just be taken as givens. □ When doing an experiment, I try to understand how the experimental setup works ….	Strongly disagree Disagree Agree Strongly agree	see ref. [59,84]

For each survey item, students were given a score of 1-4 (or 5, for the questions related to grit). For the items related to fascination, value, grit, and self-efficacy, a high score means that a student is highly fascinated by physics, values physics highly, and has a high level of grit and self-efficacy. Furthermore, for the factor of intelligence mindset, a high score means that a student has a malleable view of intelligence, whereas a low score means that a student views intelligence as a fixed ability. For the factor of physics epistemology, a high score means that a student has a more "expert-like" view of knowledge and learning in physics, whereas a low score means that a student has a "novice-like" view. We note that some of the items on the survey were reverse coded since the statement was posed in a negative way (e.g., a student who strongly agrees with the statement

“I do not expect to understand physics equations in an intuitive sense; they must just be taken as givens” would be given a score of 1). For each factor (e.g., fascination, value, grit, self-efficacy, intelligence mindset, and physics epistemology), each student was given an average score. For example, a student who answered “Yes!” to two of the fascination questions and “no” to the other fascination question would have an average fascination score of (4+4+2)/(3 total questions) = 3.33.

2.3.2 Administration of the survey

The survey was administered to students in algebra-based and calculus-based physics courses at the beginning of the fall semester of 2015, the beginning of the spring semester 2016, and the end of the spring semester 2016. The physics courses were varied in the type of pedagogy used, but the majority of the courses were lecture-based physics courses. Physics 1 courses included topics such as kinematics, forces, energy and work, rotational motion, gravitation, and oscillations and waves. Physics 2 courses included topics such as electricity and magnetism, electromagnetic waves, images, interference, and diffraction. The algebra-based physics courses are typically taken by students majoring in many different areas but planning on pursuing medicine or other health professions, for which the two-semester physics sequence is required. The calculus-based physics courses are typically taken by engineering and physical science majors as a required course for those majors. The physics courses included three lecture hours and one recitation hour per week. The recitations were mandatory and included a weekly, low stakes quiz. Most of the physics courses were in a traditional, lecture-based format, but there was one calculus-based and one algebra-based course that was in the “flipped course” format. In the “flipped” courses, students watched lecture videos before attending the lectures. In the lectures, they worked on collaborative group problem solving and clicker questions. Thus, some of the courses involved in this study involved characteristics from Factors 1, 2, and 3 in the SELF framework, i.e., providing scaffolding support, involving formative assessment, taking into account students' prior knowledge and skills, and use of collaborative learning and grade incentives to encourage students to engage with learning tools. However, we note that most of the courses did not explicitly take into account students' motivational characteristics in factor 2 and did not necessarily include features from factor 3 such as framing the importance of learning from tools and to get student buy in and embedding motivational features within tools that are conducive to effective learning.

The survey was typically administered in the first and last recitations of the course, although some instructors chose to give the survey in the lecture portion of the course. Furthermore, in the first round of administration (Fall semester 2015), some instructors chose to give the survey in a written format, whereas others chose to give the survey in an online format outside of class. In the first round of administration (Fall semester 2015), we found that the participation rates were significantly lower for students who were given the online format of the survey. Thus, in subsequent administrations of the survey (Spring semester 2016), we asked instructors to administer the survey in a written format. The survey was completed by most students in about 10-15 minutes. In general, there were no grade incentives given to students who took the survey. The instructor or teaching assistant responsible for giving the survey was given the following script to announce before administering the survey to the students: "We are surveying you on your beliefs about physics in order to improve the class. Your responses will not be evaluated for grades except to make sure the responses were done seriously, rather than randomly." This script encouraged students to take the survey seriously.

2.3.3 Validation of the survey

At the beginning of fall 2015 and the beginning of spring 2016, we analyzed the internal consistency of the subscales, i.e., fascination, value, self-efficacy, intelligence mindset, grit, and physics epistemology. After the initial reliability of the subscales in the fall 2015, we removed one of the statements in the grit subscale ("Setbacks don't discourage me") since the wording may have been confusing for students and it did not correlate well with other statements related to grit. All alphas are above 0.60 across both time points (see Table 2-1) which is considered fairly good, especially since some of the scales have only three items. The scales with five or more items all had a Cronbach's alpha of 0.70 or higher. No substantial increases in alpha for any of the scales could have been achieved by eliminating items.

To establish the separability of six different subscales along with validity of items as clear indicators of the scale to which they were assigned, we performed an exploratory factor analysis on the items in the survey based upon the data at the beginning of the spring 2016. A principal components analysis with Varimax rotation method was used, and the initial eigenvalues indicated that the first six components (all with eigenvalues greater than 1) explained a total of 49% of the variance (the 7th component explained only an additional 3.6% of the variance). Item loadings on each component are given in the appendix. Component 1 is related to self-efficacy, component 2 is related to value associated with physics, component 3 is related to physics epistemology, component 4 is related to fascination with physics, component 5 is related to intelligence mindset, and component 6 is related to grit. Thus, the data supported the existence of six separable scales, and items loaded on the scales as intended.

We hypothesized that males and females may respond differently to the intelligence mindset items on the survey. For example, two of the items in the survey related to "general" views about intelligence mindset: "You have a certain amount of intelligence, and you can't really do much to change it" and "No matter how much intelligence you have, you can always change it quite a bit." However, the other two items related to intelligence mindset in the survey were embedded in a physics context: "Anyone can become good at solving physics problems through hard work" and "Only very few specially qualified people are capable of really understanding physics." We postulated that males and females may answer questions related to "physics" intelligence mindset differently due to the stereotype that men perform better in logical, math-intensive fields and women perform better in communication and writing-intensive fields [53]. Thus, in the results section, we report students' average "general" intelligence mindset and "physics" intelligence mindset separately.

2.3.4 Connecting student demographics to the survey

To connect student demographics such as gender and race to students' responses on the survey, we collected data on students who were enrolled in physics courses from the University's data warehouse. Both sets of data were linked by an identification number for each student that was based on a hash-function of their university email; that is, the researchers only had access to the demographics data in this de-identified form. Gender and ethnicity were included in the demographic data. We removed students whose gender was missing.

For the results related to gender, we report data for students who were enrolled in Physics 1 in the fall 2015 and Physics 2 in the spring 2016 (i.e., students who were enrolled in the "on-cycle" sequence of the courses) to simplify interpretations of changes across the courses. Results are discussed separately for students who were enrolled in calculus-based and algebra-based introductory physics courses, given the differential representation across those divides, the likely differences in haven taken high school physics, and the centrality of physics to their own majors.

Table 2-2 shows the percentages of males and females in introductory physics courses who completed the survey at the three points in time. The number of students is different at different points of time due to the fact that some students did not take physics 2 courses and some students may not have been present in the lecture or recitation section in which the survey was administered (and therefore did not participate in the survey).

Table 2-2 The total number of students and the percentages of students who completed the survey at different points in time broken down by males (M) and females (F).

	Algebra-Based Physics 1	Algebra-Based Physics 2	Calculus-Based Physics 1	Calculus-Based Physics 2
Fall 2015 (N=1,095)	n=421; 64% F, 36% M	n=85; 60% F, 40% M	n=467; 31% F, 69% M	n=122; 28% F, 72% M
Beginning of Spring 2016 (N=1,149)	n=306; 59% F, 41% M	n=380, 63% F, 37% M	n=119, 38% F, 62% M	n=344; 34% F, 66% M
End of Spring 2016 (N=871)	n=243; 61% F, 39% M	n=295; 63% F, 37% M	n=59; 41% F, 59% M	n=274; 35% F, 65% M

2.4 Results

2.4.1 Gender differences at the beginning of the physics 1

We found that there were significant gender differences in male and female students' motivational factors at the beginning of Physics 1 in both calculus-based and algebra-based courses. We performed a one-way multivariate analysis of variance (MANOVA) with the dependent variables being students' average reported self-efficacy, fascination, value, intelligence mindset, physics epistemology, and grit at the beginning of Physics 1 and the independent variable being gender. For the calculus-based course, the results of the MANOVA revealed that there was a statistically significant difference in male and female students' overall motivation, $F(7,459)=10.344$, $p<0.001$; Wilks' Lambda=0.864. Follow-up univariate tests revealed that males reported statistically significantly higher levels of self-efficacy, fascination, value associated with physics, and physics intelligence mindset. See Table 2-3 and Figure 2-2 for p-values and Cohen's d effect sizes (Cohen's d is $d = (\mu_1 - \mu_2)/\sigma_{pooled}$, where μ_1 is the average score of males and μ_2 is the average score of females and σ_{pooled} is the standard deviation of the average score of all students).

For the algebra-based course, the results of the MANOVA revealed that there was a statistically significant difference in male and female students' overall motivation, $F(7,412)=10.213$, $p<0.001$; Wilks' Lambda=0.852. Follow-up univariate tests revealed that males reported statistically significantly higher levels of self-efficacy and fascination. On the other hand, females reported statistically significantly higher levels of grit. See Table 2-3 and Figure 2-3 for exact p-values and effect sizes.

Table 2-3 Effect sizes showing the comparison of factors related to males' and females' motivation throughout introductory physics course sequence. Bolded effect size values indicate statistically significant differences at the level of *p* < 0.05.

	Male > Female effect size		
Calculus-Based	Beginning of Physics 1	Beginning of Physics 2	End of Physics 2
Self-Efficacy	**0.75**	**0.80**	**0.95**
Fascination	**0.37**	**0.63**	**0.89**
Value	**0.32**	**0.35**	**0.70**
Physics Intelligence Mindset	**0.22**	**0.43**	**0.28**
Physics Epistemology	0.05	0.08	0.18
General Intelligence Mindset	-0.09	-0.01	-0.12
Grit	-0.17	-0.17	-0.02
Algebra-Based			
Self-Efficacy	**0.63**	**0.61**	**0.62**
Fascination	**0.60**	**0.83**	**0.56**
Value	0.20	**0.39**	**0.35**
Physics Intelligence Mindset	0.20	**0.24**	**0.48**
Physics Epistemology	0.02	0.13	0.11
General Intelligence Mindset	-0.04	-0.02	0.03
Grit	**-0.26**	-0.20	**-0.42**

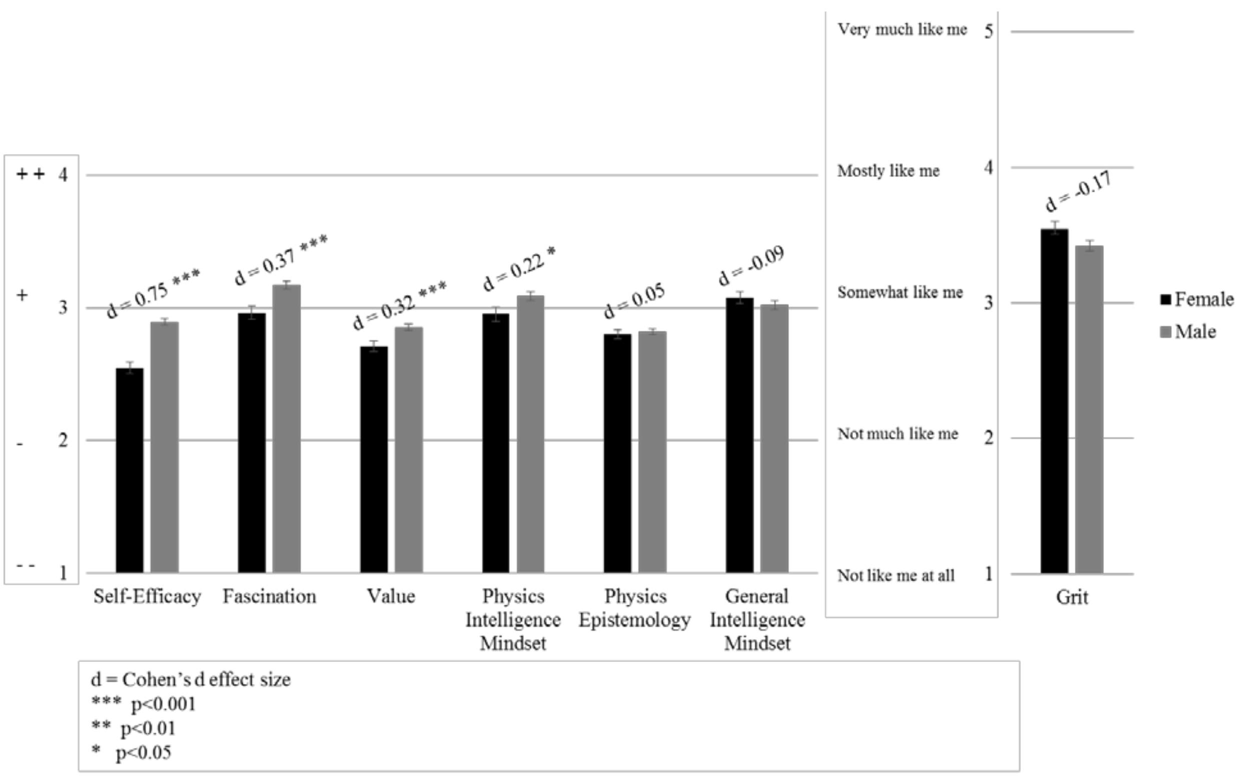

Figure 2-2 Females' (N=144) and males' (N=323) average reported motivational characteristics at the beginning of calculus-based Physics 1, with standard error bars. The "+" signs indicate positive valence responses and the " – " signs indicate negative valence responses. We separate the construct of grit since it was a 5-point Likert scale.

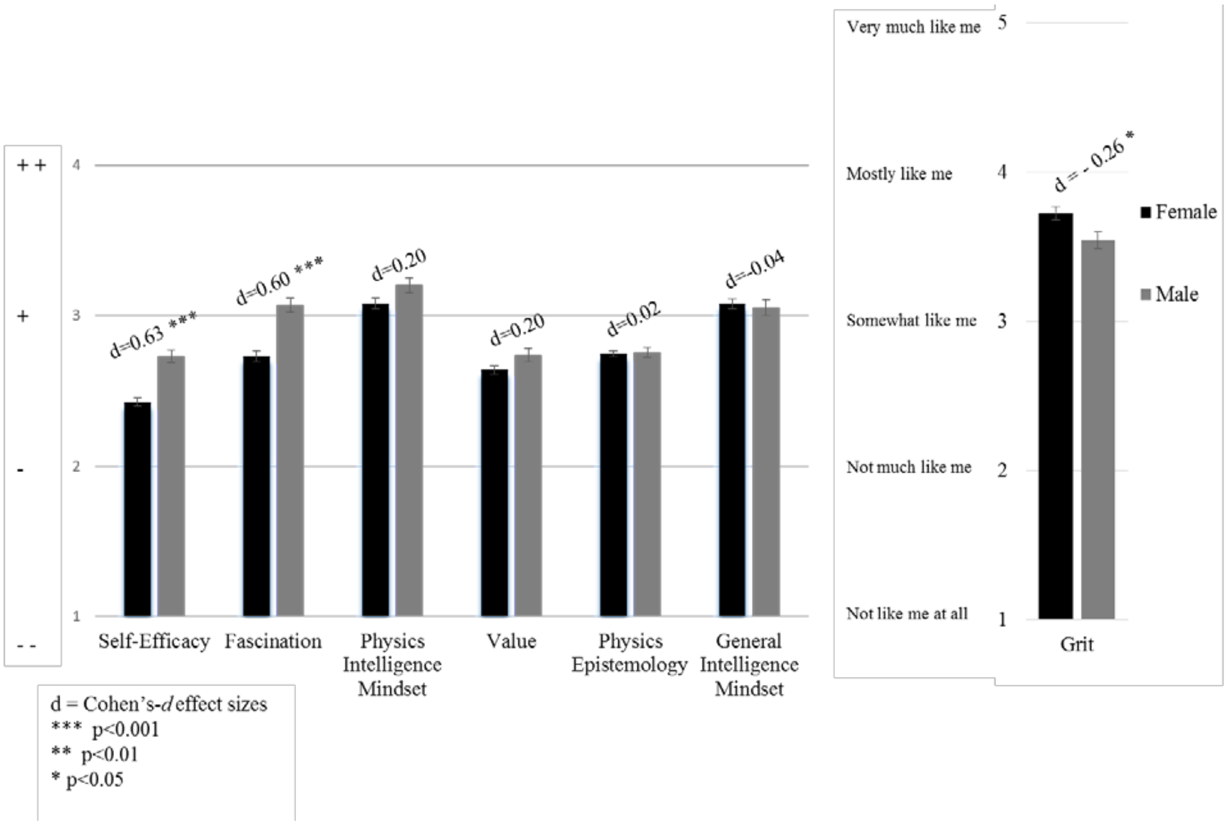

Figure 2-3 Females' (N=270) and males' (N=151) average reported motivational characteristics at the beginning of algebra-based Physics 1, with standard error bars. The "+" signs indicate positive responses and the "-" signs indicate negative responses. We separate the construct of grit since it was a 5-point Likert scale.

2.4.2 Changes in students' motivational characteristics over the course of an introductory physics course sequence

We calculated change scores for the change in students' motivational characteristics between the beginning of Physics 1 and the beginning of Physics 2. For example, in order to calculate a student's self-efficacy, change score in Physics 1, we subtracted the student's average self-efficacy score at the beginning of Physics 1 from the student's average self-efficacy score at the beginning of Physics 2. This change score allowed us to determine the extent to which students' motivational characteristics changed after taking an introductory Physics 1 course focusing on mechanics and Newton's laws. A one-way MANOVA was performed with the dependent variables being the change scores in students' average self-efficacy, fascination, value, physics intelligence mindset, general intelligence mindset, physics epistemology, and grit and the independent variable being gender. We found that there was not a significant difference in the change in male and female students' overall motivation from the beginning of Physics 1 to the beginning of Physics 2 in either calculus-based or algebra-based courses. In other words, on average, the gaps between male and female students' motivational characteristics remained the same. We did find that there was a significant difference in the change scores of males and females in regards to fascination in calculus-based Physics 1 courses (result of ANOVA gives $F(1, 280)=5.688$, $p=0.018$). That is, females' average reported fascination with physics decreased significantly more than males' after taking a Physics 1 course and the gap between male and female students' reported fascination increased. See Figure 2-4 for the change in male and female students' fascination after taking a calculus-based Physics 1 course.

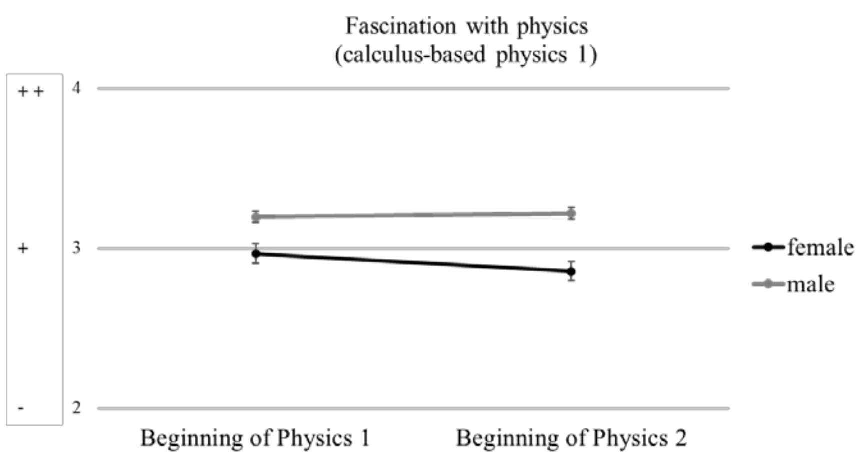

Figure 2-4 Male and female students' fascination with physics before and after taking calculus-based Physics 1, with standard error bars. The "+" signs indicate positive responses and the "-" signs indicate negative responses.

We also calculated change scores for the change in students' motivational characteristics between the beginning of Physics 2 and the end of Physics 2. For example, in order to calculate a student's self-efficacy change score in Physics 2, we subtracted the student's average self-efficacy score at the beginning of Physics 2 from the student's average self-efficacy score at the end of Physics 2. This change score allowed us to determine the extent to which students' motivational characteristics changed after taking an introductory Physics 2 course focusing on electricity and magnetism. A one-way MANOVA was performed with the dependent variables being the change scores in students' average self-efficacy, fascination, value, physics intelligence mindset, general intelligence mindset, physics epistemology, and grit and the independent variable being gender. We found that there was not a significant difference in the change in male and female students' overall motivation from the beginning of Physics 2 to the end of Physics 2 in either calculus-based or algebra-based courses. In other words, the gaps between male and female students' motivational characteristics remained approximately the same after taking a Physics 2 course. However, we did find that there was a significant difference in the change scores of males and females in regards to fascination and value in calculus-based courses. That is, females' average fascination decreased significantly more than males (ANOVA result gives $F(1,246)=6.412$, $p=0.012$) after taking a calculus-based Physics 2 course. Furthermore, females' value associated with physics decreased significantly more than males' (ANOVA result gives $F(1,246)=5.637$, $p=0.018$) in calculus-based Physics 2. In other words, the gap between female and male students' reported fascination and value associated with physics increased after taking a calculus-based Physics 2 course. See Figure 2-5 for the change in male and female students' fascination and value after taking a calculus-based physics 2 course.

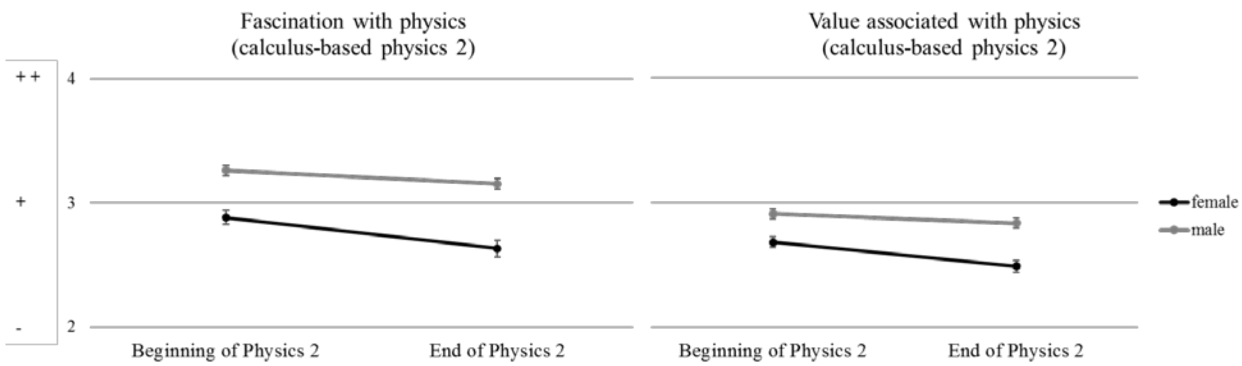

Figure 2-5 Students' fascination and value associated with physics at the beginning and end of calculus-based physics 2, with standard error bars. The "+" signs indicate positive responses and the "-" signs indicate negative responses.

In addition, algebra-based Physics 2, we found that there was a significant difference in the change scores of male and female students in regards to physics intelligence mindset. That is, females' reported physics intelligence mindset became more "fixed" relative to males and the gap between female and male students' physics intelligence mindset increased (ANOVA result gives $F(1, 278)=5.359$, $p=0.021$). See Figure 2-6 for the change in male and female students' physics intelligence mindset after taking an algebra-based physics 2 course. See the appendix Figure 2-7 – Figure 2-13 for male and female students' average self-efficacy, fascination, value associated with physics, physics intelligence mindset, physics epistemology, general intelligence mindset, and grit over the course of an introductory physics course sequence (we note that Appendix Figure 2-7 – Figure 2-13 include all students who responded to the survey, not just the common cohort of students taking both Physics 1 and Physics 2).

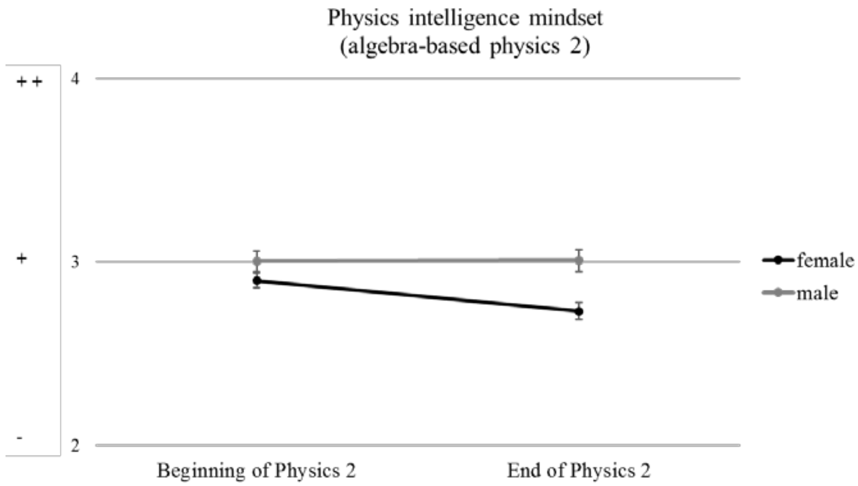

Figure 2-6 Students' physics intelligence mindset at the beginning and end of algebra-based physics 2, with standard error bars. The "+" signs indicate positive responses (i.e., a more malleable view of intelligence) and the "-" signs indicate negative responses (i.e., a more fixed view of intelligence).

2.5 Summary and discussion

We found gender differences with regards to students' views of several motivational factors throughout a two-semester introductory physics courses. At the beginning of a Physics 1 course, females' average self-efficacy, fascination, value associated with physics, and physics intelligence mindset were significantly lower than males'. We also found that females' fascination and value associated with physics decreased significantly more than males' after taking a calculus-based physics course sequence. Furthermore, after taking an algebra-based physics 2 course focusing on electricity and magnetism, females' physics intelligence mindset became more "fixed" as compared to males' physics intelligence mindset. This result suggests that women may view intelligence in physics differently than intelligence in general, e.g., that one must be "brilliant" to do physics. We found that grit was the only construct on which females reported average scores higher than males.

We discussed a holistic framework, SELF, that can be used as a guide to increase students' engagement with introductory physics courses. The SELF framework posits that instructional design, learning tools and their implementation along with student characteristics and environments play a critical role in helping students learn effectively. Since student characteristics can play a central role in whether students are engaged in learning effectively, the baseline data we discussed for students' characteristics can be a useful stepping stone for developing and implementing pedagogies and learning tools to help all students succeed in introductory physics courses.

2.6 Implications

Students' characteristics discussed here can play a central role in students' engagement with a course and impact their learning. We found important gender differences on motivational factors such as value, fascination, self-efficacy, and intelligence mindset. Researchers, curriculum developers, and instructors can take these differences into account when developing and implementing instruction and learning tools to help all students engage effectively with physics courses.

Our findings suggest that that females who were interested in physics initially (had high value and fascination associated with physics at the beginning of a physics course) lose interest over the introductory physics course sequence. Instructors, researchers, and curriculum developers may be able to boost females' engagement with physics courses by framing physics as a science that can benefit society throughout a physics course sequence. Indeed, a recent study showed that when a STEM career was presented to females as more altruistic, communal, and beneficial to people, females' interest in STEM fields increased [30,102]. Framing instruction and learning tools appropriately to get buy-in from students (factor 3 in the SELF framework), e.g., framing physics as a science that can benefit society, is one way to motivate all students in introductory physics courses to be engaged with curriculum and learning tools effectively.

In addition, we found that women tend to enter physics courses with lower self-efficacy than men, and their self-efficacy remains low after taking introductory physics courses. Low self-efficacy may be associated with women feeling like they do not belong and that they are not capable of doing physics. Certain types of instruction and learning tools in Factor 1 in the SELF framework (when implemented appropriately) can improve students' self-efficacy. For example, research has shown that instructional strategies such as collaborative learning, conceptual problems, and inquiry-based labs contributed to students' improved self-efficacy and improved classroom climate [40]. Modeling Instruction has also been shown to improve students' self-efficacy, especially for females [11]. Furthermore, the implementation of instruction and learning tools (Factor 3 of the SELF framework) can impact students' self-efficacy. For example, research suggests that when implementing collaborative group work, collaborative groups should include at least two underrepresented students in a group because an underrepresented student's ideas and contributions are sometimes ignored if he/she is the only underrepresented student in the group [103]. The timing of instructional interventions can also be taken into account—e.g., since women tend to have lower self-efficacy at the beginning of a physics course sequence, this issue should be addressed early on in the course.

Our research also suggests that women tended to view intelligence in physics as a fixed ability, compared to men. This type of mindset can affect learning. For example, Dweck found that students who believe that intelligence is fixed are more likely to attribute academic setbacks to a lack of ability than students who believe that intelligence is malleable and improvable with hard work and effort [53]. These views shape how students respond to setbacks, either by withdrawing effort or redoubling effort, seeking help, using a better strategy, etc. It may be beneficial to provide explicit support to students to improve their views about intelligence during the implementation of learning tools (Factor 3 of the SELF framework), such as discussing the importance of using effective learning strategies, attributing failure to inappropriate learning strategies as opposed to intellectual ability, and viewing mistakes as an opportunity for learning and growth throughout an introductory physics course sequence. Indeed, interventions have been developed that help students view intelligence as malleable. For example, Blackwell and colleagues designed an intervention for middle school students in which they attended an eight-session workshop to learn about study skills and scientific research showing that the brain grows connections and "gets smarter" when a person works on challenging tasks. Students who learned that intelligence is malleable earned better math grades over the course of the year than students in the control group [54]. Other interventions involved providing effort v. intelligence praise after success [104] and orchestrating email exchanges with a mentor during a yearlong program [105].

Instructors, researchers, and curriculum developers can create and implement effective instruction and learning tools in physics, in part, by considering students' motivational characteristics. Focusing on students' motivational characteristics, especially those of underrepresented students, may prove fruitful in helping more students succeed in STEM courses and increase the diversity in STEM fields.

2.7 Chapter references

1. See, e.g., https://www.nsf.gov/statistics/2017/nsf17310/data.cfm

2. S. Eddy and S. Brownell, Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines, PRPER **12**, 020106 (2016).

3. A. Traxler, X. Cid, J. Blue, and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future, PRPER **12**, 020114 (2016).

4. A. Noack, T. Antimirova and M. Milner-Bolotin, Student diversity and the persistence of gender effects on conceptual physics learning, Can. J. Phys. **87**, 1269 (2009).

5. S. Bates, R. Donnelly, C. MacPhee, D. Sands, M. Birch and N. Walet, Gender differences in conceptual understanding of Newtonian mechanics: a UK cross-institution comparison Eur. J. Phys. **34,** 421 (2013).

6. V. Coletta, J. Phillips, and J. Steinert, FCI normalized gain, scientific reasoning ability, thinking in physics, and gender effects, in AIP. Conf. Proc. (AIP Publishing, 2012), pp. 23-26.

7. J. Docktor and K. Heller, Gender differences in both Force Concept Inventory and introductory physics performance. In C. Henderson, M. Sabella, & Leon Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference* (pp. 15-18). Melville, NY: American Institute of Physics (2008).

8. A. Madsen, S. McKagen and E. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? Phys. Rev. ST. Phys. Educ. Res. **9**, 020121 (2013).

9. Z. Hazari, R. Tai and P. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, Science Education **91**, 847 (2007).

10. L. McCullough, Gender, context, and physics assessment, J. Int. Women's Stud. **5**, 20 (2004).

11. E. Brewe, V. Sawtelle, L. Kramer, G. O'Brien, I. Rodriguez and P. Pamela, Toward equity through participation in Modeling Instruction in introductory university physics, Phys. Rev. ST. PER **6**, 010106 (2010).

12. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

13. R. Lock and Z. Hazari, Discussing underrepresentation as a means to facilitating female students' physics identity development, PRPER **12**, 020101 (2016).

14. J. Day, J. Stang, N. Holmes, D. Kumar and D. Bonn, Gender gaps and gendered action in a first-year physics laboratory, PRPER **12**, 020104 (2016).

15. I. Rodriguez, G. Potvin, and L. Kramer, How gender and reformed introductory physics impacts student success in advanced physics courses and continuation in the physics major, PRPER **12**, 020118 (2016); A. Miyake, L. Kost-Smith, N. Finkelstein, S. Pollock, G. Cohen and T. Ito, Reducing the gender achievement gap in college science: A classroom study of values affirmation, Science **330**, 1234 (2010).

16. 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).

17. See the web-based interventions at http://www.mindsetworks.com/brainology/

18. B. Adegoke, Impact of interactive engagement on reducing the gender gap in quantum physics learning outcomes among senior secondary school students, Phys. Educ. **47,** 462 (2012).

19. A. Maries and C. Singh, Stereotype threat? Effects of inquiring about test takers' gender on conceptual test performance in physics AIP Conf. Proc. **1697,** 1 (2015).

20. P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, Am. J. Phys. **60**, 637 (1992).

21. M. Sabella, K. Coble and S. Bowen, Using the resources of the student at the urban, comprehensive university to develop an effective instructional environment, in AIP Conference Proceedings **1064**, 42 (2008).

22. R. Steinberg and K. Donnelly, PER-based reform at a multicultural institution, Phys. Teach. **40,** 108 (2002).

23. M. Sabella and A. Van Duzor, Cultural toolkits in the urban physics learning community, in AIP Conference Proceedings, **1513**, 34 (2013).

24. I. Halloun and D. Hestenes, The initial knowledge state of college physics students, Am. J. Phys. **53**, 1042 (1985).

25. J. Eccles, Understanding women's educational and occupational choices, Psychology of women quarterly **18**, 585 (1994).

26. Z. Hazari, G. Potvin, R. Lock, F. Lung, G. Sonnert, and P. Sadler, Factors that affect the physical science career interest of female students: Testing five common hypotheses Phys. Rev. ST Phys. Educ. Res. **9**, 020115 (2013).

27. Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, Journal of Research in Science Teaching **47**, 978 (2010).

28. M. Wang and J. Degol, Motivational pathways to stem career choices: Using Expectancy-Value perspective to understand individual and gender differences in stem fields, Dev. Rev. **33**, 304 (2013).

29. J. Blickenstaff, Women and science careers: leaky pipeline or gender filter? Gender and Education **17**, 369 (2005).

30. A. Diekman, E. Clark. A. Johnston, E. Brown and M. Steinberg, Malleability in communal goals and beliefs influences attraction to STEM careers: Evidence for a goal congruity perspective, J. Pers. Soc. Psychol. **101**, 902 (2011).

31. A. Kelly, Social cognitive perspective of gender disparities in undergraduate physics, PRPER **12**, 020116 (2016).

32. R. Ivie, R. Czujko and K. Stowe, *Women physicists speak: the 2001 international study of women in physics* (Washington, DC, American Institute of Physics, 2002).

33. R. Ivie and K. Stowe, *Women in physics 2000 (American Institute of Physics Rep. No. R-430)* (Washington, DC, American Institute of Physics, 2000).

34. E. Seymour and N. Hewitt, Talking About Leaving: Why Undergraduates Leave the Sciences (Westview, Boulder, CO, 1997).

35. A.C.K. Leung, A. Terrana, S. Jerzak, Students' opinions on the educational value of physics laboratories: a cross-sectional survey, Can. J. Phys. **94**, 913 (2016) doi: 10.1139/cjp-2016-0358

36. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemp. Educ. Psychol. **25**, 82 (2000); A. Bandura, *Self-efficacy: The exercise of control* (W. H. Freedman, New York, NY, 1997).

37. R. W. Lent, S. D. Brown and K. C. Larkin, Self-efficacy in the prediction of academic performance and perceived career options, J. Counsel. Psychol. **33**, 265 (1986).

38. J. Nilssen and J. Shemwell, Gender, experience, and self-efficacy in introductory physics, PRPER **12**, 020105 (2016).

39. R. W. Lent, S. D. Brown and K. C. Larkin, Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking, J. Counsel. Psychol. **34**, 293 (1987).

40. H. Fencl and K. Scheel, Research and teaching: engaging students - an examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course, J. Coll. Sci. Teaching **35**, 20 (2005).

41. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students, Int. J. Behav. Dev. **14**, 153 (1991).

42. G. Trujillo and K. D. Tanner, Considering the role of affect in learning: monitoring students' self-efficacy, sense of belonging, and science identity, CBE Life Sci. Educ. **13**, 6 (2014).

43. G. Quan and A. Elby, Connecting self-efficacy and views about the nature of science in undergraduate research experiences, PRPER **12**, 020140 (2016).

44. K. Miller, J. Schell, A. Ho, B. Lukoff and E. Mazur, Response switching and self-efficacy in Peer Instruction classrooms, PRST PER **11**, 010104 (2015).

45. M. Bernacki, T. Nokes-Malach and V. Aleven, Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning, Metacognition Learning **10**, 99 (2015).

46. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1096 (2012).

47. A. Duckworth, C. Peterson, M. Matthews and D. Kelly, Grit: Perseverance and Passion for Long-Term Goals, Personality Processes and Individual Differences **92**, 1087 (2007)

48. A. Duckworth and P. Quinn, Development and validation of the short grit scale (Grit-S), Journal of Personality Assessment **91**, 166 (2009).

49. V. Tinto, Classrooms as communities: Exploring the educational character of student persistence, J. High. Educ. **68**, 599 (1997).

50. T. Strayhorn, What role does grit play in the academic success of Black male collegians at predominantly white institutions? Journal of African American Studies **18**, 1 (2014).

51. K. Powell, Higher education: On the lookout for true grit, Nature **504**, 471 (2013).

52. P. Akos and J. Kretchmar, Investigating grit at a non-cognitive predictor of college success, The Review of Higher Education **40**, 163 (2017).

53. C. Dweck, *Mindset: The New Psychology of Success* (Ballentine, New York, 2006); C. Dweck, *Self-theories: Their Role in Motivation, Personality, and Development* (Psychology Press, Philadelphia, 1999).

54. L. Blackwell, K. Trzesniewski and C. Dweck, Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention, Child Development **78**, 246 (2007).

55. J. Aronson, M. Lustina, C. Good and K. Keough, When white men can't do math: Necessary and Sufficient factors in stereotype threat, Journal of Experimental Social Psychology **35**, 29 (1999).

56. C. Steele and J. Aronson, Stereotype threat and the intellectual test performance of African Americans, Attitudes and Social Cognition **69**, 797 (1995).

57. J. Aronson, C. Fried and C. Good, Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence, Journal of Experimental Social Psychology **38**, 113 (2002).

58. D. Hammer, Epistemological beliefs in introductory physics, Cognition Instruct. **12**, 151 (1994).

59. E. Redish, J. Saul and R. Steinberg, Student expectations in introductory physics, American Journal of Physics **66**, 212 (1998).

60. K. Edmondson and J. Novak, The interplay of scientific epistemological views, learning strategies, and attitudes of college students, Journal of Research in Science Teaching **30**, 547 (1993).

61. G. Kortemeyer, Correlations between student discussion behavior, attitudes, and learning, Phys. Rev. ST. PER. **3**, 010101 (2007).

62. A. Maries, S. Y. Lin and C. Singh, The impact of students' epistemological framing on a task requiring representational consistency, Proceedings of the 2016 Physics Education Research Conference, Sacramento, CA 212-215 doi:10.1119/perc.2016.pr.048

63. W.-M. Roth and A. Roychoudhury, "Physics students' epistemologies and views about knowing and learning," J. Res. Sci. Teach. **31**, 5 (1994).

64. Terrana, A.C.K. Leung and S. Jerzak, Use of online resources by physics students in Canada: A cross sectional study, Can. J. Phys. **95**, 201 (2017) doi: 10.1139/cjp-2016-077

65. H. Eshach and I. Kukliansky, Development of an instrument for assessing students' data analysis skills in the undergraduate physics laboratory, Can. J. Phys. **94**, 1205 (2016) doi: 10.1139/cjp-2016-0308

66. H. Eshach, The use of intuitive rules in interpreting students' difficulties in reading and creating kinematic graphs, Can. J. Phys. **92**, 1 (2014) doi:10.1139/cjp-2013-0369

67. E. M. Kennedy and J. R. de Bruyn, Understanding of mechanical waves among second-year physics majors, Can. J. Phys. **89**, 1155 (2011) doi:10.1139/p11-113

68. C. Kalman, M. Milner-Bolotin and T. Antimirova, Comparison of the effectiveness of collaborative groups and peer instruction in a large introductory physics course for science majors, Can. J. Phys., **88**, 325 (2010).

69. E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, NJ, 1997).

70. M. Milner-Bolotin, T. Antimirova and A. Petrov, Clickers beyond the First Year: Interactive Teaching in Upper Level Physics Courses, Journal of College Science Teaching, **40**, 18 (2010).

71. A. Wood, R. Galloway, R. Donnelly and J. Hardy, Characterizing interactive engagement activities in a flipped introductory physics class, PRPER **12**, 010140 (2016).

72. J. Stewart, S. DeVore, G. Stewart and L. Michaluk, Behavioral self-regulation in a physics class PRPER **12**, 010125 (2016).

73. E. Yerushalmi, E. Cohen, A. Mason and C. Singh, What do students do when asked to diagnose their mistakes? Does it help them? I. An atypical quiz context, Phys. Rev. ST Phys. Educ. Res. **8**, 020109 (2012).

74. E. Yerushalmi, E. Cohen, A. Mason and C. Singh, What do students do when asked to diagnose their mistakes? Does it help them? II. A more typical quiz context, Phys. Rev. ST PER **8**, 020110 (2012).

75. J. Farrell, R. Moog and J. Spencer, A Guided Inquiry Chemistry Course, Journal of Chemical Education, **76**, 570 (1999).

76. H. Hong and X. Lin-Siegler, How Learning About Scientists' Struggles Influences Students' Interest and Learning in Physics. Journal of Educational Psychology (2011) (doi: 10.1037/a0026224).

77. L. McDermott, Oersted Medal Lecture 2001: Physics Education Research—The Key to Student Learning, Am. J. Phys. **69**, 1127 (2001).

78. D. Hammer, Epistemological beliefs in introductory physics, Cognition and Instruction **12**, 151 (1994).

79. J. Piaget, *Success and Understanding* (Harvard University Press, Cambridge, MA, 1978).

80. G. Posner, K. Strike, W. Hewson and W. Gertzog, Accomodation of a scientific conception: Toward a theory of conceptual change, Science Education **66**, 211 (1982).

81. A. Collins, J. S. Brown and S. E. Newman, Cognitive Apprenticeship: Teaching the crafts of reading, writing and mathematics, in *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, edited by L. B. Resnick (Lawrence Erlbaum, Hillsdale, NJ, 1989) pp. 453-494.

82. S. Devore, E. Marshman and C. Singh, Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics, Phys. Rev. PER **13**, 010127 (2017).

83. D. Hestenes, M. Wells and G. Swackhamer, Force Concept Inventory, The Physics Teacher 30 141 (1992); D. 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); J. Li and C. Singh, Developing a magnetism conceptual survey and assessing gender differences in student understanding of magnetism, Proceedings of the Phys. Ed. Res. Conference, AIP Conf. Proc., Melville, New York **1413,** 43 (2012) doi:10.1063/1.3679989; J. Li and C. Singh, Developing and validating a conceptual survey to assess introductory physics students' understanding of magnetism, Euro J Phys **38**, 025702 (2017); C. Singh and D. Rosengrant, Multiple-choice test of energy and momentum concepts, Am. J. Phys **71**, 607 (2003); L. Rimoldini and C. Singh, Student understanding of rotational and rolling motion concepts, Phys. Rev. ST PER **1,** 010102 (2005); C. Singh, Student understanding of symmetry and Gauss's law of electricity, Am. J. Phys. **74**, 923 (2006); P. Engelhardt and R. Beichner, Students' understanding of direct current resistive electrical circuits, Am. J. Phys. **72**, 98 (2004); R. Beichner, Testing student interpretation of kinematics graphs, Am. J. Phys. **62**, 750 (1994).

84. W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006); B. Zwickl, T. Hirokawa, N. Finkelstein and H. Lewandowski, Epistemology and expectations survey about experimental physics: Development and initial results, Phys. Rev. ST Phys. Educ. Res. **10**, 010120 (2014).

85. E. Deci, H. Eghrari, B. Patrick and D. Leone, Facilitating internalization: The self-determination theory perspective, Journal of Personality **62**, 119 (1994).

86. H. Jang, Supporting students' motivation, engagement, and learning during an uninteresting activity, Journal of Educational Psychology **100**, 798 (2008).

87. P. Pintrich, A motivational science perspective on the role of student motivation in learning and teaching contexts, Journal of Educational Psychology **95**, 667 (2003).

88. P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, Am. J. Phys. **60**, 637 (1992).

89. R. Engle and F. Conant, Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learning classroom, Cognition and Instruction **20**, 399 (2002).

90. A. Brown, Transforming schools into communities of thinking and learning about serious matters, American Psychology **52**, 399 (1997).

91. W. Grolnick, S. Gurland, K. Jacob, and W. Decourcey, The development of self-determination in middle childhood and adolescence, in *Development of Achievement Motivation* (edited by A. Wigfield and J. Eccles), p. 147-171 (San Diego, CA, Academic Press).

92. B. Britton and A. Tesser, Effects of time-management practices on college grades, Journal of Educational Psychology **83**, 405 (1991).

93. S. Bates, R. Galloway, C. Lopston, and K. Slaughter, How attitudes and beliefs about physics change from high school to faculty, Phys. Rev. ST. PER. **7**, 020114 (2011).

94. V. Willson, C. Akcerman and C. Malave, Cross-time attitudes, concept formation, and achievement in college freshman physics, Journal of Research in Science Teaching **37**, 1112 (2000).

95. K. Slaughter, S. Bates and R. Galloway, A longitudinal study of the development of attitudes and beliefs towards physics in Proc. of the Physics Education Research Conference (2011).

96. E. Gire, B. Jones and E. Price, Characterizing the epistemological development of physics majors, Phys. Rev. STPER **5**, 010103 (2009).

97. C. Lindstrom and M. 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).

98. A. Traxler and E. Brewe, Equity investigation of attitudinal shifts in introductory physics, Phys. Rev. ST Phys. Educ. Res. **11**, 020132 (2015).

99. http://www.activationlab.org/tools/

100. J. Schell and B. Lukoff, Peer instruction self-efficacy instrument [Developed at Harvard University]. Unpublished Instrument coetta (2010).

101. S. Glynn, P. Brickman, N. Armstrong and G. Taasoobshirazi, Science motivation questionnaire II: Validation with science majors and nonscience majors, Journal of Research in Science Teaching **48**, 1159 (2011).

102. B. Reynolds, M. Mehalik, M. Lovell and C. Schunn, Increasing student awareness of and interest in engineering as a career option through design-based learning, International Journal of Engineering Education **25**, 788 (2009).

103. L. Carli, Gender and social influence, Journal of Social Issues **57**, 725 (2001).

104. C. Mueller and C. Dweck, Praise for intelligence can undermine children's motivation and performance, Journal of Personality and Social Psychology **75**, 33 (1998).

105. C. Good, J. Aronson and M. Inzlicht, Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat, Journal of Applied Development Psychology **24**, 645 (2003).

2.8 Appendix A

Additional plots

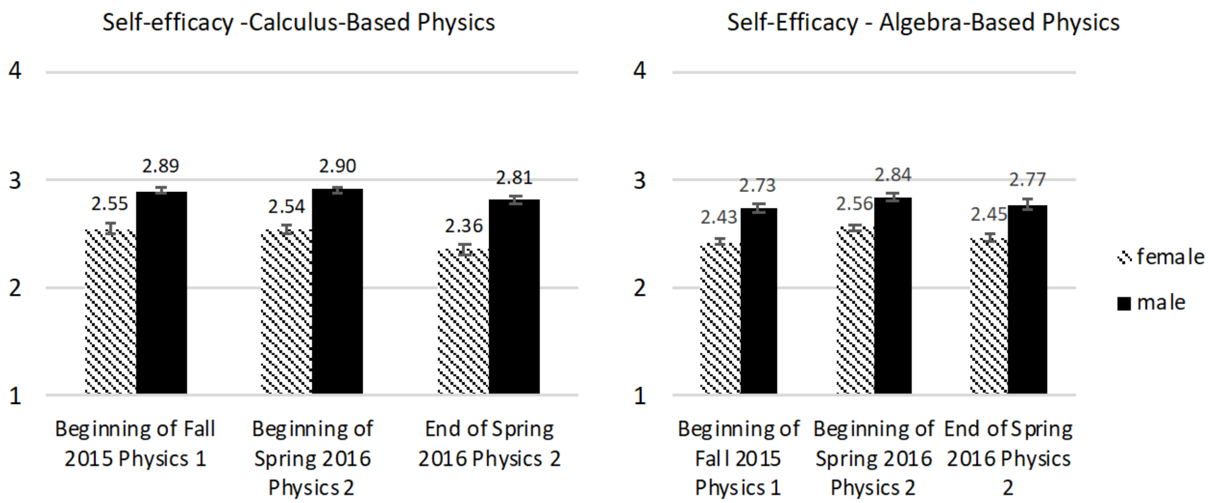

Figure 2-7 Female and male students' average self-efficacy scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

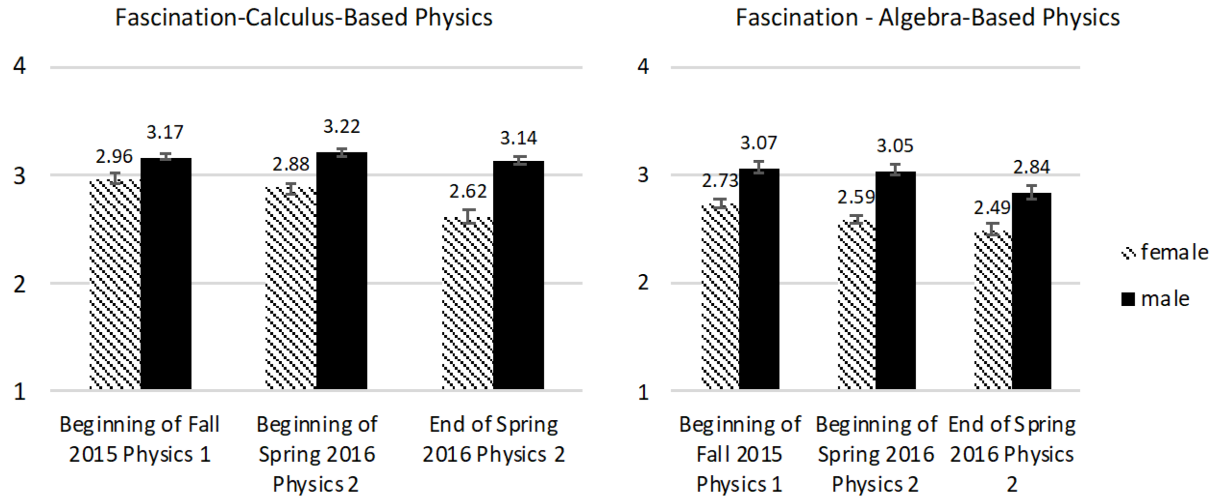

Figure 2-8 Female and male students' average fascination scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

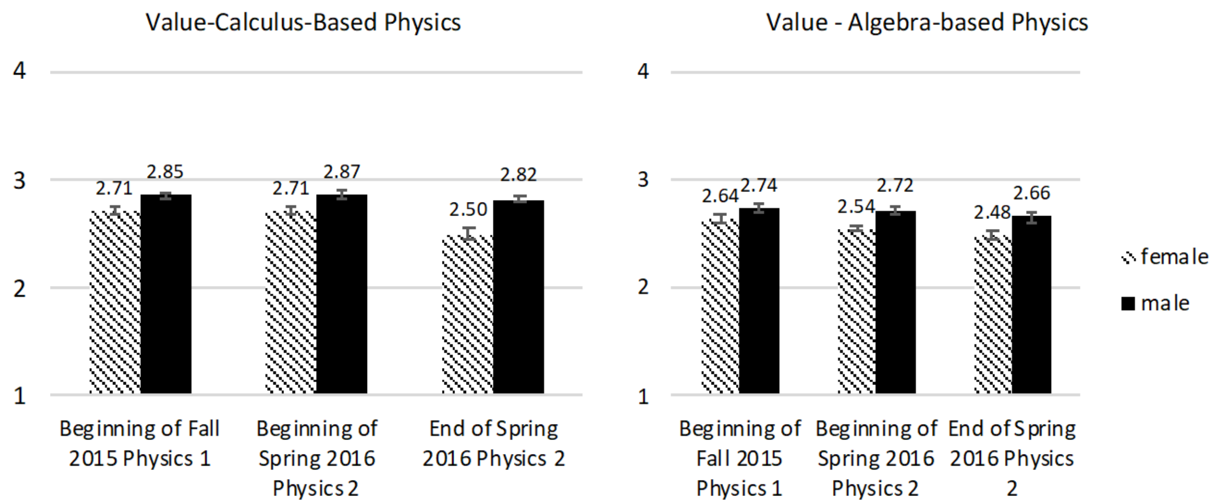

Figure 2-9 Female and male students' average value scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

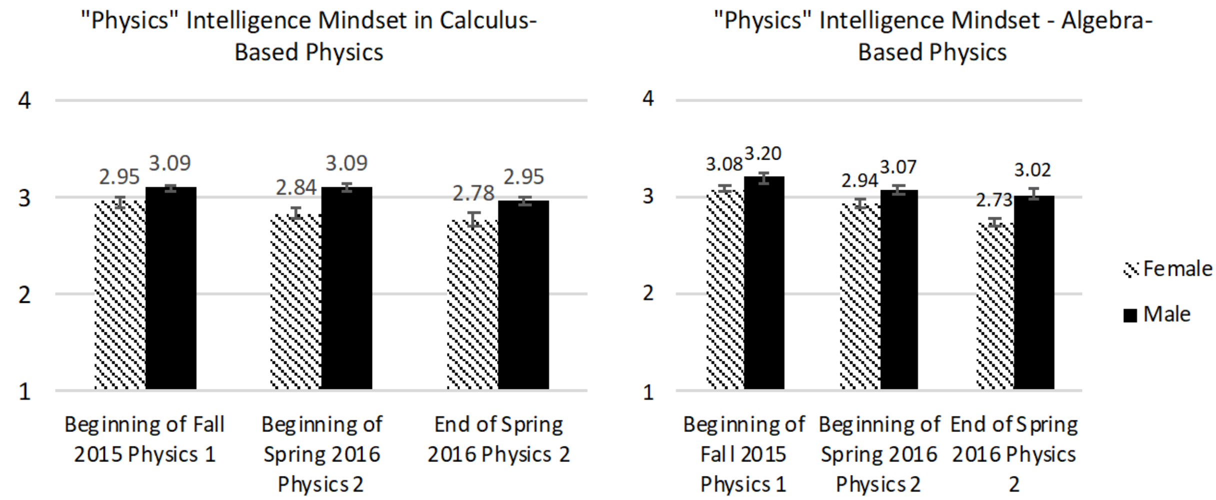

Figure 2-10 Female and male students' average "physics" intelligence mindset scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

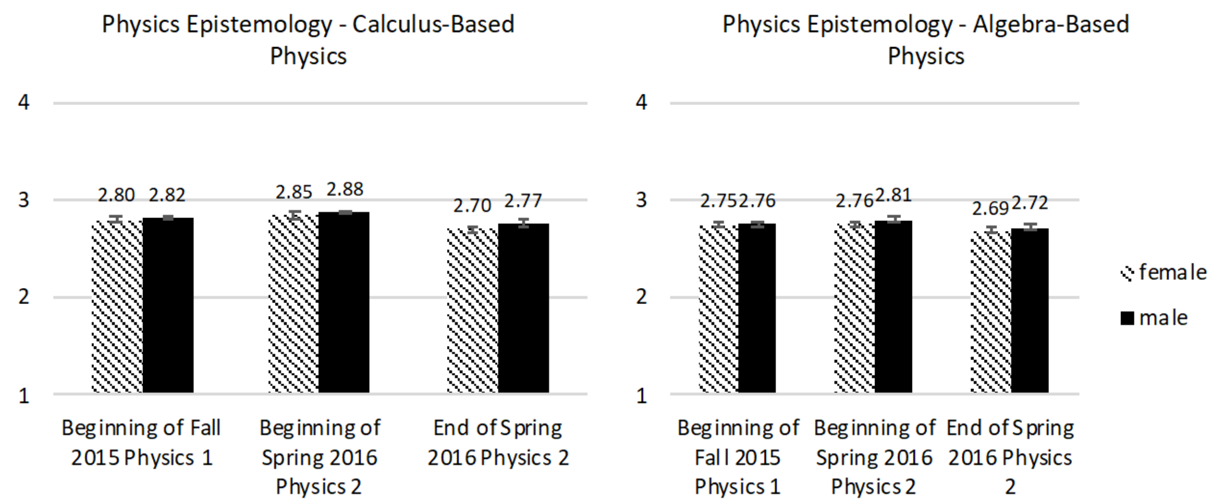

Figure 2-11 Female and male students' average physics epistemology scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

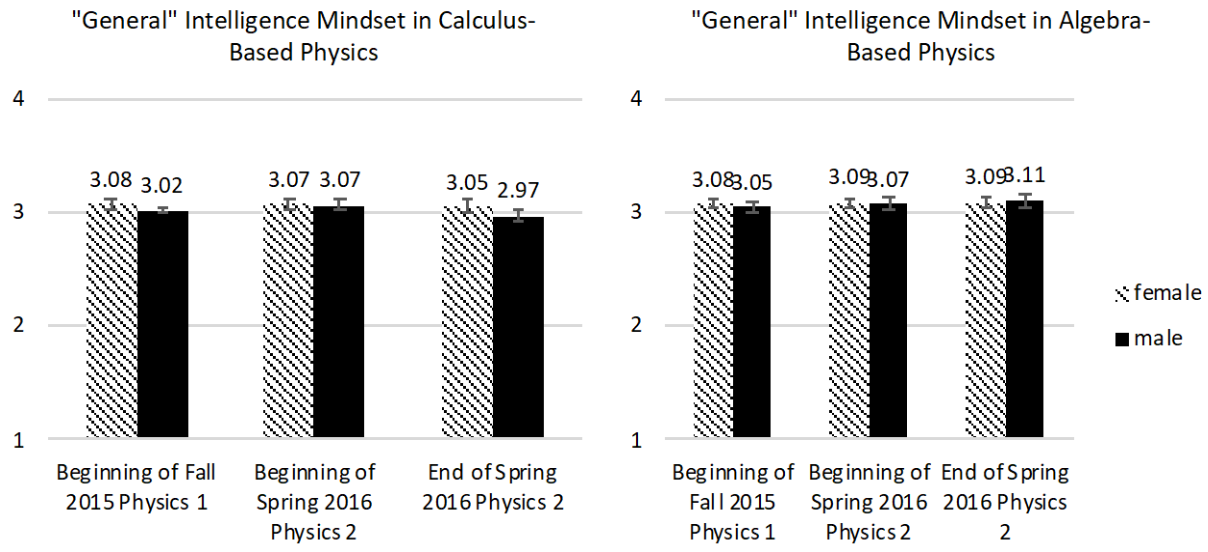

Figure 2-12 Female and male students' average intelligence mindset scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

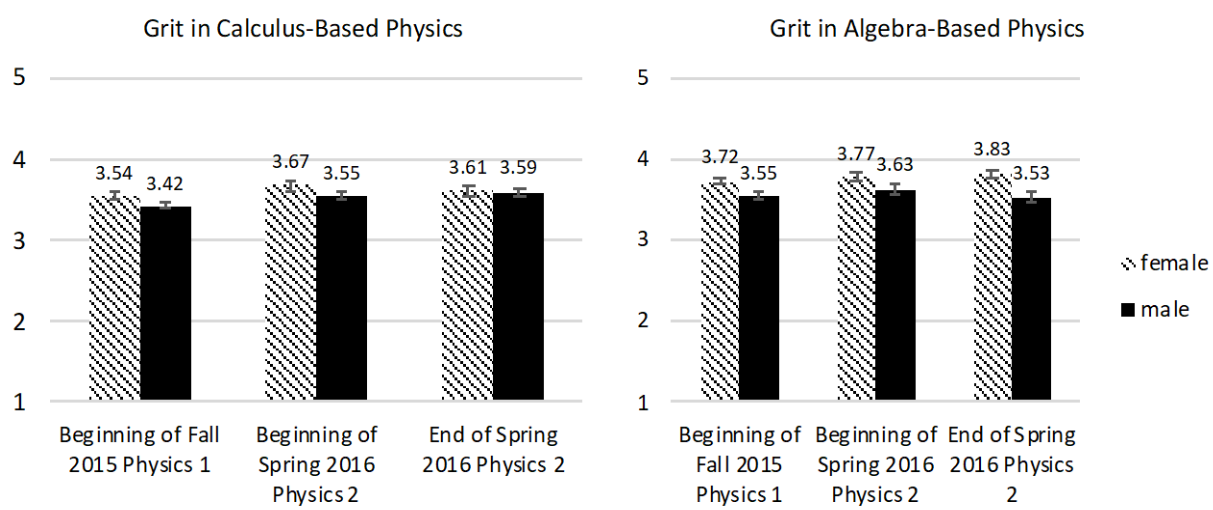

Figure 2-13 Female and male students' average grit scores throughout an introductory physics course sequence for calculus-based courses (left) and algebra-based courses (right) with standard error bars.

Survey and Exploratory Factor Analysis

1. I wonder about how nature works
    a. Never
    b. Once a month
    c. Once a week
    d. Every day
2. In general, I find physics
    a. Very boring
    b. boring
    c. interesting
    d. Very interesting
3. Knowing physics helps me understand how the world works:
    a. Never
    b. Sometimes
    c. Most of the time
    d. All of the time
4. Knowing physics is important for:
    a. No jobs
    b. A few jobs
    c. Most jobs
    d. All jobs
5. I can complete the physics activities I get in a lab class
    a. Rarely
    b. Half the time
    c. Most of the time
    d. All the time
6. If I went to a museum, I could figure out what is being shown about physics in:
    a. None of it
    b. A few areas
    c. Most areas
    d. All areas

		No!	no	yes	Yes
7.	I want to know everything I can about physics	a	b	c	d
8.	Physics makes the world a better place to live.	a	b	c	d
9.	Knowing physics is important for being a good citizen.	a	b	c	d
10.	Learning physics helps me understand situations in my everyday life.	a	b	c	d
11.	I am often able to help my classmates with physics in the laboratory or in recitation.	a	b	c	d
12.	I get a sinking feeling when I think of trying to tackle difficult physics problems.	a	b	c	d
13.	If I wanted to, I could be good at doing physics research.	a	b	c	d
14.	If I study, I will do well on a physics test.	a	b	c	d

		Strongly disagree	Disagree	Agree	Strongly agree
15.	You have a certain amount of intelligence, and you can't really do much to change it.	a	b	c	d
16.	No matter how much intelligence you have, you can always change it quite a bit.	a	b	c	d
17.	Anyone can become good at solving physics problems through hard work.	a	b	c	d
18.	Only very few specially qualified people are capable of really understanding physics.	a	b	c	d

Figure 2-14 Motivational survey questions from 1 to 18 and their response options.

		Not like me at all	Not much like me	Somewhat like me	Mostly like me	Very much like me
19.	I often set a goal but later choose to pursue a different one.	a	b	c	d	e
20.	I have difficulty maintaining my focus on projects that take more than a few months to complete.	a	b	c	d	e
21.	I finish whatever I begin.	a	b	c	d	e

		Strongly disagree	Disagree	Agree	Strongly agree
22.	I do not expect to understand physics equations in an intuitive sense; they must just be taken as givens.	a	b	c	d
23.	Knowledge in physics consists of many pieces of information each of which applies primarily to a specific situation.	a	b	c	d
24.	Learning physics is mainly about remembering the laws, principles, and equations given in class and/or in the textbook.	a	b	c	d
25.	Problem solving in physics mainly involves matching problems with facts or equations and then substituting the values to get a number.	a	b	c	d
26.	In doing a physics problem, if my calculation gives a result that differs significantly from what I expect, I'd have to trust the calculation.	a	b	c	d
27.	"Understanding" physics basically means being able to recall something you've read or been shown.	a	b	c	d
28.	When doing an experiment, I try to understand how the experimental setup works.	a	b	c	d
29.	The primary purpose of doing a physics experiment is to confirm previously known results.	a	b	c	d

Figure 2-15 Motivational survey questions from 19 to 29 and their response options.

Table 2-4 Factor analysis results for the motivational survey

Rotated Component Matrix[a]						
	Component					
	1	2	3	4	5	6
I wonder about how nature works		.276		.337		
In general, I find physics	.383	.456		.444		
Knowing physics helps me understand how the world works:	.221	.562		.350		
Knowing physics is important for:		.641				
I can complete the physics activities I get in a lab class	.678					
If I went to a museum, I could figure out what is being shown about physics in...	.572	.277				
I want to know everything I can about physics	.302	.552		.423		
Physics makes the world a better place to live	.244	.496		.228		
Knowing physics is important for being a good citizen		.719		-.210		
Learning physics helps me understand situations in everyday life		.661				
I am often able to help my classmates with physics in the laboratory or in recitation	.743					
I get a sinking feeling when I think of trying to tackle difficult physics problems	.595		.226	.286		
If I wanted to, I could be good at doing physics research	.582	.240		.322		
If I study, I will do well on a physics test	.509			.336	.230	
You have a certain amount of intelligence, and you can't really do much to change it					.825	
No matter how much intelligence you have, you can always change it quite a bit					.814	
Anyone can become good at solving physics problems through hard work				.411	.511	
Only very few specially qualified people are capable of really understanding physics			.210	.306	.444	
I often set a goal but later choose to pursue a different one						.764
I have difficulty maintaining my focus on projects that take more than a few months to complete						.772
I finish whatever I begin						.743
I do not expect to understand physics equations in an intuitive sense; they must be taken as givens	.311		.405	.358		
Knowledge in physics consists of many pieces of information each of which applies primarily to a specific situation			.584			
Learning physics is mainly about remembering the laws, principles, and equations given in class and/or in the textbook			.666			
Problem solving in physics mainly involves matching problems with facts or equations and then substituting the values to get a number			.646			
In doing a physics problem, if my calculation gives a result that differs significantly from what I expect, I'd have to trust the calculation			.392	.332		
Understanding physics basically means being able to recall something you've read or been shown.			.671			
When doing an experiment, I try to understand how the experimental setup works				.563		
The primary purpose of doing a physics experiment is to confirm previously known results			.547			

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 10 iterations.
b. Factor loadings less than 0.20 were omitted.

3.0 Female A's have similar physics self-efficacy as male students with C's in introductory courses: a cause for alarm?

3.1 Introduction

3.1.1 Underrepresentation of women in STEM fields

In many countries, there has been much focus on the underrepresentation of women in engineering and physical science fields, from elementary to high school, at the undergraduate level, in career fields, and in leadership positions. In the United States, the numbers of women studying science, technology, engineering, and mathematics (STEM) and pursuing STEM careers have not changed significantly in the last decade [1]. For example, since 2000, women have earned approximately 20% of the bachelor's degrees awarded in physics and engineering, with a similar under-representation at the Masters and PhD levels [2-4]. Some efforts have been made to increase the diversity in STEM courses and STEM occupations, yet, the reasons for the low percentage of women in science and engineering are not fully understood and progress has been slow (e.g., less than 1% change over the period from 2006 to 2014).

3.1.2 Self-efficacy and engagement in STEM fields

Deciding to pursue a STEM career and continuing a pathway to physical sciences and engineering is generally affected by several interrelated factors, including students' prior preparation and skills [5-16], quality of teaching and type of teaching approach [17-21], sociocultural factors [22-35], and motivational factors [36-49]. Motivational factors, such as intelligence mindset, interest in science, and self-efficacy in science, can also influence students' decisions to major and persist in STEM fields [36-38]. In particular, self-efficacy is the belief in one's capability to be successful in a particular task, course, or subject area [40-43] and it is one aspect of motivation. Self-efficacy can impact one's interests [44]. Because self-efficacy can shape interest, it also influences engagement during learning [45]. Furthermore, students with high self-efficacy in a domain often enroll in more difficult courses in that domain than those with low self-efficacy because they perceive difficult tasks as challenges rather than threats [46]. Self-efficacy in STEM also predicts initial career goals and enrollment in STEM courses [37,43,47] as well as persistence toward long-term career goals [44,48].

Self-efficacy in particular can have large long-term effects because it can thrust a student into a feedback loop which can impact students' self-efficacy and performance in a positive or negative way. As students complete short-term goals (e.g., complete a physics course requirement for a physical science or engineering major), they obtain feedback about their performance that shapes their self-efficacy. More surprisingly, research suggests that the reverse effect also occurs: for a variety of reasons, self-efficacy beliefs constrain performance in science courses beyond the more normative effects of students' prior knowledge and skills [49]. For example, students with a high self-efficacy are more likely to exhibit effective learning strategies such as self-monitoring [44,50,51] and tend to make more efficient use of problem-solving strategies and time-management [44]. In addition, students with high self-efficacy are less likely to reject correct hypotheses prematurely and tend to be better at solving conceptual problems than students with low self-efficacy but equal performance [52]. As a result, there are feedback loops in which higher initial self-efficacy produces higher performance which further strengthens self-efficacy; by contrast, lower initial self-efficacy produces lower performance which further weakens self-efficacy.

Unfortunately, several studies have shown that female students have significantly lower self-efficacy than male students in STEM-related domains [23,53,54] including in mathematics, engineering [14,24,55], and computer science [56,57]. In physics, female students report significantly lower self-efficacy than male students [58-62], even in interactive engagement courses [60] and even if female students had received prior physics instruction in high school [61]. Furthermore, female students experience a decline in self-efficacy throughout their engineering education at college [55]. A strong sense of self-efficacy, especially for female students in physical science and engineering courses, can help them persist in STEM fields [63-67]. Given the central role of self-efficacy in career choice and career persistence, the self-efficacy gap likely contributes to the low representation of women pursuing careers in physical science and engineering fields [21]. Therefore, understanding the source and nature of the gap is important.

3.1.3 Rationale for the study and research questions

Prior studies have mainly focused on using self-efficacy to predict performance outcomes [44,49,58]. However, this study aims to investigate whether there are differences in the self-efficacy of male and female students throughout introductory physics courses at matched performance levels. In particular, it is important to investigate whether self-efficacy is a reflection of students' actual performance, or whether there is a self-efficacy gender gap over and above performance that may be due to environmental factors such as stereotype threat. There are two types of performance that we look at: conceptual learning test results and students' grades. Differences in self-efficacy even within similarly performing male and female students can have many detrimental effects. A large underestimate of one's capability and/or performance can impact interest and goals. For example, if female students inaccurately perceive that they are not capable of succeeding in a STEM field, that could lead to decreased interest in STEM disciplines and underrepresentation of women in STEM fields.

In mathematics, prior research has shown that male students assess their own mathematical ability more favorably than female students of similar ability in high school [68]. Male students were more likely than equally performing female students to enroll in calculus courses in high school, and this difference was shown to be due to differences in male and female students' self-efficacy [68]. Since enrolling in calculus in high school has a large influence on the decision of women to choose a STEM major, female students' self-efficacy in mathematics is correlated with whether they enroll in advanced mathematics courses in high school and whether they decide to major in a STEM discipline [68]. Research also suggests that many women believe that they must achieve at exceptionally high levels in mathematics and science to be successful STEM professionals [69]. Interviews with and surveys of women in engineering programs suggest that the exit of women from engineering programs is not driven primarily by their performance or success, but partly because women have low self-efficacy and negatively interpret their grades [38]. For example, female students may view a "B" grade as a poor performance even though a "B" grade is above average [68]. Furthermore, research suggests that women who are equally successful as men at a mathematical task are less likely to compete at the task at the same rate as male students [70]. Much less research has examined whether self-efficacy in science has a similarly gendered misperception of ability.

The most common dropout points for women from physical science or engineering pathways are during the first and second years of college [38]. Since college level physics is a key prerequisite to obtaining a physical science or engineering degree and students usually take physics sequences in their first year, it is important to examine the self-efficacy of similarly performing male and female students in college level physics courses for scientists and engineers. However, research has not systematically examined the self-efficacy of male and female students who have similar performance outcomes in physics. It may be that female students simply have lower performance levels and lower self-efficacy that reflects the lower performance levels; alternatively, it may be that female students' self-efficacy in physics is below similarly performing male students in college level physics courses. Women who have high standards for achievement in physics may drop out of physics courses and leave a STEM major at a higher rate than male students if they underestimate their own capability to succeed in physics. Thus, we systematically investigated the self-efficacy of similarly performing male and female students in the context of the typical two-semester physics course sequence for engineering and physical science majors.

Matching students by performance requires selecting a common performance metric. Student performance can be assessed in many different ways, and there can be gender biases in particular assessment formats [71-77]. For example, grades based upon participation or explanation quality in open-ended responses may reflect biases in the evaluator or rater. Therefore, using standardized research-based assessments provides one strong method for examining this topic. However, performance shapes self-efficacy through explicit feedback, and so the grades that students obtain in courses provide another important performance metric for this topic. Therefore, we examine self-efficacy by gender in two different ways: once matched by research-based standardized conceptual physics assessments and again matched by course grades.

Since prior research has suggested that instructor attitudes, assessment practices, and course format can differentially influence learning, attitudes, and retention in STEM courses

[19,28,29,30,78,79], we also examine students' self-efficacy across different course types (e.g., flipped vs. traditional format) and different instructors. In particular, it has been found in some research studies that students' self-efficacy is impacted by different teaching methodologies [58,60,61,80-84]. Our investigation can shed light on whether self-efficacy gender differences at matched performance levels are broadly generalizable across different instructional contexts.

In particular, our research questions (RQ) are as follows:

RQ1. What is the self-efficacy of female and male students throughout a two-semester physics course sequence when prior knowledge differences on physics conceptual surveys are accounted for?

RQ2. What is the self-efficacy of female and male students throughout a two-semester physics course sequence when students' course performance is accounted for?

RQ3. Are the effects consistent across different instructors and different types of physics courses, e.g., flipped vs. traditional formats in physics 1 and physics 2?

3.2 Methodology

To investigate the self-efficacy of introductory students with similar performance outcomes in physics, we administered a motivation survey and physics conceptual assessments in several sections of a two-course calculus-based introductory physics sequence. We collected data across two consecutive academic years.

3.2.1 Participants and class context

Participants were students enrolled in 9 sections in Physics 1 and 11 sections in Physics 2. These two large introductory physics courses arranged in a two-semester sequence at the University of Pittsburgh, a large R1 public university. This calculus-based physics sequence is typically taken by engineering and physical science majors as a requirement. Physics 1 includes topics such as kinematics, forces, energy and work, rotational motion, gravitation, and oscillations and waves. Physics 2 includes topics such as electricity and magnetism, electromagnetic waves, reflection, interference, and diffraction. Both courses included four lecture hours and one recitation hour per week. The recitations were mandatory and included a weekly, low stakes quiz. Most of the sections of both courses were taught in a traditional, lecture-based format, but there were four sections of Physics 1 and four sections of Physics 2 that were taught in a "flipped" format [85,86]: students watched lecture videos before attending the lectures, during which they worked on collaborative group problem solving and clicker questions (in which the students answered questions individually, discussed their answers with a peer, and then answered the questions again).

Table 3-1 shows the number of students in introductory physics courses who completed the motivation survey and/or physics conceptual survey throughout the introductory physics course sequence. The number of students is different at different points of time because some students did not take Physics 2 and some students were not present in the lecture or recitation section in which the surveys were administered (however, when the analysis is performed again for matched students, the results are qualitatively similar).

The cohort of both courses was approximately 32% female. Seventy-five percent of the students were White, 13% of the students were Asian, and 12% of the students were Black, Hispanic, or Multi-racial. Most of the students were between 18 and 19 years old. Sixty-eight percent of the students were enrolled in an engineering track, and the rest were science majors or other majors requiring physics.

Table 3-1 The number of students completing the self-efficacy survey and the physics conceptual assessments at different points in time.

	Motivation Survey		Conceptual Assessment	
	Pre	**Post**	**Pre**	**Post**
Physics 1	N=1,054	data not collected	N=726	N=644
Physics 2	N=914	N=630	N=845	N=807

3.2.2 Validity and reliability of the survey

The motivation survey was adapted and validated based upon previously developed survey instruments [82,87-89] and used in prior research studies [62,90,91]. The survey includes questions focused on several aspects of motivation, including interest and value associated with physics, intelligence mindset, and self-efficacy. In this paper, we focus only on students' responses to the self-efficacy questions. Table 3-2 shows the physics self-efficacy survey items, which all involved 4-point Likert scales. We note that the scale for some of the self-efficacy questions was in the form of "NO!, no, yes, YES!" This response scale has been extensively validated in prior studies, including studies of self-efficacy [54]. Our findings are similar to the validation of these rating scales in prior investigation [54]. In particular, this scale was used as opposed to "strongly disagree, disagree, agree, strongly agree" because students interpret these rating scales appropriately and because it reduces students' cognitive load, which is especially important for non-native speakers and for questions that ask raters to consider subtle differences in survey items [54]. The four-point scale also allows printing the scale next to each item so students do not need to search the instructions or map back to the initial scale as some surveys do, thereby further reducing potential errors in mapping or misremembering scale numbers. The physics self-efficacy survey items represent a diverse set of contexts and forms of interaction with physics and are meant to capture a general sense of efficacy with physics content. The item types are purposely varied to encourage processing of each item, rather than quickly responding similarly to each item. Internal coherence and discriminability from other motivational constructs (e.g., identity, interest, and valuing of physics) was established through factor analysis, and item separation is shown in detail in our previous work [62]. We measured Cronbach's alphas for each construct to check internal consistency of the questions. The Cronbach's alpha is above 0.70 for all of the constructs and for self-efficacy questions it is 0.74, which is considered adequate [91,92]. Individual interviews with students also provided useful feedback for refining or discarding survey items.

We also ensured content validity, i.e., the degree to which the survey items reflect the domain of interest (in our case, self-efficacy), by taking steps to ensure that the respondents interpreted the survey questions as was intended. We conducted one-on-one interviews with 12 students (8 male and 4 female students) in introductory, calculus-based physics courses and 3 students in graduate-level physics courses (2 male and 1 female students) using a think-aloud protocol [93] to verify that the students interpreted each question as intended. Students voluntarily participated in the interviews and were compensated for their time. Most of the students enrolled in the introductory physics courses planned to major in engineering or physical science. The interviewed students' responses were audio-recorded. Each interview took approximately an hour. During the individual interviews, students were asked to read each question aloud and explain how they interpreted the question. They were also asked to respond to the survey questions and give an explanation for their responses. During the interviews and discussions, we paid attention to respondents' interpretations of the questions and modified them accordingly in order to clarify their intent. We note that none of the interviewees had difficulty interpreting the “NO!, no, yes, YES!” response scale.

Table 3-2 The physics self-efficacy survey. One item indicated with an (R) is reverse coded.

<table>
<tr><th>Survey Item</th><th>Response options</th></tr>
<tr><td>1. I can complete the physics activities I get in a lab class</td><td>o Rarely
o Half of the time
o Most of the time
o All of the time</td></tr>
<tr><td>2. If I went to a museum, I could figure out what is being shown about physics in [see options to the right]:</td><td>a. None of it
b. A few areas
c. Most areas
d. All areas</td></tr>
<tr><td>3. I am often able to help my classmates with physics in the laboratory or in recitation.</td><td rowspan="4">a. No!
b. no
c. yes
d. Yes!</td></tr>
<tr><td>4. I get a sinking feeling when I think of trying to tackle difficult physics problems. (R)</td></tr>
<tr><td>5. If I wanted to, I could be good at doing physics research.</td></tr>
<tr><td>6. If I study, I will do well on a physics test.</td></tr>
</table>

3.2.3 Measures

Conceptual physics knowledge. In Physics 1, the Force Concept Inventory [94] was administered. In Physics 2, the Conceptual Survey of Electricity and Magnetism [95] was administered. Both assessments have been extensively validated as measures of conceptual understanding of core physics phenomena and principles within each of the two topic areas [94,95]. They also correspond to the earlier parts of the course that would be covered by all sections; coverage of later course content often varies across sections (e.g., optics topics are not consistently covered in Physics 2).

For a number of the analyses, we grouped the students by performance into three equal-width performance bins (i.e., the number of students in each bin was approximately the same). Since students learned content from pre to post, the cutoff scores were determined relative to the time point. See Table 3-3 for the cutoff scores for each bin at each time point.

Table 3-3The cutoff scores (in assessment percentages) for each FCI pretest/posttest and CSEM pretest/posttest performance bin, along with percentage of female students within each bin.

Assessment	Time point	Bin cutoff scores	% Female
FCI (Physics 1; 33% Female)	**pretest** N=726	Low (0%-45%)	52%
		Medium (46%-67%)	33%
		High (68%-100%)	14%
	posttest N=644	Low (0%-56%)	51%
		Medium (57%-81%)	33%
		High (82%-100%)	15%
CSEM (Physics 2; 32% Female)	**pretest** N=845	Low (0%-33%)	48%
		Medium (34%-44%)	26%
		High (45%-100%)	22%
	posttest N=807	Low (0%-42%)	44%
		Medium (43%-63%)	30%
		High (64%-100%)	26%

Course grades. We also obtained students' final grades in each course. Students' final grades are mostly based on high-stakes assessments that may have more impact on their self-efficacy than their performance on the physics conceptual surveys. Since some grades occur infrequently, for analysis, students were grouped into bins by grade ranges: 1) C- and below (considered insufficient to move on to the next course), 2) C and C+, 3) B-, B, B+, and 4) A-, A. Table 3-4 shows the grade distribution for students' final grades in each course, along with the percentage of students in each bin who are female.

Table 3-4 The letter grade distributions for students' final grades in each course, including the percentage of female students within each bin.

Course	Total	C- or below	C, C+	B-, B, B+	A-, A
Physics 1	N=1,130 33% Female	N=210 34% Female	N=375 44% Female	N=368 33% Female	N=177 24% Female
Physics 2	N=1,059 32% Female	N=195 30% Female	N=308 35% Female	N=384 34% Female	N=172 22% Female

3.2.4 Procedures

The self-efficacy survey was given at three time points: at the beginning of Physics 1, the beginning of Physics 2, and the end of Physics 2; the end of Physics 1 was excluded to avoid survey fatigue/redundancy with the beginning of Physics 2 only a few weeks later. The relevant physics conceptual assessment was given as a pre/posttest in each course. The self-efficacy survey and physics conceptual assessments were typically administered in the first and last recitations of the course, although one instructor chose to give the survey in the lecture portion of the course. In the first year of administration, some instructors chose to give the survey and assessments in a written format, whereas others chose to use an online format completed outside of class. We found that the participation rates were significantly lower for students who were given the online format. Thus, in subsequent administrations, only the written format was used. The self-efficacy survey was completed by most students in a couple of minutes (embedded in a larger motivational survey taking between 10-15 minutes), and the students worked through the conceptual physics assessments in the remaining class time (approximately 35-40 minutes).

Instructors were encouraged to give a small amount of course credit to students for completing the surveys. The instructor or teaching assistant responsible for giving the survey was given the following script to announce before administering the survey to the students to encourage students to take the assessments seriously: "We are surveying you on your understanding and beliefs about physics in order to improve the class. Your responses will not be evaluated for grades except to make sure the responses were done seriously, rather than randomly."The university provided course grades and student demographics (gender, ethnicity, and major), and this information was linked to the conceptual assessments and self-efficacy surveys through an honest broker; that is, the researchers only had access to the linked data in a de-identified form. We removed a small number of students whose gender was unidentified.

3.2.5 Analysis

To determine whether there are differences in the self-efficacy of female and male students (overall, combining different instructors and across flipped vs. traditional formats) controlling for performance on physics conceptual surveys, we performed a linear regression in which the dependent variable was students' average post self-efficacy score in either Physics 1 or Physics 2 and the independent variables were gender and the students' FCI or CSEM posttest scores. This analysis was also conducted separately for each instructor and format of physics course (i.e., flipped or traditional lecture-style format). This by-section analysis allowed us to determine whether gender had a significant impact on students' end of semester self-efficacy (controlling for their post FCI/CSEM scores) across various instructors and course formats.

To determine whether there are differences in the self-efficacy of female and male students (overall, combining different instructors and across flipped vs. traditional formats) controlling for performance on physics conceptual surveys, we performed a linear regression in which the dependent variable was students' average post self-efficacy score in either Physics 1 or Physics 2 and the independent variables were gender and the students' FCI or CSEM posttest scores. This analysis was also conducted separately for each instructor and format of physics course (i.e., flipped or traditional lecture-style format). This by-section analysis allowed us to determine whether gender had a significant impact on students' end of semester self-efficacy (controlling for their post FCI/CSEM scores) across various instructors and course formats.

The same analysis was conducted using grade bins rather than conceptual assessment bins. In this analysis, we focused only on self-efficacy differences between male and female students who passed the course (i.e., received a grade of C or better in the course since a C or better is required to remain on the engineering or physical science track). For each final letter grade, we calculated the effect size between male and female students' average self-efficacy scores and we checked whether the differences were significant using a two-way ANOVA. We also repeated this analysis at the level of individual self-efficacy survey items to ensure that the gender differences in self-efficacy were robust to physics self-efficacy more broadly, rather than just efficacy about a particular aspect of physics (e.g., test taking).

3.3 Results

3.3.1 Gender differences in self-efficacy, controlling for performance on standardized conceptual physics tests

In regards to Research Question 1, at the beginning of Physics 1, a two-way ANOVA (self-efficacy is the outcome, gender and conceptual test bin are the factors) revealed a small but significant gender difference ($F(1, 690) = 8.25$, $p < 0.01$), a significant effect of FCI bin (higher performing students had higher confidence levels, $p < 0.01$), and no significant interaction effect between gender and FCI bin. However, as shown in Table 3-5 and Figure 3-1 (upper left), male students' self-efficacy was higher than female students' primarily in the low and medium FCI performance groups. The effect size differences in female and male students' self-efficacy for the medium and low pre FCI bins were greater than 0.30 (but less than 0.50); which is considered a medium effect size. To contextualize the effect size, female students in the medium FCI group had the same self-efficacy as male students in the low FCI group.

At the end of Physics 1, the main effect of gender was also statistically significant (F (1, 318) = 22.28, $p < 0.001$) and even larger overall, and the interaction of gender and FCI performance bin was statistically significant (F (2, 317) = 3.76, $p < 0.05$). As shown in Table 3-5 and Figure 3-1 (upper right), the gender effect size was largest in the low group, moderate in the medium group, and approaching large in the high group. Interestingly for male students, self-efficacy was the same in the medium and low group (i.e., their self-efficacy was not influenced by relative performance levels). As a result of both patterns, the top performing female students had roughly the same self-efficacy as the lowest performing male students.

Table 3-5 Self-efficacy scores in Physics 1 binned by FCI scores (high, medium, low) for female and male students (M=mean, SD=standard deviation, N=number of students), along with Cohen's *d* for the gender contrast in means.

Self-Efficacy	FCI Physics 1																	
	Low						Medium						High					
	Pre			Post			Pre			Post			Pre			Post		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Female	2.58	0.47	122	2.30	0.54	46	2.76	0.35	68	2.65	0.32	32	3.04	0.38	34	2.83	0.43	20
Male	2.73	0.39	112	2.78	0.39	38	2.91	0.41	139	2.77	0.43	68	3.05	0.37	216	3.04	0.44	115
d size	0.33**	—	—	0.90***	—	—	0.37*	—	—	0.30	—	—	0.03	—	—	0.48*	—	—

d sizes are Cohen's d's between female and male students
* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001probability level

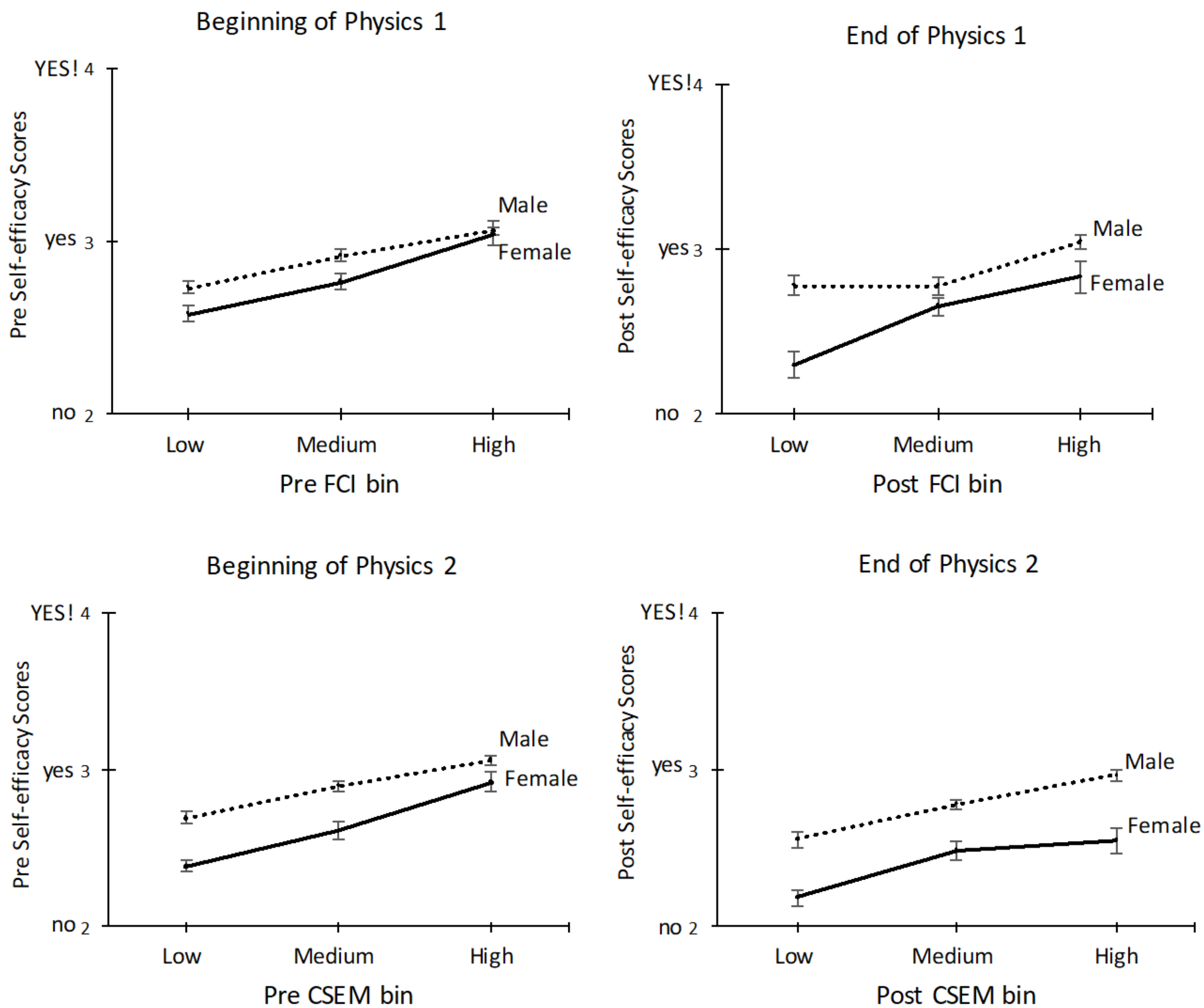

Figure 3-1 Pre self-efficacy (left column) and post self-efficacy (right column) scores of female and male students binned by temporally proximal conceptual test score in Physics 1 (top row) and Physics 2 (bottom row) courses. Error bars represent standard error of the mean.

At the beginning of Physics 2, both main effects of gender and CSEM performance bin were statistically significant (for gender, $F(1, 777) = 49.44$, $p < 0.001$; for CSEM bin, $F(2, 776) = 60.70$, $p < 0.001$), with no significant interaction effect between gender and CSEM bin. As shown in Table 3-6 and Figure 3-1 (lower left), the gender effect size was generally large (greater than 0.50), with female students in the medium CSEM group showing similar self-efficacy scores as male students in the low CSEM group, and female students in the high CSEM group having equal self-efficacy with male students in the medium group.

By the end of Physics 2, both main effects of gender and CSEM performance bin were statistically significant (for gender, $F(1, 594) = 71.94$, $p < 0.001$; for CSEM bin, $F(2, 593) = 27.00$, $p < 0.001$), with no significant interaction effect between gender and CSEM bin. In other words, students who earned higher scores in CSEM reported higher self-efficacy. As shown in Table 3-6 and Figure 3-1 (lower right), the gender differences become even larger for the high CSEM performance group, increasing from $d = 0.31$ to $d = 0.78$ (which is considered a large effect size). Importantly, at the end of this physics sequence, female students in the high CSEM group report similar post self-efficacy as male students in the low CSEM group.

Table 3-6 Self-efficacy scores in Physics 2 binned by CSEM scores (high, medium, low) for female and male students (M=mean, SD=standard deviation, N=number of students), along with Cohen's d for the gender contrast in means.

Self-Efficacy	CSEM Physics 2																	
	Low						Medium						High					
	Pre			Post			Pre			Post			Pre			Post		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Female	2.38	0.45	135	2.18	0.45	79	2.60	0.45	83	2.48	0.49	75	2.92	0.46	61	2.54	0.56	47
Male	2.69	0.45	137	2.56	0.47	90	2.90	0.40	182	2.77	0.45	154	3.06	0.43	200	2.96	0.47	150
d size	0.65***	—	—	0.74***	—	—	0.66***	—	—	0.61***	—	—	0.31*	—	—	0.78***	—	—

d sizes are Cohen's d's between female and male students
* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001probability level

3.3.2 Gender differences in self-efficacy controlling for physics course grade

In regards to Research Question 2 (self-efficacy by grade level), a two-way ANOVA (where self-efficacy is the outcome, gender and grade bin C, B, A are the factors) revealed that the differences in male and female students' self-efficacy were statistically significant ($F(1, 608) = 66.31$, $p < 0.001$), and the effect size was large (greater than 0.50; see Table 3-7 and Figure 3-2 top) in Physics 1. There was also a main effect of course grade ($F(2, 607) = 43.06$, $p < 0.001$) and no interaction effect between gender and grade ($F(2, 607) = 1.65$, $p = 0.19$). Female students had significantly lower self-efficacy compared to their male counterparts in all grade groups (A, B, C). Moreover, as the top of Figure 3-2 shows, female students receiving As have similar self-efficacy as male students receiving Cs in Physics 1.

Table 3-7 Self-efficacy scores in Physics 1 binned by course grade (A, B, C) for female and male students (M=mean, SD=standard deviation, N=number of students), along with Cohen's *d* for the gender contrast in means.

Post Self-Efficacy	Physics 1								
	C			B			A		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Female	2.33	0.42	86	2.59	0.45	80	2.90	0.41	31
Male	2.75	0.42	134	2.90	0.43	173	3.14	0.38	105
d size	0.89***	—	—	0.67***	—	—	0.60**	—	—

d sizes are Cohen's d's between female and male students' self-efficacy scores.
* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001probability level

Similarly, the gender gap in self-efficacy controlling for final course grade continues to persist in Physics 2 with equally large effect sizes. The ANOVA had significant effects of gender ($F(1, 555) = 52.44$, $p < 0.001$) and course grade ($F(2, 554) = 21.32$, $p < 0.001$), and no interaction (($F(2, 554) = 1.14$, $p = 0.32$). Students who earned higher grades reported higher self-efficacy. In other words, the effect of gender on self-efficacy does not change based on the course grade. Female students had much lower self-efficacy than did male counterparts at each matched letter grade group and the gender effect sizes are all large (greater than 0.50). Again, women receiving

As had similar self-efficacy scores to men receiving Cs in Physics 2 (see Table 3-8 and Figure 3-2).

Table 3-8 Self-efficacy scores in Physics 2 binned by course grade (A, B, C) for female and male students (M=mean, SD=standard deviation, N=number of students), along with Cohen's *d* for the gender contrast in means.

Post Self-Efficacy	Physics 2								
	C			B			A		
	M	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*
Female	2.30	0.51	70	2.43	0.49	87	2.76	0.54	27
Male	2.65	0.53	133	2.87	0.43	158	3.03	0.47	81
d size	0.62***	—	—	0.89***	—	—	0.60*	—	—

d sizes are Cohen's d's between female and male students' self-efficacy scores.
* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001probability level

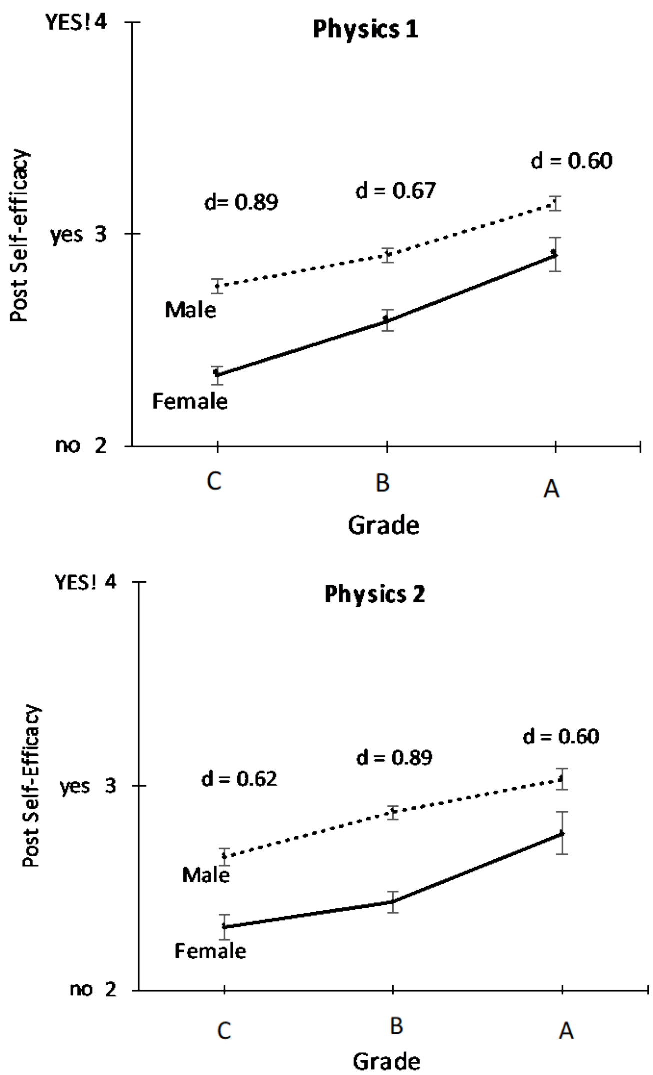

Figure 3-2 Post self-efficacy scores of female and male students binned by course grade (A, B, C) in Physics 1 (top) and Physics 2 (bottom) courses. Error bars represent standard error of the mean.

In order to explore what fraction of the self-efficacy gender gap was related to students' course performance versus students' biased perception, we first measured students' raw post self-efficacy differences between female and male students for Physics 1 and Physics 2, which are 0.39 ($p < 0.001$) and 0.40 ($p < 0.001$), respectively and which favors the male students. Next, we controlled for students' grade in Physics 1 and Physics 2 and performed an ANCOVA with gender and grades as independent variables and post self-efficacy as dependent variable. We found that the mean value of the gender gap in self-efficacy decreased only a small amount to 0.34 ($p < 0.001$) in Physics 1 and 0.38 ($p < 0.001$) in Physics 2 courses. In other words, a small part of the gender gap in self-efficacy might be attributable to differences in performance, but the gap mainly comes from biased perceptions: 87% in Physics 1 and 95% in Physics 2 (see Figure 3-3). Consistent with our results discussed earlier, existing gender gap in students' self-efficacy is largely due to students' self-perception rather than how they actually perform in the course. In other words, gender gap in self-efficacy does not come from performance differences by gender but mainly from biased beliefs about physics among female and male students.

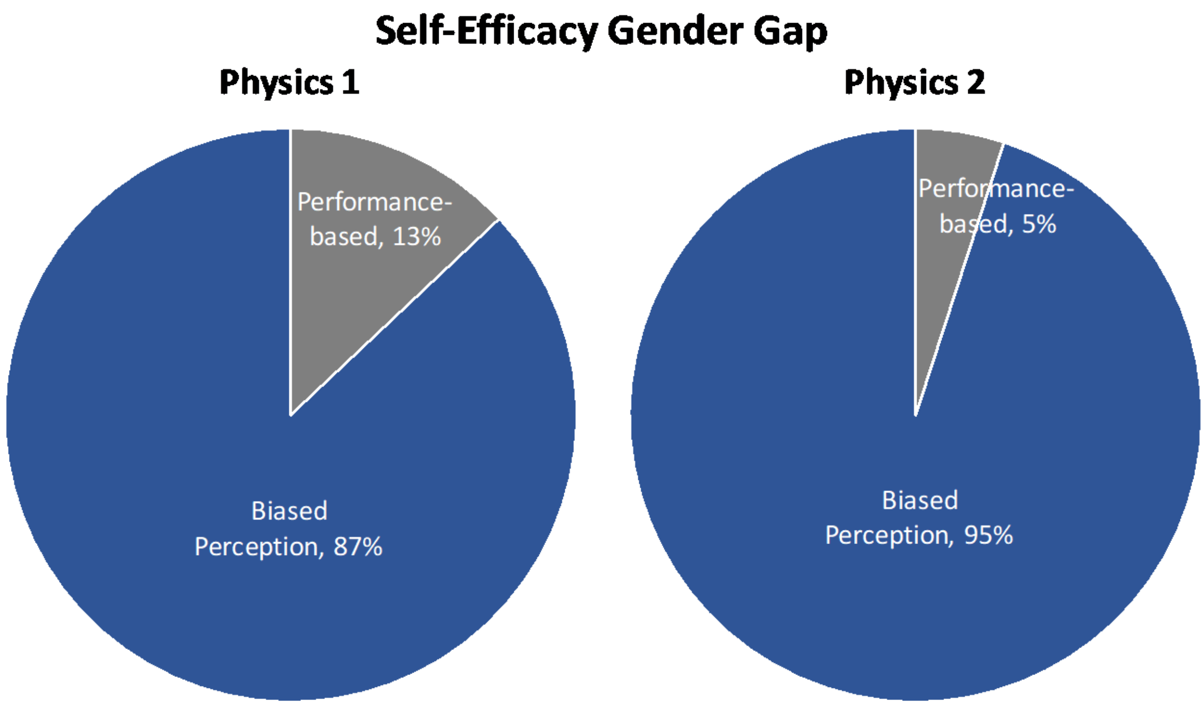

Figure 3-3 The relative contributions of performance and biased perception to the post self-efficacy gender gap in each course.

We also repeated the analysis with each self-efficacy survey question. This analysis allowed us to verify that the gender difference in self-efficacy was not dominated by only a few aspects of self-efficacy such as test-taking or in-class performance. Specifically, we performed a two-way ANOVA in which the dependent variable is the student's score on the survey item and the independent variables are gender (coded as 0 for female, 1 for male students) and course grade (0 to 4) in Physics 1 or Physics 2. The standardized regression coefficients for gender, denoted by β_1 in Eq. 1 are shown in Table 3-9. There was always a statistically significant gender effect—male students were more positive in their responses to all of the self-efficacy questions than female students in both Physics 1 and Physics 2 courses, even when taking into account students' course grades. The effects were slightly smaller for the questions about lab classes and physics research, but generally were consistent in size across questions.

$$\text{Self-Efficacy Survey Item Score} = \beta_1 \times \text{gender} + \beta_2 \times \text{grades} + \text{Constant} \quad (1)$$

Table 3-9 Gender effect β_1 values for each self-efficacy question at pre- and post-test in Physics 1 and Physics 2 courses. R indicates that the survey item was reverse coded.

Survey Item	Physics 1		Physics 2	
	Pre	Post	Pre	Post
1. I can complete the physics activities I get in a lab class	0.14*	0.21**	0.19***	0.19*
2. If I went to a museum, I could figure out what is being shown about physics in:	0.22***	0.27***	0.24*	0.28***
3. I am often able to help my classmates with physics in the laboratory or in recitation.	0.19***	0.23***	0.25***	0.25***
4. I get a sinking feeling when I think of trying to tackle difficult physics problems. (R)	0.24***	0.24***	0.25***	0.27***
5. If I wanted to, I could be good at doing physics research.	0.17*	0.17***	0.17*	0.18**
6. If I study, I will do well on a physics test.	0.18**	0.26***	0.24***	0.26***

* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001 probability level

3.3.3 Gender differences in self-efficacy across different course types and instructors

We also explored whether gender differences in self-efficacy exist for different instructors and teaching methods in introductory level physics courses to determine the generalizability of the findings of Research Questions 1 and 2. There were 9 sections in Physics 1 and 11 sections of Physics 2 courses in the dataset. All of the instructors were male, but with some having several years of teaching experience and others new to teaching. Some of these sections were led by the same instructor, and thus not a meaningful generality test, leading us to merge sections within Physics 1 or within Physics 2 that were taught by the same instructor. However, we did not merge sections across courses or if they were of different formats within a course (*traditional* vs. *flipped*).

As a result, we analyzed four sections of Physics 1 (three traditional sections and 1 flipped section) and four sections of Physics 2 (three traditional sections and 1 flipped section).

To formally quantify the gender effect on self-efficacy controlling for performance across the various physics courses, we performed a linear regression within each section in which post self-efficacy is the dependent variable, and gender and physics performance level (FCI or CSEM as appropriate) are the independent variables (see Eqs. 2 and 3). Of particular interest is the β_1 estimate (the gender effect) within each section.

$$\text{Post Self-Efficacy in Physics 1} = \beta_1 \times \text{gender} + \beta_2 \times \text{FCI pre scores} + \text{Constant} \quad (2)$$

$$\text{Post Self-Efficacy in Physics 2} = \beta_1 \times \text{gender} + \beta_2 \times \text{CSEM pre scores} + \text{Constant} \quad (3)$$

As shown in Table 3-10, the gender effects were statistically significant in all but one case across courses, instructors, and formats, and that one exception case had a similar sized β_1 and a small number of students, so the most likely explanation is low statistical power. Figure 3-4 plots the β_1 values across sections: it is clear that the gap in self-efficacy for female and male students did not diminish in *flipped*, i.e., active engagement courses in either Physics 1 or Physics 2. For example, in Physics 1, the standardized beta coefficient for gender β_1 in the flipped course (~0.3) was similar to the β_1 in the traditional courses (between ~0.2-0.4). Similarly, in Physics 2, the standardized beta coefficient for gender β_1 in the flipped course (~0.3) is similar to the β_1 in the traditional courses (between ~0.3-0.4). Even though flipped instruction might improve learning and possibly self-efficacy overall, it does not decrease this gender effect.

Table 3-10 For different instructors and class types in Physics 1 and Physics 2 courses, number of students (N), gender standardized coefficient (β_1) and *p*-values are given.

Class Number	Section/Instructor/Class Type	N	β_1 values	p-values
Physics 1	Instructor 1 Traditional	140	0.19	0.018**
	Instructor 2 Traditional	26	0.18	0.95^{ns}
	Instructor 3 Traditional	25	0.42	0.017**
	Instructor 4 Flipped	128	0.30	0.001**
Physics 2	Instructor 4 Traditional	112	0.34	<0.001***
	Instructor 4 Flipped	286	0.29	<0.001***
	Instructor 5 Traditional	96	0.35	<0.001***
	Instructor 6 Traditional	101	0.29	0.003**

* Significant at the 0.05 probability level
** Significant at the 0.01 probability level
*** Significant at the 0.001probability level
ns: Non-significant result

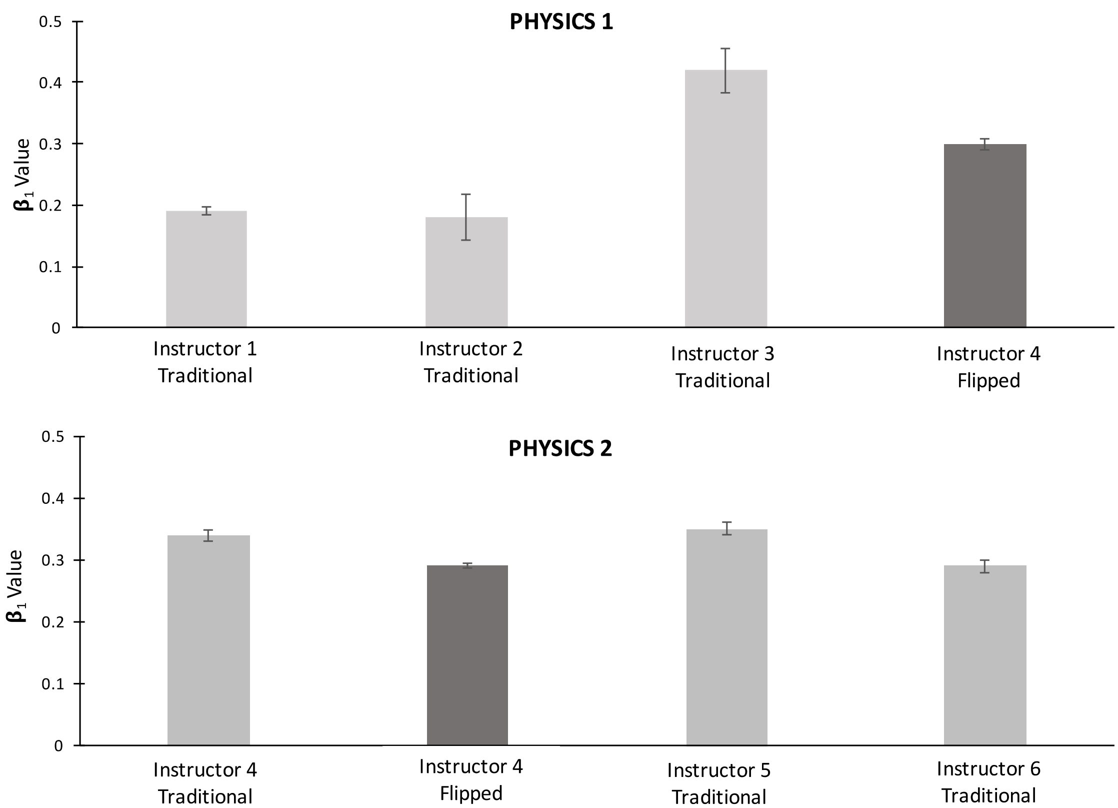

Figure 3-4 Gender standardized coefficients (β_1 values) for different instructors and class types in Physics 1 (top row) and Physics 2 (bottom row) courses. Darker and lighter bars represent flipped and traditional physics courses, respectively. Error bars represent standard error of the β_1 values.

3.4 Discussion and summary

Extensive prior research has shown that self-efficacy in science is an important driver of interest in science [40], STEM career choice [69], and persistence towards those career goals [48]. Prior research has also shown that female students have significantly lower self-efficacy than male students in STEM fields [23,53], and this gap in self-efficacy partly contributes to the underrepresentation of women in science and engineering fields [68,69]. However, few studies have focused on self-efficacy differences between female and male students controlling for performance, either to remove a natural confound or to understand how the self-efficacy gap might vary across performance levels. Therefore, we investigated self-efficacy differences of male and female students with equal performance levels in introductory physics courses, which are generally recognized as gateway courses for obtaining a physical science or engineering degree.

Our findings indicate that female students had significantly lower self-efficacy than male students throughout a two-semester introductory physics course sequence at every matched performance level. Further, the gap persists regardless of whether performance was measured through research-validated instruments or through the performance indicators provided to students (i.e., grades). Importantly, the gender self-efficacy gap grew after instruction in Physics 1 and Physics 2 courses, most notably within the highest achieving student group, in both traditional and flipped courses. The findings suggest that the self-efficacy of female students, and especially high achieving female students, is negatively impacted by their experiences in introductory physics courses.

Furthermore, the gender gap in students' post self-efficacy is found to largely stem from students' perceptions in the course and the contribution from students' actual performance is very small. There can be several possible reasons for students' incorrect judgements about their competence. Most saliently, societal stereotypes and cultural beliefs about gender and STEM achievement can bias students' self-assessment of their competence. For instance, female and male students' self-confidence is likely to be influenced by beliefs about the discipline and who can succeed in it. Further, as noted in the background section, women can be subject to implicit or explicit stereotype threat in physics courses [97,98]. Negative stereotypes about women in physics may cause women to have different perception of success than men when they initially enter in a male-dominated discipline in which contributions of "brilliant" men are over-emphasized. Female students may even assume that they have to make extra efforts to succeed in physics relative to male students and their achievement is not a reflection of how good they are in physics unlike the achievements of "successful" men. Likewise, women might undergo additional stress and struggle to demonstrate their skills to be valued equally as men in a classroom in which they are underrepresented. Also, as mentioned in the introduction, some prior research has found that teaching practices and interactions may treat female students differently than male students. For example, if female students are not called upon to answer questions or not given the same type of positive feedback as male students, this could have a negative effect on self-efficacy. In addition, students generally compare themselves to others who are similar where gender can be one determinant of being similar [24]. The thought of "there are not many people like me" can negatively influence women's self-efficacy and reinforce stereotypical beliefs about women's ability in physics. Similarly, men may portray higher confidence in their ability regardless of how they perform due to these biased perceptions which favor their gender. Correll found that men assess their math competence higher than women even though they perform similarly [68]. In this research, Correll also found that boys were more likely to pursue careers requiring math competence and skills at higher rates than girls, not because they were actually better at math but rather because they *thought* they were better. Describing her research on gender differences in math, Tobias notes that "when girls succeed at a math lesson or on a math quiz, they attribute their success to luck; boys attribute it to their own inner ability. When girls fail, they attribute their failure to a lack of ability; boys attribute theirs to a lack of effort. That's why even girls who do well in mathematics in school don't develop the kind of confidence males do." [99]. This type of dichotomy in how male and female students internalize their successes and failures may be partly responsible for the lower self-efficacy of women in physics courses. In summary, these types of environmental and sociocultural biases and gender-based beliefs about physics can have a large impact on students' self-beliefs in their competence than how they actually perform in the course. Moreover, it is possible that if female students had higher self-efficacy about physics, they would have less anxiety about learning physics and all of their cognitive resources while solving physics problems would be devoted to learning which would have the potential to boost their performance to a higher level than what we observed in this study.

3.4.1 Implications for the physical sciences and engineering

The observed effect is particularly alarming because of its large size. In a research sense, the effect can be considered large because the Cohen's *d* approached 1. In a more practical sense, the effect can be considered large because we found that female students with high scores on physics conceptual surveys (or who are receiving As) had similar self-efficacy as male students with only medium or low scores on physics conceptual surveys (or are receiving Bs and Cs). Whether this effect is framed as under-confidence among women, over-confidence among men, or some combination (perhaps the most likely interpretation), the practical outcome could be the same: worse outcomes for women in STEM.

In the introductory physics course sequence, the self-efficacy gaps may produce higher levels of anxiety during exams, negatively impacting exam performance [100]. In addition, self-efficacy problems have been shown to impact interest, and therefore high achieving female students may begin to lose interest in engineering and physical science due to their inaccurate assessment of their capability in physics. Self-efficacy is also related to self-regulated learning strategies, that is, higher self-efficacy is associated with better self-regulated learning [44]. This decreased interest and the lower self-efficacy may trigger female students to devote less time to homework or studying, or even to drop out of physics courses / decide to exit the STEM major track altogether.

In sum, inaccurate assessments of one's capability and/or performance can influence interests and career goals. Gender differences in self-efficacy, especially in fields which have been historically male-dominated, may inhibit progress toward increasing the diversity in these fields. We discovered alarming trends in female students' self-efficacy in introductory physics courses which are generally required for engineering and other STEM majors—in particular, female students with A grades often had similar physics self-efficacy as male students with C grades. These self-efficacy gaps may result in an "accumulation of disadvantage" for women in physical science and engineering domains. In other words, female students' lower self-efficacy in physics than male students may partly contribute to the underrepresentation of women (and even highly qualified women) in some STEM fields. Even minor gender differences in self-efficacy in early college experiences can add up to major inequalities later in STEM careers [101,102].

The current research examined only physics, but similar patterns are likely to occur in other courses for which there are negative gender stereotypes such as difficult mathematics, engineering, computer science, and other STEM courses. For example, research has already shown that male students have higher self-efficacy than female students of similar performance in mathematics [68]. Such broader gender gaps in STEM-related self-efficacies may result in differential educational experiences between male and female students who major in STEM-related fields and, ultimately, the underrepresentation of women in those fields. Thus, it is imperative to reflect on ways to broadly improve the environments of introductory STEM courses for which there are negative gender stereotypes in order to help female students reconcile their self-efficacy with their actual performance and capability.

3.4.2 Implications for instruction

What might be causing the differences in male and female students' self-efficacy, or alternatively, what might be done to reduce these differences? Since the effect grows with instruction, features of instruction seem particularly important to consider, both the instructional style of the instructor and the general pedagogy used in the class. Past research suggests that some instructors may have implicit (or even explicit) negative gender biases [75]. Although instructional style was not formally measured, the study included a wide range of instructors (tenure-stream and non-tenure stream; some having won teaching awards) and pedagogies (from very traditional lecture to including a number of active learning strategies as part of flipped instruction). Yet we found that the gender differences in self-efficacy were consistent across the instructors and class types, i.e., lecture-based and active-engagement courses. Thus, whatever the source, the recently-explored adjustments to pedagogy to increase learning appear not to be relevant to this gender gap in self-efficacy [103,104]. It may be that messages conveyed by students to each other [25] or the broader culture [55] may at least partly be the root cause and thus a different kind of counter-messaging is required.

Why might active learning not reduce the gender gaps in self-efficacy? While active learning may be beneficial for both male and female students in terms of performance outcomes, the nature of these in-class interactions may result in a decrease in female students' self-efficacy. Felder et al. [25] note that women usually play less active roles than men in cooperative learning groups in engineering and instead women report feeling that group work benefitted them because there were opportunities to have the material explained to them (i.e., reinforcing stereotypes of relative weakness). In addition, women also report feeling that their contributions in group work are undervalued and their contributions in active learning situations may also be ignored or discounted by other male students in the group [25]. Therefore, it is important for instructors to think carefully about how active learning is implemented and ways to help all students benefit from it. For example, instructors may need to frequently remind students that all group members' contributions are important and valuable. In addition, cooperative learning groups can be structured such that women outnumber men in any group containing women—in this way, women may not feel as intimidated by male members in the group. Investigations of the types of pedagogies and interventions that help women accurately assess their capability and performance in introductory physics courses are crucial in reducing the alarming self-efficacy gender gaps and their detrimental effects.

3.5 Chapter references

1. National Student Clearinghouse Research Center (2015). Science and Engineering Degree Attainment: 2004–2014 retrieved from https://nscresearchcenter.org/snapshotreport-degreeattainment15/

2. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016; *Women, Minorities, and Persons with Disabilities in Science and Engineering, Occupation* (2017). Retrieved from https://nsf.gov/statistics/2017/nsf17310/digest/occupation/women.cfm.

3. Degrees in engineering and engineering technologies conferred by postsecondary institutions, by level of degree and sex of student: Selected years, 1949-50 through 2013-14 (2015). Retrieved from https://nces.ed.gov/programs/digest/d15/tables/dt15_325.45.asp.

4. B.L. Yoder Engineering by the numbers Washington, DC: ASEE. 2011 Retrieved from https://www.asee.org/papers-and-publications/publications/college-profiles/15EngineeringbytheNumbersPart1.pdf

5. R. Henderson, G. Stewart, J. Stewart, L. Michaluk and A. Traxler, Exploring the gender gap in the conceptual survey of electricity and magnetism, Phys. Rev. Phys. Educ. Res. **13**, 020114 (2017).

6. A. Traxler, R. Henderson, J. Stewart, G. Stewart, A. Papak and R. Lindell, Gender fairness within the Force Concept Inventory, Phys. Rev. Phys. Educ. Res. **14**, 010103 (2018).

7. A. Traxler, X. Cid, J. Blue and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future, Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).

8. A. Traxler and E. Brewe, Equity investigation of attitudinal shifts in introductory physics, Phys. Rev. ST Phys. Educ. Res. **11**, 020132 (2015).

9. Z. Hazari, G. Potvin, R. Lock, F. Lung, G. Sonnert and P. Sadler, Factors that affect the physical science career interest of female students: Testing five common hypotheses**,** Phys. Rev. ST Phys. Educ. Res. **9**, 020115 (2013).

10. Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, J. Res. Sci. Teach. **47**, 978 (2010).

11. Z. Hazari, P. Sadler and G. Sonnert, The science identity of college students: Exploring the intersection of gender, race, and ethnicity, J. Coll. Sci. Teach. **42**, 82 (2013).

12. R. Ivie, S. White and R. Chu, Women's and men's career choices in astronomy and astrophysics, Phys. Rev. Phys. Educ. Res. **12**, 020109 (2016).

13. R. Ivie and K. Stowe, Women in Physics, 2000. AIP Report https://eric.ed.gov/?id=ED446559

14. M. Besterfield-Sacre, M. Moreno, L. Shuman and C. Atman, Gender and ethnicity differences in freshmen engineering student attitudes: A cross-institutional study, J. Eng. Educ. **90**, 477 (2001).

15. C. Seron. S. Silbey, S. Cech and B. Rubineau, Persistence is cultural: Professional socialization and the reproduction of sex segregation, Work and Occupations **43**, 178 (2016).

16. 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).

17. A. W. Astin, *What Matters In College: Four Critical Years Revisited*. San Francisco, CA: Jossey-Bass (1993).

18. N.M. Hewitt and E. Seymour, A Long, Discouraging Climb, ASEE Prism **1**, 24 (1992).

19. E. Seymour and N.M. Hewitt, *Talking About Leaving: Why Undergraduates Leave The Sciences*. Boulder, CO: Westview Press (1997).

20. S. Tobias, *They're Not Dumb, They're Different: Stalking the Second Tier*, Tuscon, AZ: Research Corporation (1990).

21. S. Cheryan, S. Ziegler, A. Montoya and L. Jian, Why are some STEM fields more gender balanced than others? Psychological Bulletin (2016). doi: http://dx.doi.org/10.1037/bul0000052

22. M. Wang and J. Degol, Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields, Dev. Rev. **33**, 304 (2013).

23. S. Cheryan and V. Plaut, Explaining underrepresentation: A theory of precluded interest, Sex Roles **63**, 475 (2010).

24. C. Hill, C. Corbett and A. Rose, Why so few? Women in science, technology, engineering, and mathematics (2010). Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf.http://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2010.tb01066.x/epdf

25. R. Felder, G. Felder, M. Mauney, C. Hamrin and E. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Engineering Education **84**, 151 (1995).

26. E. Lichtenberger and C. George-Jackson, Predicting high school students' interest in majoring in a STEM field: Insight into high school students' postsecondary plans, Journal of Career and Technical Education **28**, 19 (2013).

27. C. Nord et al., The nation's report card: America's high school graduates (NCES 2011-462). Washington, CD: U.S. Government Printing Office (2011). Retrieved from https://files.eric.ed.gov/fulltext/ED518324.pdf.

28. A. Strenta, R. Elliot, R. Adair, M. Matier and J. Scott, Choosing and leaving science in highly selective institutions, Res. High. Educ. **35**, 513 (1994).

29. E. Pascarella and P. Terenzini, *How College Affects Students: Findings and Insights From Twenty Years of Research*, San Francisco, CA: Jossey-Bass (1991).

30. A. Wigfield et al. Motivation to succeed. In Damon, W. and Eisenberg, N (Eds.) *Handbook of Child Psychology Vol. **3** Social Emotional and Personality Development*, New York: Wiley (2006).

31. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

32. R. Beichner et al., The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project Retrieved from http://www.per-central.org/per_reviews/volume1.cfm? (2007).

33. A. Diekman, E. Brown, J. Johnston and E. Clark, Seeking congruity between goals and roles: A new look at why women opt out of STEM careers, Psychological Science **21**, 1051 (2010).

34. J. Smith, C. Sansone and P. White, The stereotyped task engagement process: The role of interest and achievement motivation, J. Educ. Psy. **99**, 99 (2007).

35. S. Ahlqvist, B. London and L. Rosenthal, Unstable identity compatibility: How gender rejection sensitivity undermines the success of women in science, technology, engineering and mathematics fields, Psychological Science **24**, 1644 (2013).

36. C. Dweck, Is math a gift? Beliefs that put females at risk. In S. Ceci and W. Williams (Eds.), *Why Aren't More Women In Science? Top Researchers Debate The Evidence* (pp. 47-55). Washington, DC: APA Press (2007).

37. C. Dweck, *Mindsets and Math/Science Achievement*, New York: Carnegie Corporation of New York, Institute for Advanced Study, Commission on Mathematics and Science Education (2008).

38. I. Goodman, C. Cunningham and C. Lachapelle, The Women's Experience in College Engineering. Retrieved from https://files.eric.ed.gov/fulltext/ED507395.pdf (2002).

39. R., Lapan, A. Adams, S. Turner and J. Hinkelman, Seventh graders' vocational interest and efficacy expectation patterns, Journal of Career Development **26**, 215 (2000).

40. J. Eccles, Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices, Psychology of Women Quarterly **18**, 585 (1994).

41. A. Bandura, Self-efficacy. In R. J. Corsini (Ed.), Encyclopedia of psychology (2nd ed., Vol. **3**, pp. 368-369). New York: Wiley (1994).

42. A. Bandura, Self-efficacy: The exercise of control. New York: Freeman (1997).

43. F. Pajares, Gender differences in mathematics self-efficacy beliefs. In A.M. Gallagher and J.C. Kaufman (Eds.), Gender differences in mathematics: An integrative psychology approach (pp. 294-315). New York: Cambridge University Press (2005).

44. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemporary Educational Psychology **25**, 82 (2000).

45. D. Schunk and F. Pajares, The development of academic self-efficacy. In A. Wigfield and J. Eccles (Eds.), Development of Achievement Motivation: A Volume In The Educational Psychology Series (pp. 15-31). San Diego, CA: Academic Press (2002).

46. H. Watt, The role of motivation in gendered educational and occupational trajectories related to maths, Educational Research and Evaluation **12**, 305 (2006).

47. A. Zeldin, S. Britner and F. Pajares, 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**, 1036 (2008).

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

49. S. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, Journal of Research in Science Teaching **43**, 485 (2006); S.L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, Journal of Research in Science Teaching **45,** 955 (2008); A.L. Zeldin, S.L. Britner and F. Pajares, A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers, Journal of Research in Science Teaching **45,** 1036 (2008).

50. A. Bandura, Social cognitive theory of self-regulation, Organizational Behavior and Human Decision Processes **50**, 248 (1991).

51. P. Pintrich, A motivational science perspective on the role of student motivation in learning and teaching contexts, Journal of Educational Psychology **95**, 667 (2003).

52. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students, International Journal of Behavioral Development **14**, 153 (1991).

53. P. Huang and S. Brainard, Identifying determinants of academic self-confidence among science, math, engineering, and technology students, Journal of Women and Minorities in Science and Engineering **7**, 315 (2001).

54. 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).

55. S. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, J. Eng. Educ. **87**, 369 (1998).

56. L. Irani, Understanding gender and confidence in CS course culture. Paper presented at the SIGCSE 2004, Norfolk VA. Retrieved from http://ai2-s2-pdfs.s3.amazonaws.com/f5c2/cd1620808de342808613cd0cd4a67dbd8931.pdf (2004).

57. M. Rosson, J. Carroll and H. Sinha, Orientation of undergraduates toward careers in the computer and information sciences: Gender, self-efficacy and social support, ACM Transactions on Computing Education **11**, 1 (2011).

58. 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).

59. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1069 (2012).

60. 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).

61. 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).

62. 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).

63. R. Marra, K. Rodgers, D. Shen and B. Bogue, Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy, J. Eng. Educ. **98**, 27 (2009).

64. J. Raelin, M. Bailey, J. Hamann, L. Pendleton, R. Reisberg and D. Whitman, The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention, J. Eng. Educ. **103**, 599 (2014).

65. O. Eris et al., Outcomes of a longitudinal administration of the persistence in engineering survey, J. Eng. Educ. **99**, 371 (2010).

66. K. Meyers, S. Silliman, N. Gedde and M. Ohland, A comparison of engineering students' reflections on their first-year experiences, J. Eng. Educ. **99**, 169 (2010).

67. B. M. Jones, Pursuing diversity at state-supported residential STEM schools, NCSSSMST Journal **16**, 30 (2010).

68. S. Correll, Gender and the career choice process: The role of biased self-assessments, American Journal of Sociology **106**, 1691 (2001).

69. S. Correll, Constraints into preferences: Gender, status, and emerging career aspirations, American Sociological Review **69**, 93 (2004).

70. L. Vesterlund and M. Niederle, Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics **122**, 1067 (2007).

71. P. Rosser, The SAT Gender Gap: Identifying the Causes. Retrieved from https://eric.ed.gov/?id=ED311087 (1989).

72. R. Childs, Gender Bias and Fairness, (1990), retrieved from http://pareonline.net/getvn.asp?v=2andn=3

73. M. Spear, Sex bias in science teachers' ratings of work and pupil characteristics, Eur. J. Sci. Educ. **6**, 369 (1984).

74. D. Sadker, M. Sadker and K. Zittleman, *Still Failing at Fairness: How Gender Bias Cheats Girls and Boys in School*. New York: Scribner (1994).

75. C. Moss-Racusin, J. Dovidio, V. Brescoll, M. Graham and J. Handelsman, Science faculty's subtle gender biases favor male students, In Proc. of the National Academy of Sciences of the United States of America **41**, 16474 (2012). https://doi.org/10.1073/pnas.1211286109

76. V. Coletta, J. Phillips and J. Steinert, FCI normalized gain, scientific reasoning ability, thinking in physics, and gender effects, AIP Conf. Proc., pp. 23-26. doi: https://doi.org/10.1063/1.3679984 (2012).

77. L. McCullough, Gender, context, and physics assessment, Journal of International Women's Studies **5**, 20 (2004).

78. A. Wigfield, J. Eccles and P. Pintrich, Development between the ages of eleven and twenty-five, In D. Berliner, and R. Calfee (Eds.), The Handbook of Educational Psychology. New York: MacMillan Publishing (1996).

79. J. Blickenstaff, Women and science careers: Leaky pipeline or gender filter, Gender and Education **17**, 369 (2006).

80. J. Enfield, Looking at the impact of the flipped classroom model of instruction on Undergraduate Multimedia Students at CSUN, TechTrends **57**, 14 (2013).

81. L. P. Galway, K. K. Corbett, T. K. Takaro, K. Tairyan and E. Frank, A novel integration of online and flipped classroom instructional models in public health higher education, BMC Medical Education **14**, 1472 doi: 10.1186/1472-6920-14-181 (2014).

82. 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, Journal of College Science Teaching **35**, 20 (2005).

83. E. Brewe, Modeling theory applied: Modeling Instruction in introductory physics, *American Journal of Physics* **76**, 1155 (2008).

84. V. Sawtelle, E. Brewe and L. Kramer, "Positive impacts of Modeling Instruction on self-efficacy" in *Proceedings of the 2010 Physics Education Research Conference*, AIP Conference Proceedings Number 1289, edited by C. Singh, M. Sabella and S. Rebello (AIP, New York, 2010), p. 289.

85. G. Novak, E. Patterson, A. Gavrin and W. Christian, *Just-In-Time Teaching: Blending Active Learning with Web Technology*. Upper Saddle River, NJ: Prentice Hall (1999).

86. E. Mazur, *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall (1997).

87. Activation Lab Science Technology Engineering Art Mathematics Retrieved from http://www.activationlab.org/tools/. (2017); C. Dweck, *Mindset: The New Psychology of Success* (Ballentine, New York, 2006); C. Dweck, *Self-theories: Their Role in Motivation, Personality, and Developmen*t (Psychology Press, Philadelphia, 1999).

88. J. Schell and B. Lukoff, Peer instruction self-efficacy instrument [Developed at Harvard University]. Unpublished Instrument; E. Redish, J. Saul and R. Steinberg, Student expectations in introductory physics, American Journal of Physics **66**, 212 (1998).

89. S. Glynn, P. Brickman, N. Armstrong and G. Taasoobshirazi, Science motivation questionnaire II: Validation with science majors and nonscience majors, Journal of Research in Science Teaching **48**, 1159 (2011).

90. Z. Y. Kalender, E. Marshman, T. Nokes-Malach, C. Schunn and C. Singh, Motivational characteristics of underrepresented ethnic and racial minority students in introductory physics courses, in the *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.046.

91. T. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, in the *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

92. L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

93. J. Nunnally, *Psychometric Theory* 2nd Ed. New York: McGraw-Hill (1978). M. Chi, Thinking aloud, in *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes*, edited by M. W. Van Someren, Y. F. Barnard, and J. A. C. Sandberg (Academic Press, London, 1994).

94. D. Hestenes, M. Wells and G. Swackhamer, Force Concept Inventory, The Physics Teacher **30**, 141 (1992).

95. 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).

96. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences.* L. Erlbaum Associates (1988).

97. A. Maries, N. I. Karim and C. Singh, Impact of Evidence-based active engagement instruction on the gender gap in introductory physics, in the *Proceedings of the 2018 Physics Education Research Conference* (2018).

98. G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women's performance in physics, International Journal of Science Education **35**, 3050 (2013).

99. S. Tobias, Rethinking teaching math, science, *Education World* 2002; S. Tobias, *Banishing math anxiety* (Kendall Hunt 2012).

100. P. Pintrich and E. De Groot, Motivational and self-regulated learning components of classroom academic performance, J. Educ. Psych. **82**, 33 (1990).

101. L. Babcock and S. Laschever, *Women Don't Ask: Negotiation and the Gender Divide* (Bantam Books, New York 2007).

102. V. Valian, *Why So Slow? The Advancement of Women*. Cambridge, MA: MIT Press (1997).

103. N. Karim, A. Maries and C. Singh, Impact of evidence-based flipped or active-engagement non-flipped courses on student performance in introductory physics, Can. J. Phys. **96**, 411 doi: https://doi.org/10.1139/cjp-2017-0171 (2017).

104. N. Karim, A. Maries and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics?, Eur. J. Phys. **39**, 025701 (2018).

4.0 Enduring damage of women's lowered self-efficacy in physics learning

4.1 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. 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, like physics, have seen little progress. For instance, the percentage of bachelor and Ph.D. degrees in physics earned by women in the US is approximately 20% [1]. Even more asymmetric participation occurs for postdoc and academic leadership positions in physics [2].

Education researchers have considered several reasons to explain the gender gap in physics participation [3-12]. These reasons include biased learning tools [3,5], non-inclusive teaching methods and physics department climate [4], sociocultural beliefs in physics, and motivational factors [6]. Although there has been much interest in improving the pedagogy of physics teaching and reforming the content of the physics curriculum, there is less focus on investigating students' motivational factors and how they are related to male and female students' learning of physics. These factors can serve as a potential underlying explanation for why women do not pursue physics as often as men do. One of the central motivational factors in educational studies is students' self-efficacy, which refers to individuals' own beliefs about how well they expect to do in a particular subject or task [13]. Prior work in many areas of education has found that self-efficacy predicts students' retention and academic performance even after controlling for knowledge [6-16].

Therefore, in understanding gender disparity in physics, self-efficacy can be an important factor to examine.

This study examines the role of self-efficacy in explaining the gender gap in calculus-based physics courses. At the college level, there are typically fewer than 30% women in calculus-based physics classrooms, compared to algebra-based physics courses in which women are often in the numerical majority [9]. Therefore, it is especially important to understand the role of self-efficacy in explaining gender-based achievement gaps in calculus-based physics courses where women are under-represented. In the following sections, we present an overview of the literature on self-efficacy research in physics learning and its relation to prior knowledge, academic preparation, and achievement differences by gender.

4.2 Background

4.2.1 Gender differences in physics performance, prior physics experience, and academic preparation

Previous studies have documented large gender differences in physics performances across various institutions [6-24]. In college level calculus-based physics courses, women often score lower than men on exams [18] and standardized conceptual physics tests [19]. Interactive engagement teaching methods have been proposed to address the gender gaps [25]. While some prior work found a reduced gender gap in active-engagement courses [23], other studies reported that the performance gap remained [21] and even became larger in calculus-based physics courses despite the use of the interactive teaching methods [22].

Interestingly, the performance gap on standardized conceptual physics assessments has been found to exist on pre-tests (at the beginning of the course before instruction), which could explain part of the differences in post-test after instruction [19,20]. Therefore, 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 preparations [26,27]. In high school in the US, Advanced Placement (AP) physics courses are one of the learning opportunities that benefit students as they prepare for college physics. AP physics courses can help students develop a deeper understanding of physics principles and enhance their confidence and interest in pursuing physics-related careers. Female students enroll in AP physics at lower rates than their male counterparts. Moreover, the number of female students decreases in the more advanced AP physics, such as BC Physics [28]. The differential participation by gender also continues among those who choose to take the AP Physics exam that counts for college credit—taking the exam is optional and requires a fee [28].

Relatedly, developing robust mathematical skills can help students in college-level physics courses [27]. For example, the number of mathematics courses taken in high school is a strong independent predictor of students' college achievements in introductory science courses [27]. Likewise, research suggests that high school math grades and SAT math scores can predict college physics course success [29-31]. In one particular study, high school preparation in math was found to be the strongest predictor of students' physics grades in college [7]. Mathematics as a foundation to physics is particularly relevant because there have also been gender differences in math performance [32,33]. Despite female students' high math performance during elementary and middle school, male students score higher on high school math assessments [32,33]. This shift in math achievement during high school might be due to environmental factors such as lack of encouragement for girls in taking more advanced math classes or a belief that math is only for boys [34]. More importantly, gender performance gap in pre-college math can further impact women's performance and self-efficacy beliefs in college science courses [35].

4.2.2 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 [36]. Self-efficacy has been found to shape and be shaped by students' interests, as well as their effort and engagement in class [13]. In particular, self-efficacy can influence students' self-regulation processes, such as goal setting, time management skills and self-judgement [37]. Students with high self-efficacy become more task-centered [38], and they are more likely to exhibit advanced level learning strategies, such as self-monitoring and self-regulation [13,38]. 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 [37].

The role of self-efficacy becomes particularly salient when students tackle difficult problems. During problem solving, students with high self-efficacy interpret the struggle as an opportunity for developing their skills while those with low self-efficacy may see the challenge as a large hurdle and further evidence of their lack of competence in the subject [37]. When encountering challenging activities, students with low confidence become less interested, spend less effort and time, and eventually disengage from the class [39]. 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 have shown that self-efficacy can predict student performance in science courses when controlling for prior knowledge and academic skill differences [40-42]. Relatedly, non-physical science majors' self-efficacy belief was also shown to be a predictor of conceptual understanding and course achievement in physics [21].

4.2.3 Self-efficacy, gender, and performance in physics

Unfortunately, many studies have shown a prevailing gender gap in students' self-efficacy levels in science and math courses, and overall academic achievements. In other words, female students have consistently reported lower self-efficacy than male students in STEM courses [6,21,43-51]. 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 traditional) [45]. Another study identified a large self-efficacy gender gap for equally performing female and male students for all achievement groups (low, medium, high) [11,12]: women obtain A's in physics often had self-efficacy levels similar to men obtaining C's in physics.

One of the major factors that can influence students' self-confidence is student-environment interaction [52]. Highlighted in Bandura's influential self-efficacy theory, verbal persuasion (such as guiding or giving feedback) is key to boosting students' self-confidence [36]. Support from family, instructors, or academic advisors is especially important for women and other minority groups to maintain their interest in historically male-dominated domains (involving white-male students) and careers [36]. For instance, positive encouragement from academic course instructors can support women' self-confidence by motivating them to engage in challenging physics activities, whereas feelings of stress, anxiety and self-doubt due to lack of recognition and biases in the classroom can negatively influence women's beliefs about their competence.

Unfortunately, there can be subtle biases regarding students' competence held by science faculty, which mostly favor male students [53].

The unintentional biases held by faculty can signal who is expected to succeed in physics. These biases against women's skills in physics often stem from the overwhelmingly masculine culture of the physics community [54-56]. Pervasive masculine stereotypes continually reinforce the idea of how important gender is to succeed in physics. In particular, masculine norms in the physics community may act as a threat to students' "gender" identity, suggesting to women that their gender might not be an appropriate match for physics. Furthermore, masculine stereotypes are often associated with a "chilly" climate in physics departments and physics classrooms, producing an unwelcoming learning environment for women [55].

Further, there are not many female physicists who can act as role models for women (i.e., female instructors, teaching assistants, or mentors). Having female peer mentors early in college was found to increase women's retention and enhance their experiences in STEM majors [57]. However, it may be even more important for women to receive recognition and support from faculty and academic mentors.

Another common stereotypical belief associated with physics is the emphasis on brilliance as a key to success [58]. In a society where being a genius or gifted is associated with boys [59], girls avoid disciplines that are thought to require innate brilliance and talent [58]. Moreover, societal stereotypes are such that boys are perceived as being "smart" and "talented" in math-intensive science fields. These gender-based social beliefs about who can succeed in physics not only act as a barrier for women's participation in the field but also negatively affect their self-efficacy and alter their interest at early ages [54]. Leslie et al. found that gender stereotypes emerge at early ages at which 6-years-old boys and girls are more likely to relate to boys being "really

really smart" and girls also avoid activities that are believed to be for those who are "really really smart" [60]. These gender stereotypes about intellectual capability promote a fixed mindset, i.e., a view of intelligence as an innate and static attribute as opposed to a growth mindset, where people view intelligence as malleable and something that can be developed with effort [61].

One of the most well-researched consequences of gender-based beliefs about ability is stereotype threat. In this phenomenon, stigmatized groups such as women in physics have a fear of confirming stereotypical expectations about their gender performing poorly in physics. This fear creates further anxiety and can impact the marginalized group's performance, which becomes a self-fulfilling prophecy [62]. In other words, gender stereotypes cause women to doubt their competency and skills in learning physics. For example, stereotype threat has been shown to negatively impact women's performance in high school physics exams and college level physics course performance [63,64].

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 typically decide to stay or exit the major at the end of the freshman year. Therefore, affirmative first-year experiences in physics courses can play an important role in sustaining female students' interest and self-efficacy in STEM majors [65].

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. Although there is research indicating the gender differences in physics self-efficacy, relatively few prior studies have attempted to relate self-efficacy [6,21], gender, and performance to physics learning. Most of the previous studies have examined physics self-efficacy for physics courses taken by only biology majors or all majors combined (including STEM and non-STEM majors) [6,21,66] instead of investigating self-efficacy differences only for physical sciences and engineering STEM majors. Typically, algebra-based physics courses taken by health-related majors and some non-majors are majority female. Studies conducted even in gender-balanced physics classrooms have reported self-efficacy differences favoring male students [6,21,66], and we would expect these gender differences to become even more salient in calculus-based physics courses since women are in the minority and might be subject to an additional source of stereotype threat. Given that, understanding gender differences in students' attitudes and achievements specifically in calculus-based physics courses can shed some light on the issue of low participation rate of female students in the traditionally male-populated STEM fields.

Additionally, the past studies mostly examined the relations between the students' attitudes and performances in physics 1 courses [19,46]. However, there has not been much work investigating the correlations between student achievement and motivational factors across gender in the context of physics 2. One important rationale for studying physics 2 courses is that students mostly find the topics related to electricity and magnetism in this course more abstract and less intuitive (i.e., not being able to observe some of the concepts in their daily lives) as opposed to Newtonian physics, and therefore more dependent on resilience in the face of adversity. Furthermore, more students have prior high-school experiences in the physics 1 context as opposed to physics 2, which can influence students' initial self-efficacy in the physics 2 courses in a different way than physics 1.

Also, it is not clear from prior research to what extent gender differences in self-efficacy predict gender-based gaps in physics achievements after controlling for various prior knowledge differences. Few studies have investigated theoretical models that predict gender-performance relations controlling for SAT math and general student attitudes [19,67]. Kost-Smith et al. partially explained the gender effect in course grades, but the gap was not fully explained in the final model [67]. However, they did not directly control for students' self-efficacy in particular and instead used a predictive model that combined multiple attitudinal factors. Since the size of gender differences in attitudes is not identical across physics attitudes [9], combining attitudes has the potential to reduce a model's explanatory power. On the other hand, we specifically focus on self-efficacy, which consistently shows large differences by gender.

Moreover, it is not clear from prior research how to integrate prior differences in academic preparation. It is highly likely that prior physics experiences (e.g., AP physics performance) and knowledge (e.g., as measured by conceptual tests) as well as mathematics performance (e.g., as measured by SAT math scores) are correlated with gender, physics self-efficacy, and physics 2 performance. However, it is unclear how they should be integrated into one overall model. For example, are gender differences in physics self-efficacy the result of AP physics performance and SAT math differences or a separate effect based, e.g., on cultural and social stereotypes?

Here, our primary goal is to explore the mediational mechanism of self-efficacy in explaining gender differences in physics 2 learning outcomes, while also integrating academic skills (SAT math and AP physics test scores) and initial physics knowledge into the path model. We hypothesize that gender differences in learning outcomes (post standardized conceptual test scores or course grades) will be mediated with prior knowledge and self-efficacy. Moreover, we also explore the contributions of SAT math and prior knowledge in a standardized conceptual test as additional possible mediators of gender differences in learning outcomes (see Figure 4-1).

Therefore, our first research question is: *To what extent can gender differences in students' physics learning outcomes be explained by differences in physics self-efficacy at the beginning of the course?* Here we contrast the relative roles of self-efficacy, prior knowledge, and SAT math in explaining gender gaps in learning outcomes.

Another important related issue involves the sources of gender differences in physics self-efficacy. Previous work has documented the various ways in which men and women have different exposure to physics within both in and out-of-school learning experiences [68], as well as differential preparation in mathematics. These pre-college differences sometimes result in an initial physics knowledge gap and overall mathematics performance gap between men and women when they enter college. These experience differences could also underlie the self-efficacy differences. Therefore, we also posit a second research question: *To what extent are gender differences in physics self-efficacy based on prior physics knowledge differences and measures of pre-college academic scores?* Here we consider AP Physics (a common relevant high school experience with college-level physics) and SAT math (a common academic skill measure that strongly influences acceptance in selective STEM programs). While those factors are plausible drivers of self-efficacy beliefs, the connections to gender in this population are unclear. In particular, given selective participation of female and male students in physical science and engineering majors, it is not clear in advance whether there are gender differences in AP Physics or SAT math among this population.

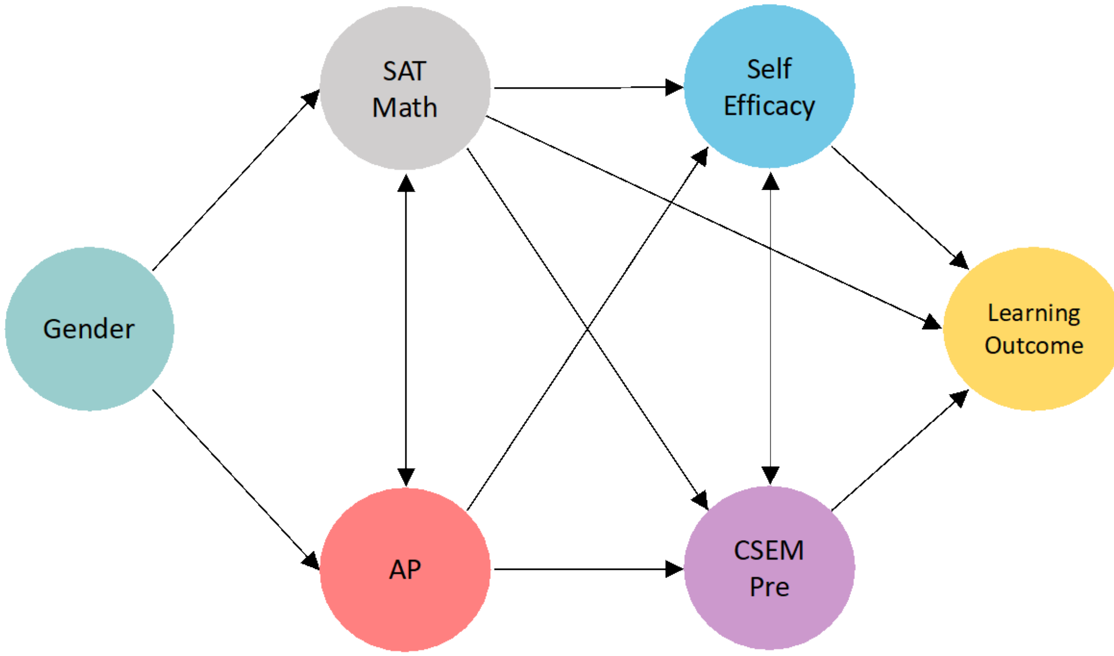

Figure 4-1 Conceptual framework connecting gender to learning outcomes (conceptual post-test and course grade) via key experience (Advanced Placement), attitude (Self-efficacy), and skill (SAT Math) variables. Single arrows correspond to linear regression relation and double arrows correspond to correlational relation.

4.3 Methodology

Data were collected from calculus-based physics courses over the course of two consecutive years. Our focus is introductory level Physics 2 courses that encompass topics in introductory level electricity and magnetism, very challenging topics even for physical science and engineering students because they have had relatively little exposure to this specific content in high school. We examine two different measures of learning outcomes: a research-based standardized conceptual test and course grades.

4.3.1 Participants' demographic information and class context

Participants were 642 students in calculus-based physics courses who intended to major in engineering or physical sciences. The demographic data (i.e., gender, ethnicity or age) were obtained from the university data warehouse that also kept extensive records about students' pre-college test scores (AP tests, SAT, etc.) and university grades. When 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 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. Students were predominantly White (78%), with the remaining students coming from a number of other ethnic/racial backgrounds: Asian (12%), African American (4%), Multi-racial (3%), Hispanic (2%) and Others (1%). Also, 90% of the students in this course were freshman with a mean age of 19.

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 course topics included electrostatics, magnetostatics, resistance, capacitance, inductance and simple electric circuits, Faraday's law of electromagnetic induction, Ampere-Maxwell's law, Maxwell's equations, electromagnetic waves, and wave optics. Regarding the class instructional format, 46% of students were in "flipped" courses (i.e., in which more active engagement tools were used) and 54% of students were in traditional lecture courses. In both traditional and flipped courses, there were weekly recitations given by graduate teaching assistants (TA). Attendance at recitations was mandatory, and students were given quizzes each week contributing to their final grade.

4.3.2 Measures

Physics Self-efficacy. We have previously developed a self-efficacy survey that was built from prior survey instruments [69-72]. Our instrument was iteratively refined and validated with exploratory factor analysis (EFA) and individual student interviews [9-12]. The self-efficacy survey included 6 items (Cronbach's alpha > 0.7) [73]. Self-efficacy questions assessed 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. Table 4-1 presents the self-efficacy items and response options for various questions. Varying response options were selected to anchor responses in more objective ways and encourage respondents to slow down while responding to read each item (see Table 4-1). Students were given a score from 1 (low) to 4 (high) for each response, with higher scores indicating higher levels of self-efficacy. For example, a student who answered "All of the time" to the first question, "All areas" to the second questions and "no" to the other four self-efficacy question would have an average self-efficacy score of (4+4+2+3—because this item is reverse coded—+2+2)/(6 total questions) = 2.83 which is between positive and neutral self-efficacy. Item Response Theory analyses verified roughly equivalent distance between response items and these averages were highly correlated with IRT factor scores, further justifying the use of simple averages.

Table 4-1 The physics self-efficacy survey with response options. One item in the survey is reverse coded and indicated as (R).

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

Conceptual Test. The Conceptual Survey of Electricity and Magnetism (CSEM) [74] 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 CSEM has been extensively validated as a measure of conceptual understanding of core physics phenomena and principles within the course topic areas of electricity and magnetism [74], and it has also been successfully used for comparing different teaching methods on a standardized basis [11,67]. The test consists of 32 multiple-choice questions. The test was given at the beginning (pre) and end (post) of the course. We calculated the proportion correct in pre and posttests. Typical mean scores on the CSEM for calculus-based physics students (out of 1) was approximately 0.32 in pre-test and 0.47 in post-test [11,74]. As it is appropriate for scales based upon dichotomous items (e.g., correct/incorrect), we use Armor's θ values to report reliability [75]. The θ values were 0.76 for pre-test and 0.84 for post-test, indicating good reliability [75].

Course grades. Students' course grades were also used as a measure of their learning outcome. The final course grade was largely determined by students' midterm and final exam scores. Weekly given homework, students' class participation, concept quizzes, attendance and recitation quizzes also contributed to the course grade. The final course grades were obtained from the data obtained from the university records. The conversion between the letter grade and corresponding grade point is given in Table 4-2. 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. Although a students' course grade is fine-grained thereby noisier measure influenced by attendance, TA, and peer support of homework completion, and uneven test quality, this measure is better aligned to exact content covered in each course and also represents an important outcome of learning for students (e.g., whether they must repeat the course).

Table 4-2 Letter grades and corresponding grade points.

	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A/A+
Grade Point	0	0.75	1.00	1.25	1.75	2.00	2.25	2.75	3.00	3.25	3.75	4
Definitions	Failure					Minimum level to graduate			Adequate level to graduate			Superior attainment

Pre-college test scores. The university provided a wide range of scores that are used to determine admission to the university, including high school GPA, standardized assessment scores for mathematical and verbal ability (SAT), and standardized assessment scores for advanced coursework. In prior research, we established that two scores were the primary predictor of student performance in the introductory Physics courses. The first is the Scholastic Assessment Test (SAT) Math scores, which ranged from 400–800. The SAT is designed to predict first year university performance. Prior research suggests that students may overgeneralize the implications of their performance on the SAT, believing that lower Math SAT scores imply lower ability for physical sciences [31].

The second measure is the standardized final exam score obtained in an Advanced Placement Physics course. Many students in physical science and engineering majors already took the equivalent of Physics 1 in their high school via the Advanced Placement program. However, there are two versions offered: algebra-based (AP Physics 1) and calculus-based (AP Physics AB). Based on a nationally administered and scored final exam, students receive a score of 1 to 5. Many universities treat a score of 4 or 5 (and sometimes also a 3) as the equivalent of passing their own university course. To create one composite measure across the two possible courses, we performed an analysis to test if there was any difference between the two test takers' subsequent performance in the college level physics 1 courses. The analysis was conducted separately for algebra-based and calculus-based AP physics takers. We calculated the mean grade of physics 1 for each AP score group (1 to 5) and a sixth group for those who had not taken the course. The physics 1 achievement patterns among the algebra-based and calculus-based AP physics test takers showed similarities across all AP test score groups indicating that the course grades in physics 1 had similar trends. Therefore, we combined the two groups of AP test scores (algebra-based and calculus-based; see Appendix Figure 4-4 and Figure 4-5). Relatedly, we also analyzed the physics 1 grades of students who did not have an AP test score (i.e., did not take the course or took the course but did not take the exam). The results indicated that non-AP test takers performed similarly in physics 1 to those who had an AP test score of 1. Therefore, we gave those students a score of 1 on this measure.

4.3.3 Procedures

Motivational and conceptual tests were given during recitation. Both surveys were administered by the responsible recitation TAs at the beginning and end of the Physics 2 courses. The self-efficacy survey was completed by most students in a couple of minutes (embedded in a larger motivational survey taking between 10-15 minutes), and the students worked through the conceptual physics assessments in the remaining class time (approximately 35-40 minutes).

Instructors were encouraged to give a small amount of course credit to students for completing the surveys. The instructor or teaching assistant responsible for giving the motivation and conceptual physics surveys was given the following script to announce before administering the surveys to the students to encourage students to take the assessments seriously: “We are surveying you on your understanding and beliefs about physics in order to improve the class. Your responses will not be evaluated for grades except to make sure the responses were done seriously, rather than randomly.”

4.3.4 Analysis

An initial examination compared female and male students' scores in predictors and outcomes for statistical significance using t-tests and for effect sizes using Cohen's *d* [76]. 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) as a statistical tool by using R (lavaan package) with a maximum likelihood estimation method [77]. SEM enables calculation of interrelated dependence between variables within a single analysis, which has greater statistical power and better controls for indirect correlations through third variables. SEM involves several commonly used fit parameters to test the goodness of the fit: Comparative Fit Index (CFI), which compares the fit of the proposed model to the null model; Tucker-Lewis Index (TLI), which is similar to CFI but takes into account a more complex model – TLI is more strict than CFI; Root Mean Square Error of Approximation (RMSEA), which refers to residuals and measures how closely the model fit to the data; and Standardized Root Mean Square Residual (SRMS), which is the standardized difference between the observed correlation and the predicted correlation. There are commonly used thresholds for deciding whether the fit is acceptable or not: CFI and TLI > 0.90; SRMR and RMSEA < 0.08. Here, we tested the proposed theoretical model and examined the resulting structural paths between constructs. In creating a final acceptable 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.

Finally, 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 learning outcome. If the predictor had more than one path between gender and learning outcome, we summed each path's contribution.

4.4 Results

4.4.1 Correlations

Zero-order pair-wise Pearson correlations are given in Table 4-3. Pearson's r values signify the strength of relationship between the variables, uncontrolled for other correlated variables. Focusing on the correlational relations among the predictors (self-efficacy, SAT math, AP physics, the CSEM pre), there were medium-level correlations ranging from 0.3 to 0.4, showing that the predictors are not so correlated as to be impossible to separate in the regression analyses, but also sufficiently inter-correlated that simple Pearson correlations with outcomes can be artificially higher than the true direct relationships. The strongest correlation was between the students' initial self-efficacy and CSEM pre-test scores, which assessed prior knowledge of physics 2 topics in the curriculum. 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 two rows of Table 4-3 present the correlation values between the learning outcomes (the CSEM post and course grade) and the main predictors. The CSEM pre-test was most closely correlated with students' CSEM post-test results and self-efficacy, with smaller but potentially important roles of SAT math and Physics AP. Course grades were roughly equally correlated with CSEM pre-test, Math SAT and self-efficacy.

The correlation between the two outcomes variables was sizeable but far from identical, supporting the need to separately analyze the relationship of the predictors to the two outcomes. Further, the pattern of simple correlations with the predictor variables was also different, further suggesting that separate analyses are warranted.

Table 4-3 Zero-order Pearson inter-correlations within the predictors, and between the outcomes and the predictors with given *p*-value statistics.

		AP Physics Test Score	SAT Math	CSEM Pre	Self-efficacy	CSEM Post
Predictors	AP Physics Test Score	–				
	SAT Math	0.33 < 0.001	–			
	CSEM Pre	0.40 < 0.001	0.36 < 0.001	–		
	Self-efficacy	0.34 < 0.001	0.28 < 0.001	0.43 < 0.001	–	
Outcomes	CSEM Post	0.30 < 0.001	0.31 < 0.001	0.50 < 0.001	0.42 < 0.001	–
	Course Grade	0.20 < 0.001	0.34 < 0.001	0.37 < 0.001	0.30 < 0.001	0.45 < 0.001

4.4.2 Gender differences in predictors and outcomes

Statistically significant gender differences in favor of male students were found on all of the variables (see Table 4-4). A very large gender gap occurred in students' initial self-efficacy reports [76]. While men reported approximately a mean score of 3 in self-efficacy beliefs which corresponded to a positive confidence level, women more typically reported a neutral level of confidence (~ 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, with medium differences in CSEM (pre and post), 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 / knowledge.

It is interesting that the gender gap was smaller in students' course grades than in CSEM performance. Students' course grades were composed of multiple formative assessments (i.e., weekly homework assignments, short quizzes or short lab reports) in addition to midterm and final exams that can be considered high-stakes assessment since they counted for a significant portion (more than half in all courses) of the course grade. The variation in gender difference may reflect a differential likelihood to complete assignments, quizzes, and lab reports or a reduced effect of stereotype threat in low-stakes assessments.

Table 4-4 Mean predictor and outcome values by gender, along with statistical significance (t-test *p*-value) and effect sizes (Cohen's *d*) for the gender contrast. Score ranges for each variable are also shown.

Predictors and Outcomes	Mean		t-test *p*	Cohen's *d*
	Female (N=174)	Male (N= 368)		
AP Physics Test (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
CSEM Post (0-1)	0.48	0.56	< 0.001	0.44
Course Grade (0-4)	2.58	2.72	< 0.05	0.20

4.4.3 SEM path model

To understand whether the gender differences in students' learning outcomes in physics were mediated by differences in students' initial self-efficacy, prior knowledge in physics, and pre-college academic skills, path models were applied to each of the learning outcomes.

4.4.3.1 Using CSEM as a learning outcome

After iterations to remove non-significant links, the final mediation model produced acceptable fit parameters: CFI = 0.99 (> 0.95), TLI = .98 (> 0.95), RMSEA = 0.04 (< 0.08), and SRMR = 0.013 (< 0.08). In this model, SAT math, self-efficacy and CSEM pre-test had direct effect on students' CSEM post scores (see Figure 4-2), where there were no direct connections with gender or AP Physics. Students' initial CSEM scores had the strongest effect ($\beta = 0.35^{***}$) on the conceptual test results. This result is not surprising especially considering the earlier results that showed medium level correlation between pre and post CSEM scores. Self-efficacy was the second strongest variable that had a direct effect on the CSEM post-test. In particular, self-efficacy remained a strong and significant predictor of learning outcome even after controlling for pre-college academic skills and prior knowledge differences in physics.

More interestingly, we find that the direct association with gender became non-significant for AP physics test scores, CSEM pre- and post-test. The only direct connections to gender involved a large relationship with self-efficacy ($\beta = 0.30^{***}$) and small relationship with SAT math ($\beta = 0.14^{**}$). This finding suggests a substantial and powerful impact of students' self-efficacy on learning outcomes even after adjusting for prior knowledge differences. More importantly, the gender gap in self-efficacy strongly mediated the gendered differences in pre- and post- physics test achievements. We note that self-efficacy differences were measured at the beginning of the course, which can be one indication of initial stereotype threat in physics courses that may negatively impact female students' performance.

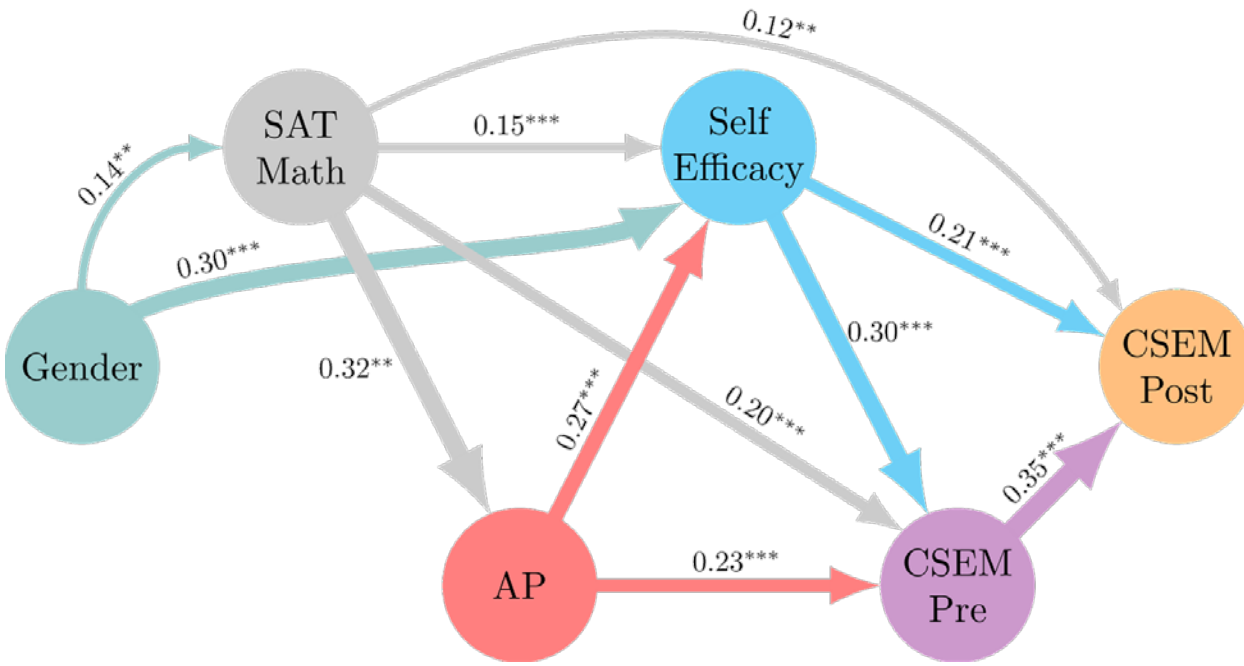

Figure 4-2 Results of the structural equation modeling between gender and standardized post-test (the CSEM post-test) through self-efficacy, SAT math, AP physics test score and the CSEM pre-test. The line thicknesses correspond to the magnitude of β values. All p-values are indicated by *** for $p < 0.001$ and ** for $p < 0.01$. The arrow below the diagram indicates the direction of regressions.

4.4.3.2 Using course grade as a learning outcome

For course grade, a similar model proved to fit the data well, and in fact provided an even stronger fit: CFI = 0.99, TLI = 0.99, RMSEA = 0.02, and SRMR = 0.01. As with CSEM, direct predictors of course grade were CSEM pre-test score, SAT math, and self-efficacy variables. CSEM pre-test again had the strongest connection ($\beta = 0.27^{***}$), but this time SAT math ($\beta = 0.13^{**}$) and self-efficacy ($\beta = 0.12^{**}$) were of similar size (see Figure 4-3), reflecting a smaller effect of self-efficacy. The weaker connection of grade to self-efficacy is consistent with the smaller gender difference in grade than in CSEM post scores.

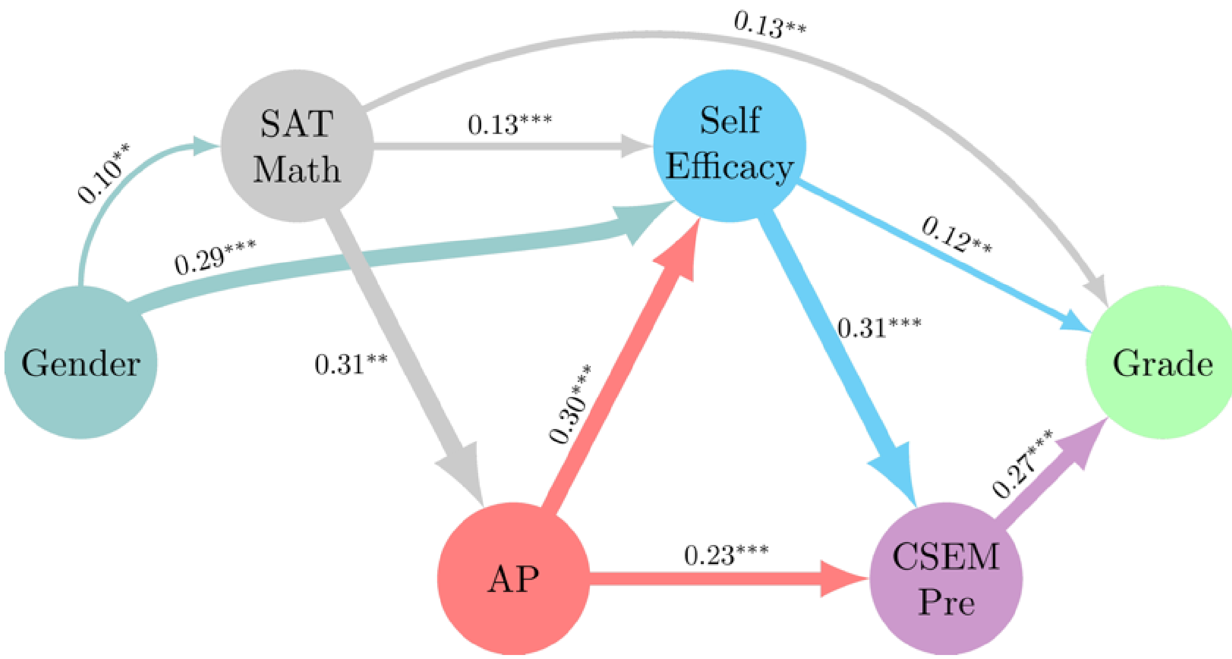

Figure 4-3 Results of the structural equation modeling between gender and course grade through self-efficacy, SAT math, AP physics test score and the CSEM pre. The line thicknesses correspond to the magnitude of β values. All p-values are indicated by * for $p < 0.001$ and ** for $p < 0.01$. The arrow below the diagram indicates the direction of regressions.**

4.4.4 Total mediated effects

Since gender is not directly connected to either CSEM-post or grade in the final path models, it is possible to examine the relative contribution of gender to outcomes via different mediators (i.e., calculate the total mediated effects). For instance, one of the path that SAT Math had was the connection between gender and physics course grade, therefore we multiplied the coefficients from gender to SAT Math and SAT Math to course grade (i.e. 0.10 x 0.13 = 0.013). Similarly, one of the self-efficacy paths between gender and grade flowed through SAT math so the calculation was: (gender → Sat Math) x (SAT Math → self-efficacy) x (self-efficacy → grade), i.e. 0.10 x 0.13 x 0.12 = 0.002. The other self-efficacy indirect paths were also calculated with a similar procedure and then added to calculate the total effect.

Total mediated effect calculations were conducted separately for both outcome variables (CSEM post and course grade) and shown in Table 4-5 along with each predictor's direct relationship to the outcome. All of the total mediated effects were found to be significant. Self-efficacy is by far the largest mediator of the gender effect for CSEM, with half as large contribution via CSEM pre and one quarter the contribution from SAT Math. The gender effect on grades is smaller, so it is not surprising that the mediated effects are all smaller for course grades. Again, self-efficacy was the largest effect, but this time similar in size to CSEM-pre. Overall, gender had the highest mediated effect through self-efficacy for both of the learning outcomes, followed by CSEM pre-test and SAT math. In other words, the gender gap in students' final course performances was predominantly mediated by the large gender differences in self-efficacy.

Table 4-5 Direct contribution of predictor to each outcome and the size of the total mediated contribution of the gender effect to each outcome via each predictor.

	CSEM Post		Course grade	
Predictors	Direct Predictor to Outcome	Gender to Outcome via Predictor	Direct Predictor to Outcome	Gender to Outcome via Predictor
CSEM pre	0.35***	0.04***	0.27***	0.03***
Self-efficacy	0.21***	0.08***	0.12**	0.04**
SAT Math	0.12**	0.02**	0.13**	0.01*

* Significant at the 0.05 probability level (*p*-value)
** Significant at the 0.01 probability level (*p*-value)
*** Significant at the 0.001 probability level (*p*-value)

4.5 Discussion

4.5.1 Summary of results

In Physics 2 courses which are required for most students, there were small gender differences in course grade and medium differences in conceptual test performance (CSEM effect size d ~ 0.40). One possible explanation for the smaller effect on course grades is that students' overall grades are composed of multiple other assessments, such as class participation, homework, quizzes or reports from short lab experiments in addition to high-stakes exams like midterm exams and the final exam. Prior research suggests that women perform at higher levels in low-stakes assessments and in activities that include more interactive learning [78].

We modeled the relationship between gender of students and learning outcomes 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 learning outcome differences because there were statistically significant gender differences for all of the variables (although of varying effect sizes), with male students scoring higher. Consistent with prior work [31-33], SAT math showed a medium level gender gap in the favor of male students (*d* ~ 0.30) but the gender performance gap was much smaller in AP physics test scores (*d* ~ 0.20).

The most striking result in our initial gender difference analysis was the large gender gap in students' self-efficacy at the beginning of the course (*d* ~ 0.8). These self-efficacy belief differences across gender are well beyond differences in their actual physics performances, as we previously found [11,12]. For example, we find 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 [11,12].

Prior research has shown a relationship between self-efficacy and students' achievements [6,9-13,21,37-46] with self-efficacy predicting students' learning after controlling for prior knowledge [11,46]. 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 conceptual or grade outcome measures. Further, the analysis of indirect effects revealed that the gendered patterns in conceptual test performance and course grades were mainly associated with students' self-efficacy, with a modest contribution of prior physics preparation and a very small contribution from mathematics skill (see Table 4-5). 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, prior mathematics and physics learning appears to play a small direct role in shaping later physics learning outcomes but plays an indirect role in shaping physics learning outcomes via undermining/supporting student self-efficacy, which then itself influences learning.

4.5.2 Implications

Research suggests that self-efficacy is related to students' learning/performance even after controlling for their prior academic performance differences [11,13,46]. There are several mechanisms that explain the strong impact of self-efficacy on students' self-motivation, academic achievement, goal orientation, and academic outcome expectations [13-16,79]. Students with positive self-efficacy can engage in more challenging tasks and they are more likely to persist when they face failure in such activities. Furthermore, students with higher self-efficacy may use constructive study strategies more effectively and engage more in classroom or with the instructor, which can increase their interest, enjoyment, sense of belonging and persistence. Students' perceptions about their physics self-efficacy can be related to students' past and current learning experiences, and student-environment interactions.

In addition to being driven by prior preparation differences, students' self-efficacy is related to their interactions with peers and classroom experiences [65]. In a male-dominated classroom environment such as in calculus-based introductory physics, a woman might experience a lower level of sense of belonging and higher anxiety with low confidence [80]. In addition, non-supportive instructional pedagogies, lack of recognition and classroom interactions can further decrease women's self-efficacy in physics. With that in mind, the instructor's focus on inclusivity, and approaches to recognizing students in poorly gender-balanced classrooms become even more vital in supporting women's confidence and promoting engagement in the classroom. Since women have sometimes been found to feel more comfortable working in interactive learning environments, instructors' implementation of explicitly inclusive active-engagement strategies might help women feel more confident and competent in physics. These inclusive strategies that provide a supportive environment in which women feel recognized and valued might also bolster women's interest in taking more physics-related courses.

We emphasize that the gender gap in physics course performance can increase in active engagement classrooms [22], so the new reforms towards active-engagement courses in physics will not generally be sufficient on their own to address performance gender gaps. In particular, attention to other factors, such as inclusivity, the extent to which women feel valued and recognized and details of the support for classroom discussions will matter in order to benefit all students equally. For example, during classroom activities, instructors will likely need to make sure that all students' opinions are valued and respected by all of the group members. In group activities during labs or lectures, female students typically have tasks that require a low level of cognitive engagement with the subject matter, such as notetaking or simply reporting the work [80]. Male students might dominate the conversations in these group discussions, which can cause female students' self-confidence to drop even more. Therefore, in such active-engagement activities, instructors need to assign each student a role and later rotate the student's role and ensure that all students have a sense of belonging and contribute to the task equally. These types of inclusivity considerations in which all students feel recognized and valued and have a more equal opportunity to learn and feel that they are equally contributing to group work, are likely to improve outcomes for all students.

Another common social and cultural aspect within physics needs to be considered seriously in identifying possible sources of self-efficacy differences across gender. The domain itself, at all levels, is mainly populated by male faculty and students [81], which creates a gender-based biases about who can succeed in physics. While this stereotypical portrayal of a "male" scientists can boost students that fit in to this biased image, it stands as a major barrier for others who do not identify themselves as such. The over-representation of men in the discipline of physics not only generates a lack of sense of belonging and a stereotype threat against women's performance but also impacts the way in which women form identities in physics-related fields. For instance, male-dominated culture and practice of physics can signal to women that their gender is not appropriate to be in the field and force female students to negotiate between their gender and physics identities [82-85]. For this reason, there needs to be a radical change within the culture of the physics community to eliminate "gender-related ability" biases so that it can become more inclusive for all students.

One primary cause of these biased beliefs is the field-specific intelligence attributions. As Leslie et al. found, women recede away from the domains that are thought to require innate ability and brilliance for success in the field [58]. Physics is one of the exemplar fields that illustrates the negative correlation between the number of women and high expectations for brilliance [58]. These biases provoke fixed intelligence mindset attributions regarding how success is achieved only via innate ability. Individuals discerning intelligence as a fixed characteristic then perceive struggles as a threat to their ability and failure as an indication of a lack of ability [61]. By contrast, fostering a growth mindset view encourages students to view struggle as a stepping stone that enhances learning, enabling students to become more enthusiastic about spending effort to develop their skills in physics. There are several classroom interventions designed to create better student engagement with growth mindset [86-90]. Some of these interventions have focused particularly on minority groups as they aim to normalize students' struggles in academic life and increase their sense of belonging [86].

We also note that failure to support women's self-efficacy especially during their first-year college experiences will not only have meaningful short-term results, but is likely to lead to long-term effects, such as gendered-patterns of retention in STEM domains. For instance, initial low self-efficacy of women can increase their anxiety in the exams [80] and cause them to perform worse than they actually otherwise would [91]. Since negative self-efficacy can impact students' interest [80], women's desire to engage in and dedicate time for physics learning can also diminish. Similarly, female students with low self-efficacy and higher anxiety are more likely to relate their poor performance to the lack of ability [80]. In other words, women's low self-efficacy can prompt them to become more fixed in their mindset, and to internalize the reasons of failure to lack of their talent in the field. These women can prematurely withdraw from the course or decide to leave the STEM field due to their biased self-judgements about their competence. Felder et al. found that among women and men who failed the course, men are more likely to repeat the course and stay in the engineering track while women switch out of the engineering degree [80].

Women's biased self-efficacy in physics can further have a negative impact on their choices of academic career. Some engineering fields are mostly male dominated and contain more physics-focused topics throughout their curricula, while other engineering degrees appear to be more gender-balanced and have less focus on physics materials [80]. Due to first-year college experiences in physics 2 courses, women might switch out from physics-intensive engineering majors, such as mechanical or electrical engineering despite having initial interest in these majors. Equally importantly, due to stereotypes, women's career choices might rely on fixed mindset and the ability-related negative beliefs such as beliefs about women having low ability in physics. Therefore, supporting women's achievement also necessitates promoting and supporting positive recognition and endorsement for their competence from mentors, academic advisors and course instructors as well as their families [82]. In particular, academic encouragement, support and recognition have the potential to enhance their self-efficacy and interest, and help them develop positive identities in the physics-related fields [92,93].

4.6 Limitations and future directions

This study investigated gender differences in a calculus-based introductory physics 2 course. The gender data obtained by university records included binary response options, however we recognize that gender is a sociocultural and multidimensional construct. In future studies, we hope to incorporate gender measurement with multiple options, which can allow us to measure masculinity and femininity in more nuanced ways.

In this paper, SEM was chosen as the data-analytic strategy due to its advantages, such as the availability of multiple fit measures to identify strong models, and more effective methods for controlling for random-error and latent variable variances [94]. However, there are also certain limitations in using this analysis technique. For instance, SEM does not handle non-linear relations and is fundamentally based on correlations and therefore is an indirect measure of causality [94].

In the theoretical model, we did not control for students' physics 1 grades, which could also explain the self-efficacy gender gap at the beginning of physics 2. However, we have shown in our previous work that 87% and 95% of students' self-efficacy gender gap comes from biases rather than their actual performance in physics 1 and physics 2, respectively [11]. Moreover, since conceptual test scores and students' grades were correlated with each other, models including both variables produce problematic model fits.

Also, we have not tested the extent to which self-efficacy gender gaps moderate the learning outcome differences across different instructors using different teaching methods. However, in some previous work we have shown that self-efficacy gender gap exists in calculus-based physics 2 courses for different instructors regardless of the teaching methods (i.e., whether they used active engagement versus traditional teaching methods) [11].

Moreover, our study focuses only on the calculus-based physics courses. In the future, we intend to do similar analysis of algebra-based physics courses. Apart from the differences in curriculum content and math-intensity, there are differences in student demographics and classroom structure. First, calculus-based physics classes are taken by engineering and physical science majors, whereas health-sciences, biology, and pre-medical students enroll in algebra-based physics courses. Students' academic goals and level of interest in understanding physics may differ in the two kinds of physics classes. Equally importantly, algebra-based courses are generally taken in sophomore and junior years as opposed to introductory calculus-based physics courses, which are typically taken by freshmen students in physical science and engineering majors. While the younger students in the calculus-based physics courses may still maintain high-school-based beliefs about physics, those who are in the algebra-based course have gone through several years of college and campus life experience and their attitudes might have shifted since high school. Finally, while women are a minority in calculus-based physics classes (approximately 30% of the cohort), this proportion is double in algebra-based courses. Each of these factors may change the core determinants of self-efficacy and the role of self-efficacy in learning outcomes for students in the algebra-based introductory physics courses.

Another future plan is to design, implement and evaluate instructional strategies to address the issue of gender gap in physics self-efficacy. As discussed earlier, there are some interventions in literature for improving women's self-efficacy [88-90,95,96] in non-physics contexts, which have the potential to further promote higher interest and sense of belonging with a better performance and engagement in the physics classrooms.

4.7 Chapter references

1. American Physical Society (2017). Bachelor Degrees in Physics and STEM Earned by Women, retrieved from https://www.aps.org/programs/education/statistics/womenstem.cfm

2. American Physical Society (2017). Physics Degrees Earned by Women, retrieved from https://www.aps.org/programs/education/statistics/womenphysics.cfm

3. P. V. Engelhardt, in *Getting started in PER*, Reviews in PER Vol. 2, edited by C. Henderson and K. A. Harper (American Association of Physics Teachers, College Park, MD, 2009).

4. E. Seymour, Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, Science Education **86**, 79 (2001).

5. L. McCullough, presented at the ASQ Advancing the STEM Agenda in Education, the Workplace and Society, University of Wisconsin-Stout, 2011, WWW Document, (http://rube.asq.org/edu/2011/06/continuous-improvement/gender-differences-in-student-responses-to-physics-conceptual-questions-based-on-question-content.pdf).

6. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1096 (2012).

7. 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).

8. R. Ivie and K. Stowe, *Women in physics 2000 (American Institute of Physics Rep. No. R-430)* (Washington, DC, American Institute of Physics, 2000).

9. 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).

10. T. J. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, *Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

11. 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).

12. 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?, *Proceedings of the 2018 Physics Education Research Conference* (2018).

13. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemporary Educational Psychology **25**, 82 (2000).

14. E. L. Usher and F. Pajares, Sources of self-efficacy in school: Critical review of the literature and future directions, Review of Educational Research **78**, 751 (2008).

15. M. M. Chemers, L. Hu and B. F. Garcia, Academic self-efficacy and first year college student performance and adjustment, Journal of Educational Psychology **93**, 55 (2001).

16. R. W. Lent, B. H. Sheu, D. Singley, J. A. Schmidt, C. L. Schmidt and C. S. Gloster, Longitudinal relations of self-efficacy to outcome expectations, interests and major choice goals in engineering students, Journal of Vocational Behaviors **73**, 328 (2008).

17. G. Taasoobshirazi and M. Carr, A structural equation model of expertise in college physics, Journal of Educational Psychology **101**, 630 (2009).

18. P. M. Sadler and R. H. Tai, Success in introductory college physics: The role of high school preparation, Science Education **85**, 111 (2001).

19. 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).

20. 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).

21. 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).

22. N. I. Karim, A. Maries and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics, European Journal of Physics **39**, 025701 (2018).

23. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

24. A. T. S. Wee, B. E. Baaquie and A. C. H. Huan, Gender differences in undergraduate physics examination performance and learning strategies in Singapore, Physics Education **28**, 158 (1993).

25. R. R. Hake, Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, American Journal of Physics **66**, 64 (1998).

26. 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).

27. 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).

28. American Institute of Physics, retrieved from https://www.aip.org/sites/default/files/statistics/highschool/hs-studfemale-09.pdf - page=4

29. G. E. Hart and P. D. Cottle, Academic backgrounds and achievement in college physics, The Physics Teacher **31**, 470 (1993).

30. 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).

31. P. Vincent-Ruz, K. Binning, C. Schunn and J. Grabowski, The effect of math SAT on women's chemistry competency beliefs, Chem. Educ. Res. Pract. **19**, 342 (2018).

32. 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).

33. 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).

34. C. Benbow and J. S. Stanley, Sex differences in mathematical ability: Fact or artifact, Science **210**, 1262 (1980).

35. 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).

36. A. Bandura, Self-efficacy. In R. J. Corsini (Ed.), Encyclopedia of Psychology (2nd ed., Vol. **3**, pp. 368-369). New York: Wiley (1994).

37. D. Schunk and F. Pajares, The development of academic self-efficacy. In A. Wigfield and J. Eccles (Eds.), Development of Achievement Motivation: A Volume In The Educational Psychology Series (pp. 15-31). San Diego, CA: Academic Press (2002).

38. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students, International Journal of Behavioral Development **14**, 153 (1991).

39. 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).

40. S. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, Journal of Research in Science Teaching **43**, 485 (2006).

41. S.L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, Journal of Research in Science Teaching **45,** 955 (2008).

42. K. Boden, E. Kuo, T. Nokes-Malach, T. Wallace and M. Menekse, AIP Conference Proceedings, edited Ding, Traxler and Cao, (pp. 60) (2017).

43. P. Huang and S. Brainard, Identifying determinants of academic self-confidence among science, math, engineering, and technology students, Journal of Women and Minorities in Science and Engineering **7**, 315 (2001).

44. S. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, J. Eng. Educ. **87**, 369 (1998).

45. 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).

46. 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).

47. R. Marra, K. Rodgers, D. Shen and B. Bogue, Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy, J. Eng. Educ. **98**, 27 (2009).

48. J. Raelin, M. Bailey, J. Hamann, L. Pendleton, R. Reisberg and D. Whitman, The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention, J. Eng. Educ. **103**, 599 (2014).

49. A. Zeldin, S. Britner and F. Pajares, 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**, 1036 (2008).

50. H. Watt, The role of motivation in gendered educational and occupational trajectories related to maths, Educational Research and Evaluation **12**, 305 (2006).

51. S. Cheryan and V. Plaut, Explaining underrepresentation: A theory of precluded interest, Sex Roles **63**, 475 (2010).

52. S. Devore, E. Marshman, and C. Singh, Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics, Phys. Rev. PER **13**, 010127 (2017).

53. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham and J. Handelsman, Science faculty’s subtle gender biases favor male students, Proceedings of the National Academy of Sciences **109**, 16474 (2012).

54. S. Cheryan, S. Ziegler, A. Montoya and L. Jian, Why are some STEM fields more gender balanced than others? Psychological Bulletin **143**, 1 (2016). doi: http://dx.doi.org/10.1037/bul0000052

55. E. Seymour and N.M. Hewitt, *Talking About Leaving: Why Undergraduates Leave The Sciences*. Boulder, CO: Westview Press (1997).

56. S. Tobias, *They’re Not Dumb, They’re Different: Stalking the Second Tier*, Tuscon, AZ: Research Corporation (1990).

57. T. Denneheny and N. Dasgupta, Female peer mentors early in college increase women’s positive academic experiences and retention in engineering, Proceedings of the National Academy of Sciences **114**, 5964 (2016).

58. S. J. Leslie, A. Chimpian, M. Meyer and E. Freeland, Women are underrepresented in disciplines that emphasize brilliance as the key to success, Science **347**, 262 (2015).

59. S. Upson and L. F. Friedman, Where are the female geniuses?, Scientific American Mind **23**, 63 (2012).

60. L. Bian, S. J. Leslie and A. Chimpian, Gender stereotypes about intellectual ability emerge early and influence children’s interest, Science **355**, 389 (2017).

61. C. Dweck, Mindset: The New Psychology of Success (Ballentine, New York, 2006); C. Dweck, Self-theories: Their Role in Motivation, Personality, and Development (Psychology Press, Philadelphia, 1999).

62. J. W. Osborne, Linking stereotype threat and anxiety, Educational Psychology **27**, 135 (2007).

63. G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women’s performance in physics, International Journal of Science Education **35**, 3050 (2013).

64. A. Maries, N. I. Karim and C. Singh, Is agreeing with a gender stereotype correlated with the performance of female students in introductory physics? Phys. Rev. Phys. Educ. Res. **14**, 020119 (2018).

65. 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).

66. K. A. Shaw, The development of a physics self-efficacy instrument for use in the introductory classroom, *Physics Education Research Conference* **720**, 137 (2003). doi.org/10.1063/1.1807273

67. L. E. Kost-Smith, S. J. Pollock and N. D. Finkelstein, Gender disparities in second-semester college physics: The incremental effects of a "smog of bias", Physical Review Special Topics: Physics Education Research **6**, 0201212 (2010).

68. M. G. Jones, A. Howe and M. Rua, Gender differences in students' experiences, interests, and attitudes toward science and scientists, Science Education **84**, 180 (2000).

69. W.K. Adams, K.K. Perkins, N.S. Podolefsky, M. Dubson, N.D. Finkelstein and C.E. Wieman, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006). doi:10.1103/ PhysRevSTPER.2.010101; B. Zwickl, T. Hirokawa, N. Finkelstein and H. Lewandowski, Phys. Rev. ST Phys. Educ. Res. **10**, 010120 (2014).

70. Activation Lab. Tools: measures and data collection instruments. 2017.Available from: http://www.activationlab.org/tools/.

71. J. Schell and B. Lukoff. Peer instruction self-efficacy instrument. Developed at Harvard University. Unpublished Instrument coetta. 2010.

72. S. Glynn, P. Brickman, N. Armstrong and G. Taasoobshirazi, Journal of Research in Science Teaching **48**, 1159 (2011)

73. L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

74. 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).

75. D. J. Armor, Theta reliability and factor scaling, Sociological Methodology **5**, 17 (1973).

76. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers (1988).

77. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (2013), http://www.R-project.org/

78. R. Chennabathni and G. Rejskind, Gender issues in collaborative learning, Canadian Women Studies **17**, 44 (1998).

79. D. H. Schunk, Attributions as motivators of self-regulated learning, D. H. Schunk and B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (New York, Taylor & Francis 2008) pg 245.

80. R. M. Felder, G. N. Felder, M. Mauney, C. E. Hamrin Jr. and E. J. Dietz, A longitudinal study of student performance and retention. III. Gender differences in student performance and attitude, Journal of Engineering Education **84**, 151 (1995).

81. H. Pettersson, Making masculinity in plasma physics: Machines, labour, and experiments, Science Studies **24**, 47 (2011).

82. Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, J. Res. Sci. Teach. **47**, 978 (2010).

83. J. Eccles, Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices, Psychology of Women Quarterly **18**, 585 (1994).

84. J. Eccles, Where are all the women? Gender differences in participation in physical sciences and engineering, in S. J. Ceci and W. M. Williams (Eds.), *Why aren't more women in science?: Top researchers debate the evidence* pp. 199-210 Washington, DC, US: American Psychological Association (2007).

85. A. J. Gonzalves, A. Danielsson and Helena Pettersson, Masculinities and experimental practices in physics: The view from three case studies, Physical Review Physics Education Research **12**, 020120 (2016).

86. D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They are magic, Review of Educational Research **81**, 267 (2011).

87. D. Paunesku, G. M. Walton, C. Romero, E. N. Smith, D. S. Yeager and C. S. Dweck, Mind-set interventions are a scaleable treatment for academic underachievement, Psychological Science **26**, 784 (2015).

88. G.M. Walton, C. Logel, J. M. Peach, S. J. Spencer and M. P. Zanna, Two brief interventions to mitigate a "chilly climate" transform women's experience, relationship, and achievement in engineering, Journal of Educational Psychology **107**, 468 (2015).

89. D. S. Yeager, G. M. Walton, S. T. Brady, E. N. Akcinar and D. Paunesku, Teaching a lay theory before college narrows achievement gaps at scale, PNAS **113**, 3341 (2016).

90. A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock and G. L. Cohen, Reducing the gender achievement gap in college science: A classroom study of values affirmation, Science **330**, 1234 (2010).

91. P. Pintrich and E. De Groot, Motivational and self-regulated learning components of classroom academic performance, Journal of Educational Psychology **82**, 33 (1990).

92. G. Potvin and Z. Hazari, The development and measurement of identity across the physical sciences, *Proceedings of the 2013 Physics Education Research Conference* (2013) doi:10.1119/perc.2013.pr.058

93. A. Godwin, G. Potvin, Z. Hazari and R. Lock, Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice, Journal of Engineering Education **105**, 312 (2016).

94. A. J. Tomarken and N. G. Waller, Structural equation modeling: Strengths, limitations, and misconceptions, Annual Review of Clinical Psychology **1**, 31 (2005).

95. P. Haussler and L. Hoffmann, An intervention study to enhance girls' interest, self-concept and achievement in physics class, Journal of Research in Science Teaching **39**, 870 (2002).

96. J. M. Harackiewicz, E. A. Canning, Y. Tibbetts, S. J. Priniski and J. S. Hyde, Closing achievement gaps with a utility-value intervention: Disentangling race and social class, Journal of Personal Social Psychology **111**, 745 (2016).

4.8 Appendix B

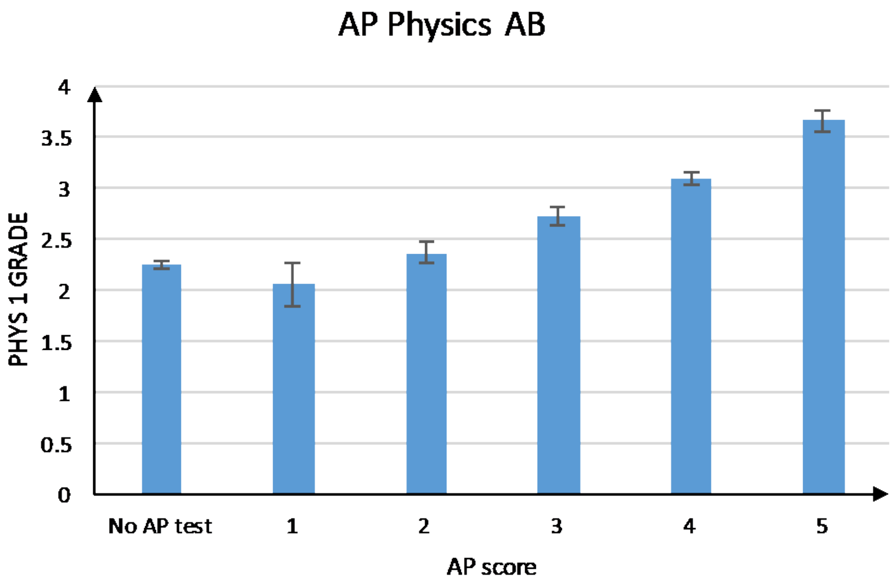

Figure 4-4 Representation of students' physics 1 grades in groups of AP Physics AB scores from 1 to 5. The "No AP test" group corresponds to the group of students who did not have an AP Physics AB test prior to college. Error bars represent standard error of the Physics 1 grade.

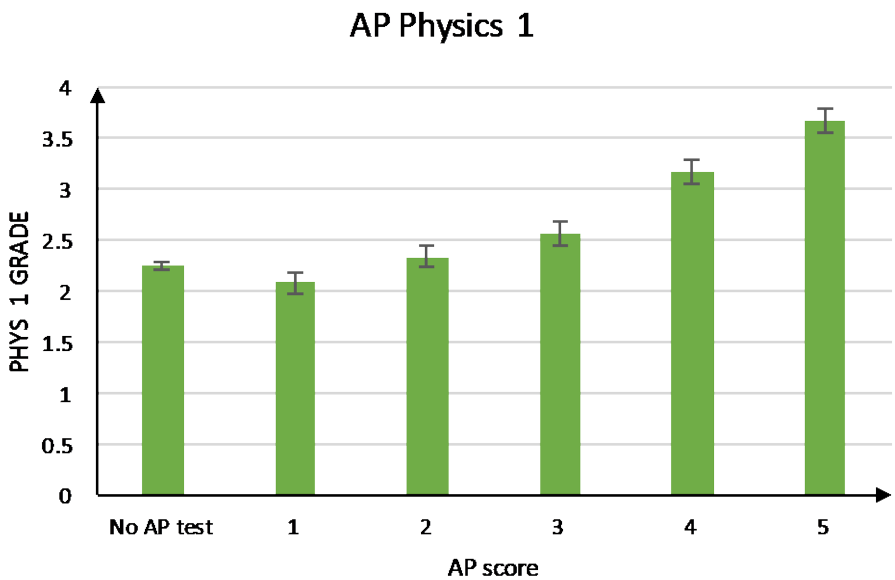

Figure 4-5 Representation of students' physics 1 grades in groups of AP Physics 1 scores from 1 to 5. The "No AP test" group corresponds to the group of students who did not have an AP Physics 1 test prior to college. Error bars represent standard error of the Physics 1 grade.

5.0 Why female STEM majors don't identify with physics: they don't think others see them that way

5.1 Introduction

There has been much interest and effort to enhance diversity in science, technology, engineering, and mathematics (STEM) majors and careers, yet women remain under-represented in many STEM disciplines [1,2]. Among the natural sciences, physics has shown the slowest progress in increasing the representation of women compared to other disciplines such as biology and chemistry [3,4]. Several interventions have been proposed to address the issue of low representation of women in physics including developing better pedagogical methods [5]; improving the efficacy of physics teaching by making the curriculum relevant to all students [6]; and investigating and improving students' attitudes in the physics classroom [5]. The majority of this work has focused on students' cognitive difficulties such as understanding the physics content (e.g. problem solving, reasoning and meta-cognitive skills), and on developing pedagogical strategies to address these difficulties. There have been several studies investigating motivational aspects of learning physics; e.g., self-efficacy, interest, and recognition [7,8], however the implications of these studies have not translated into practice and reform of physics teaching and learning environments. These motivational factors can have a significant influence on learning physics [8] as well as class enrollment and decisions about students' selection of major [9].

In explaining participation in STEM careers, identity has been argued to be a particularly important motivational construct [10]. There are many forms of identity, but most relevant here is identifying with an academic domain, such as physics: students' views about whether they see themselves as a "physics person" [9-16]. Physics identity, which can shape and be shaped by students' learning experiences and classroom interactions, has been shown to predict students' career choices and outcome expectations [9-16].

Unfortunately, prior studies investigating students' physics identity indicate clear gender differences showing that male students are more likely to see themselves as a physics person [9], which may partially explain the differential propensity to major in physics and career aspirations in physics-related fields [11-15]. Therefore, investigating the factors that influence physics identity may play an important role in understanding women's underrepresentation and their underperformance in introductory physics courses. Although there are several studies and frameworks regarding the nature of science identity generally [11-15], additional work is required to understand the motivational factors that influence women's physics identity. If we can understand the relations between the motivational factors that influence physics identity, then future work can focus on developing interventions to scaffold and support its development.

We build on prior work by Hazari and colleagues on physics identity in which they developed a framework for physics identity in high school by adapting the well-known science identity framework by Johnson et al. [11]. Previous studies have focused on documenting gender differences in some of the motivational factors hypothesized to be related to physics identity such as interest in physics, beliefs about conceptual understanding of physics and performance, and perceived recognition by others as well as how these identities impact students' career choices for female and male students [12,13]. However, it is not clear from the prior work to what extent these factors relate to and interact with each other and gender in calculus-based physics courses since the context of the course can be important for the relationships between these factors. Are some factors more important than others in their influence on physics identity? Do these motivational factors meditate the relation between gender and physics identity and if so, how?

To investigate these questions, we examined students' physics identities as well as the hypothesized underlying motivational constructs within the physics identity framework. Specifically, we focused on the college-level calculus-based introductory physics courses at a large state-related research university. Students in these classes are generally physical science or engineering majors and they are typically freshman students, so this study allows us to monitor their motivational attitudes early in their academic program. Typical for courses aimed at such majors, female students in calculus-based physics courses comprise approximately 30% of the classroom.

Our research methodology uses quantitative methods to explore which motivational factors cause change in other factors and mediate the relation between gender and physics identity. As part of this analysis, we administered a motivational attitude survey at the end of the first physics course, a point at which students have now revised their attitudes towards physics based on their first university-level experience and is also the point at which they are less likely to continue in physical sciences and engineering (i.e., it is a very consequential moment). This work can help inform the selection of strategies needed for improving students' physics learning via a lowering anxiety and creating a more supportive environment.

5.2 Background

5.2.1 A sources of gender disparity in physics

Many research studies focus on low representation of women in many STEM fields in terms of societal biases. Seymour and Hewitt argued that the masculine image of current STEM culture is a deterring factor for those who do not identify as masculine [17]. In general, the image of a physicist is portrayed as male, which can make a woman feel less accepted in the physics community. Likewise, the lack of female role models (i.e., famous female physicists or female physics instructors) and being one of the few female students in a physics classroom can communicate to women that their gender is not appropriate for the field. In addition, physics is one of the disciplines that is believed to require a natural ability to be successful in it [18,19], and the concept of being a "genius" or "inherently brilliant" is generally associated with men [20]. These two factors — the perceived masculine nature of the field and gender-based beliefs about brilliance in physics — foreground societal biases about who can excel and thrive in physics, which can impact female students' identity and dissuade them from pursuing study in physics-related disciplines. Archer et al. investigated the impact of physics-related cultural attributions on girls/women's decisions to pursue physics with ten-years of longitudinal data [21] and reported that science-keen girls/young women who name physics as their favorite subject slowly lose their interest due to alienation, discrimination, and gender-biased beliefs about physics.

Further, gendered beliefs, biases, and discrimination can negatively impact women and act as a stereotype threat against them, constraining their performance in the field. In particular, women may fear confirming the negative stereotypical beliefs about their gender in physics if they perform poorly and it can increase their anxiety [22-27]. Faculty's gendered beliefs regarding the students' ability can be one reason for the negative threat that women experience [25]. One study showed that science faculty in biological and physical sciences exhibit biases against female students by rating men significantly more competent when the curriculum vitae are identical except for the name of the student being male or female [25]. Gender-biased culture in physics and negative experiences of female students can create a "chilly" environment for women which can undermine their motivation to engage in learning within the "unwelcoming" culture of physics and other STEM disciplines. For example, women are typically assigned by their male lab mates to "female roles" (e.g., recording data as opposed to collecting data or doing other menial work) in labor division in research lab experiments [26].

Underlying these macro-level factors (e.g., gender biases, traditional gender roles, discrimination against women), students' motivational attitudes and beliefs, which are shaped by biases and discrimination, can at least partly provide a mechanistic explanation for women's under-representation in STEM disciplines. Several studies in physics education have focused on students' motivational characteristics (e.g., factors such as students' interest in and value of science, beliefs about their competence, intelligence mindset views, sense of belonging and identity formation) [9,28-30]. Individual variation in these factors can impact students' learning, persistence in degree attainments, and professional career choices [30-35]. While having positive attitudes can lead to better grades on exams [35] or higher rates of retention in challenging domains [7], negative attitudes, such as low self-confidence or not feeling recognized, can lower interest and increase disengagement from class participation [36].

In physics, researchers have developed and validated various instruments to assess students' motivational beliefs and attitudes about learning [37-47]. Investigating students' motivation in physics can provide education researchers information about how students engage (or not) in learning activities and why some students persist while others do not in physics courses. Although there are several studies of gender differences in students' competency belief (or self-efficacy) [7,8,40-44] and interest [8,40,46] in physics, there are relatively few studies on gender differences in other motivational constructs such as identity, which likely plays a central role in students' retention in physics [11]. Prior work in reducing the gender gaps in self-efficacy and promoting interest in all students has focused on using more interactive teaching approaches [48], adopting active-engagement teaching pedagogies such as Modeling Instruction [49] or describing specific classroom interventions such as Value-Affirmation [50] Further, prior work has investigated some of these motivational factors separately, but it is not clear how they relate to each other and which has the strongest relation to students' physics identities, which is central to students thriving in a discipline.

5.2.2 Prior work on science identity and physics identity

Gee's seminal framework describes identity as "the 'kind of person' one is seeking to be and enact in here and now" [51]. As an example, a physics person typically associates his/her physics identity with being good at physics and math or enjoying solving physics problems which often involves applying mathematical concepts in diverse physical concepts. Identity therefore is hypothesized to be influenced by motivational characteristics which can change over time with individuals' in-class and out-of-class experiences and interactions with peers, teaching assistants and instructors in different circumstances and learning environments [11-16].

Researchers have considered a number of different perspectives on identity formation. One aspect under debate is whether identity is predominantly internal (the individual's private view of themselves) or whether it is a combination of internal and external components (i.e., includes perceptions of how others also view the individual) [52]. For example, external identities can pertain to individuals' manifestation of their identities when performing or acting out who they are in a particular classroom (e.g., in a science classroom, how students perform on the given tasks, how they display their understanding of the subject or the way they communicate in the context of science) [53].

Broadening this internal/external debate, researchers have argued about the full set of underlying factors that influence identity. One critical question is which motivational beliefs are central to an individual adopting a science identity. Carlone and Johnson's science identity framework [10] includes three interrelated dimensions: competence ("I think I can"), performance ("I am able to do"), and recognition ("I am recognized by others"). They tested their science identity framework through ethnographic study to understand the science experiences of successful women in science and science-related careers.

Hazari et al. modified the framework by adapting it to physics specifically rather than science more generally [11]. First, "competence" and "performance" were defined as students' beliefs in their ability to *understand* the subject, and students' beliefs in their ability to *perform* physics tasks [11]. Specifically, competence and performance dimensions focus on students' beliefs and perceptions about their physics-related skills, performances, or conceptual understandings rather than how students can practice and exhibit them in class, as in Carlone and Johnson's science identity framework. Second, the "recognition" variable was renamed as "self-recognition" which, again, stresses the internalized feature of this identity component. Hazari et al. also added a fourth component: "interest" [11]. Carlone and Johnson had investigated *scientists*' identity among professionals who already had developed a certain level of interest in the science domain. But Hazari et al. studied *students*' identity formation in physics classrooms. Students typically have highly varying levels of interest in physics, and therefore it was thought to be a substantial factor in developing students' identities [11-15].

After bringing these modifications together, Hazari et al.'s physics identity framework hinges on four factors: students' interest, beliefs about their competence and performance, and their views on being recognized by others [11]. Those four dimensions of the physics identity framework together make up students' "internalized" identities. Past studies have focused on physics identities' impact on students' physics-related career choices and gender differences in high-school identities [11-16]. Also, some other work has adapted Hazari et al.'s framework to engineering and math contexts and investigated general populations of college students' physics, math, and engineering identities to relate students' engineering choices [16]. Since physics identities are influential in students' career choices and trajectories in physics-related fields, understanding how students' classroom experiences impact the affective factors that forms physics identity is vital. The prior work, however, has not investigated how these factors are related to one another in a predictive framework. Below, we describe our modified framework based on these foundational theories and discuss how to address these questions (i.e., connecting beliefs about competence to major theoretical frameworks in motivation.)

5.2.3 Theoretical model of physics identity

Prior data and theoretical considerations suggest that some further revision to the framework is required. Here, we propose a slightly reframed version of the physics identity framework (see Figure 5-1). Hazari et al.'s performance (i.e., belief in one's ability to perform required physics tasks) and competence (i.e., belief in ability to understand physics content) dimensions are treated as two separate variables, but the factor analysis they conducted on the data of general population of college students suggested that performance/competence is actually a single latent variable [54]. Carlone and Johnson's separation of competence and performance was justified for scientists, but it may be reasonable to represent them with one single construct for a population which has less experience within the discipline and competence and performance become difficult to dis-entangle. Moreover, challenges associated with these issues also suggest that more research with different student populations in different physics learning contexts is needed. In this investigation, our student population consists of freshmen year college students in engineering and physical science majors for whom we propose to combine these two constructs of performance and competence in a more broadly-used motivational construct of self-efficacy. This construct of self-efficacy is defined as students' beliefs in their capability to succeed in a certain situation, task, or particular domain [55]. Self-efficacy has been a central focus in many educational studies and predicts students' actual performance even after controlling for their prior knowledge [56,57]. Students' self-efficacy has also been found to be related to career goals and enrollment in STEM courses [58], as well as the level of persistence in their academic track and long-term goals [34,59]. Equally importantly, recent studies in self-efficacy have shown large gender differences in physics (i.e., female students with lower average self-efficacy than male students) [7,8,40-43]. The gender gap in self-efficacy also exists when comparing similarly performing female and male students [44,45]. These alarming trends show that female students feel less competent in physics than male students regardless of their actual performance. Because self-efficacy can influence students' interest and engagement in the classroom [34], female students' choice of pursuing physics-related careers can be negatively impacted by their lower self-efficacy. With these issues in mind, we replace competence/performance with self-efficacy for the following reasons: 1) to make the connection to the very large literature on self-efficacy;

2) to deal with the empirical finding in both our data and in Hazari et al.' data that competence and performance (at least as perceived by students) is really one construct rather than two separate constructs [54]; and 3) the construct is really about beliefs (which is clearer when self-efficacy is used as a construct) rather than about objective success (which is implied in performance and competence as labels) and may be more suited at least for the students in introductory physics courses.

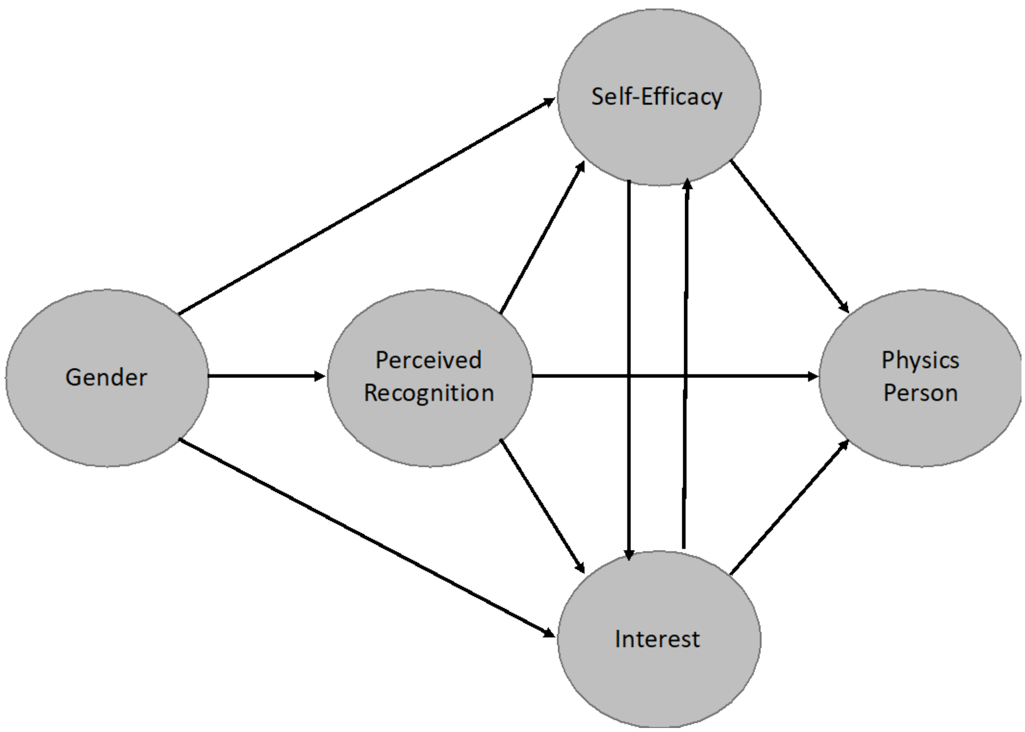

Figure 5-1 Schematic representation of the theoretical framework regarding physics identity.

Each factor is hypothesized to affect identity. For example, the second factor hypothesized to influence physics identity is interest. Much prior work has examined interest and its effect on motivation and learning. Making the science courses more relevant to students' lives and reforming curricula to promote interest in learning can improve students' achievement [60]. Interest is also well paired theoretically with self-efficacy dimensions as connected constructs that predict students' academic outcome expectations and career aspirations within Expectancy-Value Theory [61]. In the theory, Eccles et al. propose that expectancies of success (i.e., self-beliefs) for the domain such as physics are related to students' persistence and engagement that further impacts their course (in the short-term) and career choices (in the long-term). Value of the task is the other aspect of the theory that can enhance students' motivation and persistence in the particular task or field. Value is composed of four components: intrinsic value, attainment value, utility value, and cost. While *intrinsic value* explains individuals' personal interests and enjoyment in engaging with the task, *attainment value* refers to the importance of the task for individuals and relate to their identities. *Utility value*, pertains to how students can relate the task to their lives or to what extent the task can help them succeed in various fields. Finally, *cost* corresponds to negative aspect of engagement such as the amount of anxiety or opportunity cost due to the time spent on the task. These components of value together with expectations can impact cognitive and affective factors of academic learning. In particular, perceiving the task as valuable and connected with individuals' goals, and having an interest in the task and enjoying the task promote better engagement and motivation. We focus on intrinsic value (interest and enjoyment) and label it as "interest".

As the third and final factor of physics identity, we investigate the recognition by others. In particular, we use the construct of "perceived recognition," which refers to students' perception about whether others see them as a physics person or not. We believe the addition of "perceived" is important: while the perceptions held by others about the student is important, we argue that it influences the individual primarily by the extent to which those perceptions of others influence the perceptions of the students of themselves and about being differentially viewed or recognized by others. Individuals can have biased perceptions. They can mis-perceive the intent of others' actions and words, particularly when viewed through the lens of societal expectations and biases. So, an act meant to support an identity (e.g., offering to help) might be viewed as denying an identity (e.g., implying lack of sufficient competence to be in the 'club'). And from a measurement perspective, if researchers are collecting information from individuals about how others recognize them rather than directly from the 'others' doing the recognizing, the individuals are necessarily reporting on their *perceptions* of the recognition provided by others, not the *actual* recognition. Thus, while prior researchers might have intended to measure actual recognition, in reality they measured perceived recognition.

In this framework, recognition may be an important driver of both personal identity and of self-efficacy: students' classroom experiences and interactions with instructors, course assistants and peers (which shape their perception of recognition) can in turn impact their self-efficacy to achieve and class participation. In general, the motivational factors that comprise physics identity relate to and can interact with each other in meaningful ways. However, it is not clear which directions of influence are strongest/generally act as pre-cursors of the others. For example, does interest primarily drive self-efficacy (e.g., by increasing meaningful participation which then builds self-efficacy) or does self-efficacy primarily drive interest.

Students' expectancies and values can also help them build a field-specific identity [61]. In Hidi and Renninger's four phase model of interest development, interest is influenced by self-efficacy, and it later develops into something that is recognized by others [62-64]. In research literature on other motivational constructs (interest, career preferences, etc.), it is indeed common to view self-efficacy as an input, and there is much empirical support for that "causal" direction [59]. However, within the central and well-supported literature on self-efficacy, the focus is on the reverse causal direction: what other attitudes and experiences supports changes in self-efficacy [55]. For example, Bandura's seminal self-efficacy theory proposes that four factors are major sources that influence one's self-efficacy. Of particular relevance to the current work, one of the sources for strengthening individuals' self-efficacy is the social/verbal encouragement and persuasion that one receives from others such as mentors, teachers or family members (or lack of these types of positive enforcements) [55]. Therefore, based upon Bandura's theory, we test the effect of Perceived recognition on self-efficacy. In other words, we hypothesize that these motivational factors are intimately enter-twined in that recognition by others has an effect on the other two factors, self-efficacy and interest, which have an impact on identity. Both expectancy-value theory and the four phase model prominently include self-efficacy (expectancy) and interest as related to identity and its development. However, expectancy-value theory does not appear to include recognition by others. The four-phase model mentions that others (e.g., teachers) can play a role during the emerging individual interest phase and there is focus on how others can support understanding or encouragement. We propose that these beliefs are particularly important for a domain such as physics that has strong socio-cultural biases, e.g., pertaining to a field consisting of brilliant men, and not feeling recognized might have strong effects on one's views of self-competence and interest. We also note that if we find that recognition by others or lack thereof mediates self-efficacy and interest, it may be possible to change the patterns of recognition by instructors and mentors via appropriate professional development activities.

The current research was conducted in the context of introductory level calculus-based physics courses. Carlone and Johnson examined scientists' science identities whereas Hazari et al. studied physics identity of high school students as well as a general population of college students. In the context of physics at the college level, with many stereotypes about the high difficulty level, the primary pathway is unclear. Thus, some model testing work is required that examines variations of the inter-relationships between the components of physics identity. In addition, prior research has tended to use only qualitative methods [10] or quantitative methods [13]. Psychological processes are sometimes implicit, in that individuals can be unaware of the factors that are influencing their actions and beliefs [65]. At the same time, quantitative measures of attitudes can have biases and high levels of noise, and correlations among attitudes can be sufficiently high that separating causal connections from common-cause relationships can be difficult. Thus, cross-validation research that uses quantitative methods which is supported by prior or concurrent qualitative data can be important in this area [10].

We use Structural Equation Modeling (SEM) method to investigate the origins of gender differences in identity in calculus-based introductory physics classes in which women are underrepresented. In other words, the identity framework was used to directly test the underlying relationship of gender to identity components through a mediation model. In a similar study conducted by Godwin et al, students' high school identities and choices in engineering careers were investigated by using SEM for a general population of students [16]. Here, we specifically focus on college female and male students' physics identities in introductory calculus-based physics courses. In particular, we investigate the relation between gender and physics identity via a mediation analysis by using SEM technique. Additionally, we note that Godwin et al. compared the impact of female and male students' physics and math identities on their engineering career choices during the transition period from high school to college with their framework of critical engineering agency. In our study, we are interested in explaining the possible motivational constructs that create gendered patterns in students' physics identity in specific physics courses, i.e., calculus-based introductory physics courses for physical science and engineering majors— in which women are severely underrepresented. Also, in our analysis, self-efficacy is used as a measure of students' beliefs about how competent they are in physics instead of performance and competence dimensions as in previous identity works for the reasons described earlier including the fact that self-efficacy is a broadly used construct in education literature and it is possible to connect the findings with other studies focused on this construct.

We hypothesize that physics identity components in the framework will be correlated to each other and further have a mediating role in the gender and physics identity relationship. Within this framework, we posit the following central research question: *To what extent can gender differences in student's identification with physics be explained in terms of gender differences in physics self-efficacy, physics interest, and perceived recognition as a physics-person by others in the introductory level calculus-based physics courses?* In particular, can the origins of differences in identification with physics be traced to differences in these components? As noted, Godwin et al. have examined the interaction between the motivational constructs, identity and engineering choices using SEM multi-group analysis for a general population of college students [16]. However, prior work did not investigate the possible variations related to the predicting strength of each identity component (i.e., self-efficacy, recognition and interest) on physics identity for students in the calculus-based introductory physics course which is a population of majority male students. Therefore, we aim to explore potential variations between key components of identity across gender.

5.3 Methodology

For the current study, a survey covering the components of the theoretical framework was administered to 559 students at the end of the semester of a calculus-based physics course.

Primarily quantitative methods were used to provide converging analysis in addressing the central research question.

5.3.1 Participants

The 559 participants completing the surveys were students enrolled in one of four different sections of introductory calculus-based physics, which is generally taken by engineering and physical science students in their first year of undergraduate studies. The university provided students' demographic information such as age, gender, ethnic/racial, and academic major information as part of a larger research study using an honest broker process. Both sets of data — demographic and survey responses– were linked by the honest broker representative. During this process, an identification number was given for each student that was based on a hash-function of their university email or student ID number. Therefore, the researchers only had access to the demographics data in this de-identified form.

Note that, the gender data provided by the university records included only binary options given as "female" or "male". We understand that the gender identity is a socio-cultural and non-binary construct and can be described by multi-level categories, but here we are limited with our binary gender dataset for this study. Based on this university data, the survey participants were 33% female and 67% male students; one student did not have a reported gender status and was excluded from further analysis. In terms of ethnic/racial distribution, students were 77% White, 11% Asian, 4% Hispanic, 3% Black, 4% Multi-racial and 1% Other. Regarding academic majors, 61% of students were in engineering track and 39% of students were in physical science majors. Also, 90% of the students were in their first year with an average age of 19.

5.3.2 Survey instruments

We used the motivational surveys of identity, perceived recognition, and self-efficacy, and interest based on prior instruments related to students' motivational characteristics associated with physics [8]. There were other constructs in the survey, but we focus on these four. The survey questions for each construct are given in Table 5-1. The scales were adapted from existing motivational research in physics which was discussed in detail in our prior works [8,40], and the validity of several of the scales examines here were reported in our previous work [8]. The physics identity scale was added later, and therefore additional psychometric validation (CFA results) for that scale is reported in Table 5-1. The prior survey validation and refinement work involved iterative use of Exploratory Factor Analysis (EFA) and one-on-one student interviews both with introductory and graduate level students [8,40,44].

We also performed IRT to check the response option distances for survey constructs. Some of the items have response scales of "Strongly Disagree, Disagree, Agree, and Strongly Disagree" while other items had response scale "NO!, no, yes, YES!". The first group—"Strongly Disagree, Disagree, Agree, and Strongly Disagree"— response scale discrimination was 1.23 and 1.38 while the second group—"NO!, no, yes, YES!"—had 1.55 and 2.11.

Additionally, we checked the inter-reliability for the added identity items. The perceived recognition aspect of the theoretical framework relates to students' perception of how others view them as a "physics person". The survey included three separate items (Cronbach's alpha = 0.86) for family, friends, TA/instructors, respectively. Moreover, this construct focuses on the respondents' beliefs about being recognized, and therefore is called perceived recognition and is appropriately answered by the respondent rather by another individual.

Each of the identity and perceived recognition items involved a four-point Likert response on the scale: Strongly Disagree, Disagree, Agree, and Strongly Agree. "I see myself as a physics person" constitutes the core internal physics identity construct and corresponds to students' beliefs and self-perception in how they designate themselves as a physics person. The survey initially had two identity-related items. However, after an initial factor analysis, we omitted one of the items: "I see myself as scientist/engineer" since it was not loading with the identity construct (or any other construct in the survey).

Table 5-1 Survey questions for each of the motivational scales, along with CFA item loadings (beta and *p*-values of the significance test for each item loading).

Construct and Item	Beta	p
Physics Identity		
I see myself as physics person	1.000	0.00
Physics Perceived Recognition		
My parents see me as physics person	0.914	0.00
My friends see me as physics person	0.899	0.00
My TA/Instructor see me as physics person	0.660	0.00
Physics Self-Efficacy		
I am able to help my classmates with physics in the laboratory or in recitation	0.662	0.00
I understand concepts I have studied in physics	0.723	0.00
If I wanted to, I could be good at physics research	0.722	0.00
If I study, I will do well on a physics test	0.720	0.00
If I encounter a setback in a physics exam, I can overcome it	0.710	0.00
Physics Interest		
I wonder about how physics works	0.652	0.00
In general, I find physics	0.772	0.00
I want to know everything I can about physics	0.786	0.00
I am curious about recent physics discoveries	0.803	0.00
I want to know about the current research that physicists are doing	0.746	0.00

The self-efficacy component of our framework was captured by five self-efficacy items (Cronbach's alpha = 0.83) sampling different manifestations of perceived competence in academic physics (e.g., understanding and performing in various ways); they were all answered on a "NO!, no, yes, YES!" scale that has good psychometric properties and a low reading load, which is important for accurately measuring attitudes of students for whom English is not their native language or have other reading difficulties.

The final construct in the framework is interest, which refers to positive affect towards doing physics-related activities and being curious about the physical world [66] and was measured with five items (Cronbach's alpha = 0.86). The question "In general, I find physics:" had response options "Very Boring, boring, interesting, Very interesting" whereas the question "I wonder about how nature works" had temporal response options: "Never, Once a month, Once a week, Everyday". Three remaining three items were answered on the "NO!, no, yes, YES!" scale. Varying the response scale can lead respondents to slow down and read the items more carefully.

For each survey item, students were given a score from 1 (low) to 4 (high), with higher scores indicating greater levels of interest, self-efficacy, and (recognized) identity. Mean scores across items were calculated items in each scale. For example, a student who answered "Yes!" to three of the self-efficacy questions and "no" to the other two self-efficacy questions would have an average self-efficacy score of (4+4+4+2+2)/5= 3.2. Item Response Theory analyses were previously conducted with these scales to show that the psychological distance between adjacent response items and across items was roughly similar; further more complex factor scores derived from IRT or CFA are so highly correlated with the mean score that there is no practical advantage in using the more complex methods [8].

5.3.3 Quantitative analysis of survey data

In this section, we describe our analysis approach to answer our research question. To examine this question as to whether there are gender differences in identity, we initially conducted MANOVA on students' identity, self-efficacy, interest and recognition scores comparing female to male students. Simple Pearson correlations between constructs provided an initial validation of the proposed theoretical framework.

To provide quantitative validation test of the instruments / separability of the constructs in the current dataset, we performed a Confirmatory Factor Analysis (CFA) on identity and the identity components (perceived recognition, self-efficacy, and interest). The model provides a good fit to the data if fit parameters are above certain threshold measures. Commonly reported fit parameters are the Comparative Fit Index (CFI), which compares the fit of the proposed model to the null model; Tucker-Lewis Index (TLI), which compares the fit of the proposed model the null model but also taking into account the complexity of the proposed model; Standardized Root Mean Square Residual (SRMR), which is the standardized difference between the observed correlation and the predicted correlation; and Root Mean Square Error of Approximation (RMSEA), which is the absolute fit of the model to the data taking into account the amount of data available. CFI > 0.90, TLI > 0.90, SRMR < 0.08 and RMSEA < 0.08 (mediocre fit < 0.10) are considered acceptable fit parameters [67]. However, recent literature considers more stringent cut off points where RMSEA < 0.08 is moderate and < 0.06 is acceptable fit [68].

To quantify the significance and relative strength of the hypothesized path links between gender, identity components, and physics identity with the survey data, we used Structural Equation Modeling (SEM), with a maximum likelihood estimation method vis-à-vis lavaan package in R [67]. Simultaneously estimating all the model links within SEM (rather than separately with different regression models) has increases the statistical power of the analysis and produces estimates for the strength of different paths within the model (e.g., the contribution of self-efficacy to the gender differences in identity). The same model fitness parameters as with CFA are examined to find a model which produced an acceptable fit to the data: CFI, TLI, SRMR, and RMSEA. We began with the saturated model and then created model variations by dropping connections or variables of low strength. In the final model, data from 517 students were included in the analysis because 7% of the initial population had some missing responses.

During testing our hypothesis with SEM model, we used modification indices which suggest certain links between the constructs that can change the structure in order to make the fit better, but we only used the suggestions that were theoretically plausible. In particular, modification indices basically give suggestion to modify the model in order to improve the fit. R lavaan package has built-in function that was used to calculate modification indices based on a given fit measures and give suggested links (positive or negative) between the variables so we could add (if positive) or remove (if negative) those relations from our model. The suggested changes can be regarding covariates or regression models.

We also calculated the indirect effects between gender and physics identity after we run the SEM model. Indirect effect calculations estimate the strength of the mediation between the causal variable (e.g., gender) and the outcome variable (in our case, “physics person”) through hypothesized mediators (e.g., perceived recognition). The indirect effect of gender is calculated by multiplying the coefficients along a given path (i.e., from independent variable to mediator and from mediator to dependent variable). Since there can be multiple paths between a given independent and dependent variable through different mediators, the total indirect effect is calculated as the sum of the contributions through each path. The total indirect effect shows how much of an overall effect is produced by the given set of mediating variables. Examination of the strength of different paths shows the relative contribution of each mediator.

5.4 Results

5.4.1 Test of the measurement model factor structure

The CFA conducted on the proposed model produced an acceptable fit to the data: CFI = 0.975 (> 0.90), TLI = 0.968 (> 0.90), SRMR = 0.035 (< 0.08) and RMSEA = 0.054 (< 0.08). Thus, there is quantitative support for dividing the four constructs as proposed. Further, the factor loadings for each component were higher than 0.7 which is considered as acceptable (see Table 5-1). Therefore, each of the items in the survey were good items for each of their respective constructs.

5.4.2 Gender and physics identity mediation model

As shown in Table 5-2, female students had lower scale means on physics identity with a medium effect size [68]. On average, female students' responses were close to 'no' on the scale whereas on average male student's responses were close to 'yes' on the scale. All three motivational factors also showed lower scores for the female students with similar effect sizes. MANOVA Wilk's Lambda test showed significant gender differences with F (4, 542) = 16.70 p < 0.001. In addition, as shown in Table 5-3, each of the variables was strongly correlated with each other as expected within the theoretical framework. But the inter-correlations were not so high that the constructs could not be separately examined in the SEM. Thus, the pattern of gender differences and inter-correlations were consistent with the overall framework and the SEM would be needed to unpack whether each of the foundational constructs contributed towards explaining the gender differences in physics identity.

Table 5-2 Descriptive statistics for female and male students in which M stands for construct mean value, SD is the standard deviation and N is the number of students. Effect sizes and p-values are presented in the right-most column where * indicates *p* < 0.001 and a minus sign indicates that male students have higher scores than female students.**

	Females N = 182		Males N = 365		*Statistics*	
Construct	*M*	*SD*	*M*	*SD*	Cohen's *d*	p value
Physics Person	2.1	0.9	2.6	0.8	-0.6	***
Recognition	2.2	0.8	2.6	0.7	-0.5	***
Self-efficacy	2.6	0.6	2.9	0.5	-0.6	***
Interest	2.7	0.7	3.1	0.6	-0.7	***

Table 5-3 Zeroth order correlation coefficients of the constructs in the mediation model.

Observed Variable	1	2	3	4
1. Physics Person	--	--	--	--
2. Perceived Recognition	0.79	--	--	--
3. Self-efficacy	0.64	0.65	--	--
4. Interest	0.65	0.60	0.54	--

We initially tested a moderation relation between variables and did multi-group SEM analysis between female and male students. There was no group difference as a result of moderation analysis at the level of weak and strong measurement invariance and at the level of regression coefficients, so we proceeded with mediation analysis using SEM (see Appendix C).

The results of the SEM are presented visually in Figure 5-2. The model fit indices suggest an excellent fit to the data (acceptable fit thresholds in parentheses): CFI = 0.942 (> 0.90), TLI = 0.927 (> 0.90), RMSEA = 0.075 (< 0.08) and SRMR = 0.046 (< 0.08). Figure 5-2 presents the standardized coefficients between each of the variables, and all were statistically significant. From a mediation perspective, it is noteworthy that when all of the predictor variables are entered in the model, the fit of model with the gender variable connected directly to the physics identity gives poor fit parameters (RMSEA = 0.11 > 0.08). That is, the relation between gender and physics identity is mediated by the intervening variables that are given in our initial theoretical framework; female students appear to have a lower physics identity because they have lower levels of physics perceived recognition and interest which in turn strongly drives self-efficacy, which, together with perceived recognition and interest, strongly drives physics identity. Since the strength of the relationships between the components to physics identity were unequal, we also calculated the amount of the gender effect flowing through each of the three key motivational constructs by comparing indirect effects to one another. We calculated the indirect effects by multiplying the coefficients of different paths for each of the model constructs that had a direct connection to physics person. For instance, the self-efficacy's indirect gender effect was calculated by adding two paths: one via perceived recognition and one via interest. For example, for the first path, we calculated 0.27 (gender $\rightarrow$ perceived) x 0.54 (perceived $\rightarrow$ self-efficacy) x 0.15 (self-efficacy $\rightarrow$ physics person) = 0.022. All of the indirect effects are statistically significant, and perceived recognition had by far the largest indirect effect (0.16 ***), followed by interest at approximately half the size (0.07 ***), and self-efficacy with less than one fourth the size (0.034 ***).

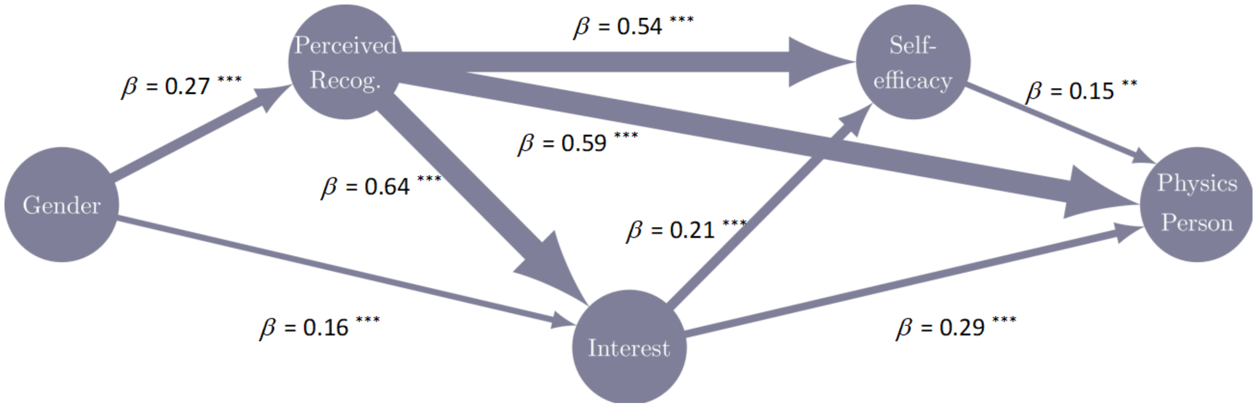

Figure 5-2 Results of the structural equation modeling between gender and being a physics person through self-efficacy, interest and perceived recognition. The line thickness corresponds to the magnitude of β values and p-values <.001 are indicated by * and p-values < .01 are indicated by ****

5.5 Summary and discussion

Students' motivational characteristics in introductory physics courses can inform educational researchers and instructors about how to sustain student engagement and persistence in physics classrooms and beyond [35]. Physics identity is a particularly important motivational characteristic because positive physics identities can foster students' interest in learning, reduce anxiety, increase focus and time-on-task as well as improve performance and retention in STEM majors and careers [11-15]. In this study, we investigated the role that motivational characteristics such as self-efficacy, interest, and perceived recognition play in forming physics identities for male and female students in the calculus-based introductory physics classes that are recognized as gateway courses in college for physical science and engineering STEM majors and careers. This study, in part, adds to the growing evidence on the possible reasons for the large and pervasive gender disparity in STEM participation in physical science and engineering.

For this study, our slightly reframed physics identity model aims to capture the inter-relationships between several closely related motivational constructs and supports earlier work by highlighting the nuances for an important population consisting of college calculus-based introductory physics students. In the context of early university engineering and physical science students, student survey responses were separable according to those four factors as hypothesized. Consistent with prior work [69], we have found statistically-significant gender differences, where male students scored higher than female students in all of the four factors in the modified physics identity framework. These differences suggest that female and male students' identity development may show differences since students' attitudes in physics learning like self-efficacy are influenced in a different way throughout the course.

Most importantly, the modified identity framework provides support for a mediational explanation of the gender disparity in physics in terms of the underlying identity components. Not only did all four components show similarly-sized gender differences; the path model partially explained the gender gap in physics identity. The analysis demonstrated how gender was related to recognition and interest and more importantly how recognition is related to interest and self-efficacy. It is also interesting to note that some relations were not present in the final path model. For example, there was no direct relation between gender and self-efficacy (much of that relation appears to flow through recognition by others). This relation is very consistent with theory that external feedback plays critical role in self-efficacy development. Women appeared to not see themselves as a physics person because they thought others did not seem them as a physics person.

As discussed above, our findings suggest that there are perceived differences between female and male students' experiences and interactions in physics classrooms. These gender differences in students' perceived classroom experiences impact their motivational characteristics.

Female students may experience a decrease in interest and develop lower self-efficacy than male students that may further shape their physics identities in negative ways and can cause gendered patterns in STEM majors and career decisions. These results have implications for both theory of identity and motivational theory. Building on the prior work, this work adds to the theory of physics identity forward to understand the influence of motivational factors on women's physics identity formation. Consistent with early work, we also have shown connections between motivational constructs like recognition by others, self-efficacy, interest, and identity in the context of calculus-based introductory physics courses which could be integrated/revised in the four-phase model and expectancy-value theory.

5.6 Implications

Our identity model suggests a causal relationship between the variables impacting students' physics identities. Through this analysis, we have shown that gender differences in students' perceived recognition influence their self-efficacy, and interest. These gender differences in students' perceived recognition, self-efficacy, and interest ultimately affect their identity in physics. One of the results in this model is the strong effect of perceived recognition in students' other motivational beliefs (i.e., self-efficacy, interest, and identity). Furthermore, students' perceived recognition exhibited a large gender gap where women perceived recognition scores corresponded to negative attitude (i.e., women mostly disagree that others recognized them as a physics person). This result highlights the need for designing and implementing specific classroom interventions targeting these issues. A recent study found that students' physics identity was positively impacted by teachers' implementation of strategies in classroom to increase students' perception of recognition [71]. Getting meaningful recognition from instructors becomes especially important factor to support women's motivation and engagement when we consider the lack of few female role models and teachers in the physics community. Equally importantly, teaching assistants (TA) who facilitate lab courses and recitations also can play a key role in supporting female students' belongingness and perceived recognition. In fact, TAs may have more opportunities to interact with students at a personal level because they teach smaller student groups in recitations and labs.

During the semi-structural interviews that we conducted earlier during survey validation [8,45], female students gave very detailed explanations for why they did not think their TA/ Instructor see them as physics person. For example, one female student noted "I disagree [with the statement], neither of them has really got to know anybody in the classroom on a personal level. We just started this class [Physics 2] but even last semester [in Physics 1], my TA did not really make any effort to know anybody!" Another female student gave the following explanation for why she does not think that her TA/Instructor see her as someone who is good at physics: "That is going to be a 'hard no!' because I don't believe compared to other people I stand out that much in physics, so I don't think they [instructors/TA] really see my name and [say/think] 'ah yes, she is wonderful in physics', no!". Discussions with her suggest that this female student's perception of not being recognized might partly hint at her lower self-efficacy, i.e., feelings of low recognition can further impact women's self-efficacy and interest which is supported by the model in Figure 5-2. Similarly, another female student during the survey validation gave an explanation for her disagreement with the TA/Instructor recognition as her not having innate ability that others in the class had: "Strongly disagree, I am mostly thinking that some talents still yet to come out [to be recognized]." Additional qualitative data collection is needed to better understand the underlying factors related to interactions between instructors/TAs and students in physics courses and how the perceived recognition or lack thereof relates to self-efficacy and interest in physics.

These statements suggest that women may have had negative experiences regarding receiving recognition from their instructors and TAs. The quantitative results given in Table 5-2 showed that both female and male students reported a mean score of recognition and physics person below a positive threshold (i.e., score of 3). However, women had much lower scores for the constructs of perceived recognition and identity as a physics person, which corresponded to a negative response (i.e., score of 2 corresponds to disagreeing with others see them as a physics person). One female student shared her experience in the recitation by stating that when she asked the TA about a particular problem, her TA responded with a comment like "oh, that is an easy question!". That student noted that she felt that her question was devalued in front of the entire class and she felt stupid. She never asked another question in the recitations for the rest of the semester, which emphasizes the essential role of instructors' and TAs' explicitly supportive or perceived non-supportive attitudes in maintaining women's active engagement and sense of belonging in physics courses. We note that to give the benefit of doubt to the TAs, they may not mean anything negative about a student's intelligence when saying that the problem is easy and may respond similarly to men and women. However, due to stereotypes about women in a physics class, male and female students can internalize such responses very differently. Moreover, in calculus-based introductory physics courses in which women are underrepresented, such stereotypes may cause an even larger threat for female students when their questions are devalued (negative perceived recognition), and then their self-efficacy and physics identity can plummet.

At a broader level, the American Association of Physics (APS) and the American Association of Physics Teachers (AAPT) have started to create organizational spaces for women in physics to come together and discuss the current issues to improve diversity, equity and inclusion in physics [71,72]. Moreover, the APS has also been giving Women in Physics Awards to acknowledge the success of accomplished women in the field in order to support women's advancement. However, we argue that there are instructional level strategies that need to be implemented and improved to recruit and retain more women at the undergraduate level and to recognize them.

The investigation presented here suggests that students' perception of receiving recognition from TA/Instructor and peers played an important role in their positive physics identity, and female students had lower perceptions of being recognized than did male students. While small gestures of recognition can have a large positive effect on women's persistence in the field, not receiving any recognition can be interpreted by female students as an indication of a lack of skills or ability since they are the minoritized groups in physics and more susceptible to negative interpretation in such ambiguous situations. Eileen Pollack, who graduated from Yale with a B.S. degree in Physics in 1978, discusses why she pursued a career in writing and being a novelist instead of her dream: getting a PhD from Princeton in Physics. "Not a single professor—not even the adviser who supervised my senior thesis—encouraged me to go to graduate school," and she continues "Certainly this meant I wasn't talented enough to succeed in physics, I left the rough draft of my senior thesis outside my adviser's door and slunk away in shame" [73,74].

So, how can we improve mentor-mentee or student-teacher interactions in physics? One approach involves organizing professional development workshops to train graduate teaching assistants and course instructors, in which they learn about under-represented groups' experiences and their vulnerability in unsupportive classroom environments. These training programs can give some guidance and provide certain strategies in terms of how the TAs communicate with minority groups. Currently TA/Instructor training workshops do not typically focus in any depth the motivational constructs that can severely impact students' learning and have major implications for diversity, equity and inclusion in physics classrooms.

Another approach involves using brief social-psychological classroom interventions. These interventions, e.g., mindset or sense of belonging interventions, are typically done at the beginning of the semester and have been found to boost women's self-confidence and interest in physics, and to reduce the possible stereotype threat in the classroom [75-78]. Such interventions can replace students' stereotypical beliefs that most cannot succeed in physics with the view that brains grow through effort and learning. After such short classroom interventions, students in collaborative learning environment might have more positive approaches when they struggle in group activities despite some students with better prior preparation performing class activities in a quicker way.

Research suggests that developing and implementing field-tested classroom interventions can improve certain minority groups' persistence and belonging in a particular course or their academic achievement in general [50,75-78]. Miyake et al. conducted a values affirmation intervention in college level physics courses as an in-class activity and observed a reduced achievement gap between female and male students' course grade [50]. Similarly, social belonging intervention was conducted for first-year engineering students and the findings show increased GPAs for women in male-dominated engineering majors [75]. Incorporating these types of classroom interventions into teaching can increase women's sense of belonging and self-efficacy and may help them develop positive feelings of recognition by peers and instructors. To address the issue raised in this study, we have recently implemented a 25-minute mindset-belonging intervention in the second week of a calculus-based Physics 1 course during the recitations. Our preliminary results showed promising outcomes in terms of women's performance: While the control group showed large gender gap in Physics 1 performance, this gap was eliminated in the intervention group [78].

Academic and personal support such as family encouragement or peer and faculty relationship and mentoring by faculty can increase the persistence of women and particularly women of color in STEM fields [79-83]. Ong et al. suggests that lack of supportive mentorship relations with the faculty can be crucial but rarely provided for minority groups in STEM disciplines such as women of color [81]. Relatedly, cultural attributions or gender roles in society can be another important aspect that can generate differences in students' motivational characteristics and identities in particular STEM fields [79,84,85]. These biased attributions include sociocultural expectations (e.g., culturally endorsed image of a physicist as "male") and negative stereotypes (e.g., beliefs about men being better at physics than women). Women experience implicit and explicit biases and stereotypes far before they come to college [23]. Instructors' inclusive teaching strategies can help to mitigate this negative threat that many women may experience in physics [71]. Relatedly, one study has found that incorporating discussion sessions about the underrepresentation of women in physics improved female students' physics identities whereas it did not change male students' physics identity [86]. In contrast, women can have lower sense-of-belonging and confidence in physics if instructors exhibit unsupportive, biased attitudes, and discriminating behaviors. In such an "unwelcoming" learning environment, many women can lose their interest further and disengage from course activities which can result in withdrawing from the course and possibly from a STEM major altogether because of these types of experiences [87]. In fact, one study shows that women with equal or better credentials than men exit STEM disciplines due to their negative and non-inclusive classroom experiences, and interactions in their first-year college STEM courses [87].

Unfortunately, negative stereotypes partially arise from the persistent masculine culture and practice of physics, which can act as identity threats for women. Physics is a traditionally male-dominated field. Hence, there is an issue of few female role models for girls and young women to identify with in the physics discipline. In other words, this historical low representation of women can reinforce biased beliefs and stereotypes, and it can send inaccurate messages about how important gender is to succeed in physics [85]. Due to the minority status of women, biases and masculine culture of physics can actually turn into internalized threats for women that negatively affect their behavior pertaining to their academic choices and careers.

Hulleman and Harackiewicz' values affirmation intervention in classroom increased students' interest and exam performance in the course especially for low performers [88]. The positive impact of such interventions on students' interest and course performance can also increase in students' identity in the domain [76]. Female and male students exhibit differences in the value they attribute to certain practices and activities. When teaching, it is important to take into account both group's interest, values, and achievement goals. However, the mainstream teaching and practice of physics generally does not represent the interests, goals and values of girls [87]. In particular, women are more likely to have communal goals and their career choices may align better with using science to help others or the environment [89]. Prior research suggests that when a STEM career is introduced to female students as beneficial to society, their interest in the field increased [90-95]. However, these types of communal goals are not strongly valued in physics which can impact their physics identities negatively. In order to increase women's interest, it may be useful to clearly point out how physics is directly related to helping others and society (such as the applications and practice of physics in medical fields, energy, and environment).

In summary, gender-based differences in physics identities and other motivational factors can reflect the gendered trends in physics and other related STEM disciplines. Gender biased attributions about brilliance, stereotype threat and the emphasis on masculinity in the culture of physics can negatively impact physics identity and potentially lead to less participation and retention of women in physics-related majors and careers. Enhancing gender inclusivity in the physics community and other male-dominated STEM fields can improve women's physics identity and identity in other STEM fields and may yield benefit to women and the STEM discipline itself. Therefore, the issue surrounding the lack of diversity in physics and other science and engineering fields needs urgent action from educational researchers, practitioners, and policymakers. In this regard, fostering women's discipline-based identities in STEM can help increase the retention and advancement of talented and competent women in the STEM disciplines.

5.7 Future directions and limitations

With the findings presented here, it is also imperative to ask what might be causing the differences in female and male students' motivational characteristics and identities and how we may be able change these patterns. In the future studies, we also intend to investigate how measures of performance, such as high school GPA or standardized test scores (i.e., SAT or AP) or physics course grades in college, can interact with students' physics identities.

Although it is a common practice in the research literature on identity, using only one item to measure students' identity may incompletely measure the construct, and future studies should consider including other related items for measuring students' physics identity.

5.8 Chapter references

1. National Student Clearinghouse Research Center (2015). Science and Engineering Degree Attainment: 2004–2014 retrieved from https://nscresearchcenter.org/snapshotreport-degreeattainment15/

2. National Science Board Science and Engineering Indicators. National Science Foundation Report, retrieved from website: https://www.nsf.gov/statistics/2018/nsb20181/digest/sections/preface

3. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016

4. *Women, Minorities, and Persons with Disabilities in Science and Engineering, Occupation* (2017). Retrieved from https://nsf.gov/statistics/2017/nsf17310/digest/occupation/women.cfm

5. P. V. Engelhardt, in *Getting started in PER,* Reviews in PER Vol. 2, edited by c. Henderson and K. A. Harper (American Association of Physics Teachers, College Park, MD, 2009).

6. E. Seymour, Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, Science Education **86**, 79 (2001).

7. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1096 (2012).

8. 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).

9. 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).

10. Carlone and A. Johnson, Understanding the science experience of successful women of color: Science identity as an analytical lens, Journal of Research in Science Teaching **44**, 1187 (2007).

11. Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, J. Res. Sci. Teach. **47**, 978 (2010).

12. Z. Hazari, G. Potvin, R. Lock, F. Lung, G. Sonnert and P. Sadler, Factors that affect the physical science career interest of female students: Testing five common hypotheses**,** Phys. Rev. ST Phys. Educ. Res. **9**, 020115 (2013).

13. Z. Hazari, P. Sadler and G. Sonnert, The science identity of college students: Exploring the intersection of gender, race, and ethnicity, J. Coll. Sci. Teach. **42**, 82 (2013).

14. C. Monsalve, Z. Hazari, D. McPadden, G. Sonnert and P. Sadler, Examining the relation between career outcome expectations and physics identity, *Proceedings of the 2016 Physics Education Research Conference* (2016) doi:10.1119/perc.2016.pr.052

15. G. Potvin and Z. Hazari, The development and measurement of identity across the physical sciences, *Proceedings of the 2013 Physics Education Research Conference* (2013) doi:10.1119/perc.2013.pr.058

16. A. Godwin, G. Potvin, Z. Hazari and R. Lock, Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice, Journal of Engineering Education **105**, 312 (2016).

17. E. Seymour and N.M. Hewitt, Talking about leaving: Why undergraduates leave the sciences. Boulder, CO: Westview Press (1997).

18. L. Bian, S. J. Leslie and A. Chimpian, Gender stereotypes about intellectual ability emerge early and influence children's interest, Science **355**, 389 (2017).

19. S. J. Leslie, A. Chimpian, M. Meyer and E. Freeland, Women are underrepresented in disciplines that emphasize brilliance as the key to success, Science **347**, 262 (2015).

20. S. Upson and L. F. Friedman, Where are the female geniuses?, Scientific American Mind **23**, 63 (2012).

21. L. Archer, J. Moote, B. Francis, J. DeWitt and L. Yeomans, The "Exceptional" physics girl: A sociological analysis of multimethod data from young women aged 10-16 to explore gendered patterns of post-16 participation, American Educational Research Journal **54**, 88 (2017).

22. C. Steele and J. Aronson, Stereotype threat and the intellectual test performance of African Americans, Attitudes and Social Cognition **69**, 797 (1995).

23. G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women's performance in physics, International Journal of Science Education **35**, 3050 (2013).

24. A. Maries and C. Singh, Stereotype threat? Effects of inquiring about test takers' gender on conceptual test performance in physics AIP Conf. Proc. **1697,** 1 (2015).

25. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham and J. Handelsman, Science faculty's subtle gender biases favor male students, Proceedings of the National Academy of Sciences **109**, 16474 (2012).

26. A.T. Danielsson, Exploring woman university physics students 'doing gender' and 'doing physics', Gender and Education **24**, 25 (2012).

27. C. Hill, C. Corbett and A. Rose, Why so few? Women in science, technology, engineering, and mathematics (2010). Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf.

28. M. Wang, and J. Degol, Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields, Dev. Rev. **33**, 304 (2013).

29. C. Dweck, Mindset: The New Psychology of Success (Ballentine, New York, 2006); C. Dweck, Self-theories: Their Role in Motivation, Personality, and Development (Psychology Press, Philadelphia, 1999).

30. L. Blackwell, K. Trzesniewski, and C. Dweck, Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention, Child Development **78**, 246 (2007).

31. V. Tinto, Classrooms as communities: Exploring the educational character of student persistence, J. High. Educ. **68**, 599 (1997).

32. J. Blickenstaff, Women and science careers: Leaky pipeline or gender filter? Gender and Education **17**, 369 (2005).

33. R. Ivie, and K. Stowe, *Women in physics 2000* (American Institute of Physics Rep. No. R-430) (Washington, DC, American Institute of Physics, 2000).

34. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemp. Educ. Psychol. **25**, 82 (2000).

35. P. Pintrich, A motivational science perspective on the role of student motivation in learning and teaching contexts, Journal of Educational Psychology **95**, 667 (2003).

36. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students, International Journal of Behavioral Development **14**, 153 (1991).

37. W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006).

38. E. Redish, J. Saul and R. Steinberg, Student expectations in introductory physics, American Journal of Physics **66**, 212 (1998).

39. K. Edmondson and J. Novak, The interplay of scientific epistemological views, learning strategies, and attitudes of college students, Journal of Research in Science Teaching **30**, 547 (1993).

40. T. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

41. T. Nokes-Malach, Z. Y. Kalender, E. Marshman, C. Schunn and C. Singh, The impact of prior preparation and motivational characteristics on learning outcomes in introductory physics courses, *Proceedings of the 2018 Physics Education Research Conference (2019)* doi:10.1119/perc.2018.pr.Nokes-Malach

42. 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).

43. 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).

44. 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? *Proceedings of the 2018 Physics Education Research Conference* (2018).

45. 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).

46. P. Haussler and L. Hoffmann, An intervention study to enhance girls' interest, self-concept and achievement in physics class, Journal of Research in Science Teaching **39**, 870 (2002).

47. G. Kortemeyer, Correlations between student discussion behavior, attitudes, and learning, Phys. Rev. ST. PER **3**, 010101 (2007).

48. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

49. E. Brew and V. Sawtelle, Modeling instruction for university physics: examining the theory in practice, European Journal of Physics **39**, 054001 (2018).

50. A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock and G. L. Cohen, Reducing the gender achievement gap in college science: A classroom study of values affirmation, Science **330**, 1234 (2010).

51. J. P. Gee, Identity as an analytical lens for research in education, Review of Research in Education **25**, 99 (2000).

52. J. E. Stets and P. J. Burke, A sociological approach to self and identity, Handbook of Self and Identity, 128-152 (2003).

53. Z. Hazari, C. Cass, and C. Beattie, Obscuring power structures in the physics classroom: Linking teacher positioning, student engagement, and physics identity development, Journal of Research in Science Teaching **52**, 735 (2015).

54. R. M. Rock, Z. Hazari and G. Potvin, Physics Career Intentions: The Effect of Physics Identity, Math Identity, and Gender, *Proceedings of the 2012 Physics Education Research Conference* (2012).

55. A. Bandura, Social cognitive theory of self-regulation, Organizational Behavior and Human Decision Processes **50**, 248 (1991).

56. S. L. Britner and F. Pajares, Sources of science self-efficacy beliefs of middle school students, Journal of Research in Science Teaching **43**, 485 (2006).

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

58. R. Lapan, A. Adams, S. Turner and J. Hinkelman, Seventh graders' vocational interest and efficacy expectation patterns, Journal of Career Development **26**, 215 (2000).

59. D. Schunk and F. Pajares, The development of academic self-efficacy. In A. Wigfield and J. Eccles (Eds.), Development of Achievement Motivation: A Volume In The Educational Psychology Series (pp. 15-31). San Diego, CA: Academic Press (2002).

60. P. Haussler & L. Hoffmann, An intervention study to enhance girls' interest, self-concept and achievement in physics class, Journal of Research in Science Teaching **39**, 870 (2002).

61. A. Wigfield, and J. S. Eccles, The development of achievement task values: A Theoretical analysis, Developmental Review **12,** 256 (1992).

62. S. Hidi, Interest: A unique motivational variable, Educational Research Review **1**, 69 (2006).

63. S. Hidi and K.A. Renninger, The four-phase model of interest development, Journal of Educational Psychologists **41**, 111 (2006).

64. K. A. Renninger and S. Hidi, Revisiting the conceptualization, measurement and generation of interest, Journal of Educational Psychologists **46**, 168 (2011).

65. R. E. Nisbett and T. D. Wilson, Telling more than we can know: Verbal reports on mental processes, Psychological Review **84,** 231 (1977).

66. http://www.activationlab.org/tools/

67. R. C. MacCallum, M. W. Browne and H. M. Sugawara, Power analysis and determination of sample size for covariance structure modeling, Psychological Methods **1**, 230 (1996).

68. D. Hooper, J. Coughlan and M. Mullen, Structural equation modeling: Guidelines for determining model fit, Electronic Journal on Business Research Methods **6**, 53 (2008).

69. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (2013), http://www.R-project.org/

70. J. Cohen, L. Erlbaum Associates (1998).

71. R. M. Lock, Z. Hazari and G. Potvin, Physics career intentions: The effect of physics identity, math identity, and gender, AIP Conference Proceedings **1513**, 262 (2013).

72. J. Wang and Z. Hazari, Promoting high school students' physics identity through explicit and implicit recognition, Physical Review Physics Education Research **14**, 020111 (2018).

73. https://www.aps.org/programs/women/

74. https://www.aapt.org/aboutaapt/organization/women.cfm

75. E. Pollack, *The only woman in the room: Why science is still a boys' club*, Beacon Press, (2015).

76. https://www.nytimes.com/2013/10/06/magazine/why-are-there-still-so-few-women-in-science.html

77. D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They are magic, Review of Educational Research **81**, 267 (2011).

78. J. M. Harackiewicz, E. A. Canning, Y. Tibbetts, S. J. Priniski and J. S. Hyde, Closing achievement gaps with a utility-value intervention: Disentangling race and social class, Journal of Personal Social Psychology **111**, 745 (2016).

79. G.M. Walton, C. Logel, J. M. Peach, S. J. Spencer and M. P. Zanna, Two brief interventions to mitigate a "chilly climate" transform women's experience, relationship, and achievement in engineering, Journal of Educational Psychology **107**, 468 (2015).

80. K. R. Binning, N. Kaufmann, E. McGreevy, O Fotuhi, S. Chen, E. Marshman, Z. Y. Kalender, L. Limeri, Lisa, L. Betancur and C. Singh, Changing social norms to foster the benefits of collaboration in diverse workgroups, unpublished manuscript, d-scholarship.pitt.edu/id/eprint/36006

81. M. Ong, T. L. Ko and A. K. Hodari, Agency of women of color in STEM: Individual and institutional strategies for persistence and success. In E. H. Branch (Ed.), Pathways, potholes, and the persistence of women in science: Reconsidering the pipeline. Lanham, MD: Lexington Books

82. R. Kachaf, L. Ko, A. Hodari and M. Ong, Career-life balance for women of color: Experiences in science and engineering academia, Journal of Diversity in Higher Education **8**, 175 (2015).

83. M. Ong, C. Wright, E. Espinosa and G. Orfield, Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics, Harcard Educational Review **81**, 172 (2011).

84. M. Ong, The status of women of color in computer science, In Communications of the ACM **54**, 32 (2011).

85. M. Ong, Challenging cultural stereotypes of 'scientific ability', In M. Pollock (Ed.), Everyday Antiracism: Getting real about race in school, New York: New Press (pp. 114-119, 2008).

86. G. Hackett and N. Betz, A self-efficacy approach to the career development of women, Journal of Vocational Behavior **18**, 326 (1981).

87. C. Ganley, C. E. George, J. R. Cimpian and M. B. Makowski, Gender equity in college majors: Looking beyond the STEM/Non-STEM dichotomy for answers regarding female participation, American Educational Research Journal **55**, 543 (2017).

88. R. Lock and Z. Hazari, Discussing underrepresentation as a means to facilitating female students' physics identity development, Physical Review Physics Education Research **12**, 020101 (2016).

89. R. Felder, G. Felder, M. Mauney, C. Hamrin and E. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Engineering Education **84**, 151 (1995).

90. C. S. Hulleman and J. M. Harackiewicz, Promoting interest and performance in high school science classes, Science **326**, 1410 (2009).

91. E. Reuben, P. Sapeinza and L. Zingales, How stereotypes impair women's careers in science, Proceedings of the National Academy of Sciences **111**, 4403 (2014).

92. J. Eccles, Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices, Psychology of Women Quarterly **18**, 585 (1994).

93. J. Eccles, Where are all the women? Gender differences in participation in physical sciences and engineering, in S. J. Ceci and W. M. Williams (Eds.), *Why aren't more women in science?: Top researchers debate the evidence* pp. 199-210 Washington, DC, US: American Psychological Association (2007).

94. S. J. Ceci, D. K. Ginther, S. Kahn and W. M. Williams, Women in academic science: A changing landscape, Psychological Science in the Public Interest **15**, 75 (2014).

95. S. J. Ceci and W. M. Williams, Understanding current causes of women's underrepresentation in science, Proceedings of the National Academy of Sciences (PNAS) Early Edition **108**, 3157 (2011).

96. S. J. Ceci and W. M. Williams, National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track, Proceedings of the National Academy of Sciences (PNAS) **112**, 5360 (2015).

97. A. Diekman, E. Clark. A. Johnston, E. Brown and M. Steinberg, Malleability in communal goals and beliefs influences attraction to STEM careers: Evidence for a goal congruity perspective, J. Pers. Soc. Psychol. **101**, 902 (2011); A. Diekman, E. Clark and A. Belanger, Finding common ground: Synthesizing divergent theoretical views to promote women's STEM pursuits, Social Issues and Policy Review **00**, 1 (2018).

5.8.2 Appendix C

Multi-group Structural Equation Modeling

We did a moderation analysis to test the group differences in the path analysis. We used the R software package "lavaan" to conduct multi-group SEM. We initially tested for measurement invariance. In other words, we looked at whether the factor loadings, intercepts, and residual variances of the observed variables are equal in the model where we measure the latent constructs so we can confidently perform multi-group analysis. The analysis involves introducing certain constraints in steps and testing the model differences from the previous step. In each step, we compare the model to both the previous step and the freely estimated model, that is, the model where all parameters are freely estimated for each gender group. First, to test for "weak" or "metric" measurement invariance, we ran the model where only factor loadings were fixed to equality across both gender groups, but intercept and errors were allowed to differ. The model was not statistically significantly different than the freely estimated model according to a likelihood ratio test, so weak measurement invariance holds when we compared to freely estimated model where Chi-square difference ($\Delta\chi^2$) = 5.02, degree of freedom difference (Δdf) =10, and non-significant *p*-value = 0.88. Next, we tested for "strong" or "scalar" measurement invariance by fixing both factor loadings and intercepts to equality across gender groups. This model was not statistically significantly different than either the metric invariance model ($\Delta\chi^2$ = 12.264, Δdf = 10, p = 0.2678) or the freely estimated model ($\Delta\chi^2$ = 17.28, Δdf = 20, p = 0.634), so strong measurement invariance holds. Finally, to test for "strict" measurement invariance we fixed factor loadings, intercepts, and residuals to equality. In this step, there was a statistically significant difference from the previous models, therefore “strict invariance” did not hold when we compared to scalar measurement ($\Delta\chi^2 = 37.728$, $\Delta df = 13$, $p = 0.0003$). However, strict invariance is unlikely to hold in most situations. Therefore, since strong measurement invariance holds for this model, we continued on to perform other group comparisons. Next, we ran a multi-group SEM where all regression estimates were fixed to equality for female and male students in addition to the factor loadings and intercepts and we compared this model with freely estimated model. There was no statistically significant difference between two models, so we report the model where regression pathways are equal for men and women. The model fit parameters for this case were not moderate but acceptable (RMSEA = 0.082, SRMR = 0.060, CFI =0.930, TLI = 0.924). The multi-group SEM results suggest that regression pathways (e.g., from self-efficacy to physics identity or interest to identity) did not show differences across gender when we compared to freely estimated model ($\Delta\chi^2 = 25.86$, $\Delta df = 26$, $p = 0.4707$) or to the scalar model ($\Delta\chi^2 = 8.57$, $\Delta df = 6$, $p = 0.1988$). However, the means of the latent variables showed the same gender differences that are reported in the mediation model. That is, there were initially large differences in students’ perceived recognition and slight differences in interest that mediated the effect of gender on identity.

6.0 Navigating your identity in physics courses: the complexity and vulnerability of being a woman

6.1 Introduction

Among science, technology, engineering and math (STEM) degrees awarded in the United States, there is a pervasive issue of low diversity that signals an alarming lack of participation, retention, and advancement of women and underrepresented racial and ethnic minorities [1-4]. Recent work has identified some of the factors and mechanisms underlying the lack of diversity in many STEM disciplines, see a review in Ref. [5], but much remains to be investigated. These ongoing efforts have shown some progress, such that the numbers of female students in some STEM majors (e.g., biology) have dramatically increased in the last few decades [1]. However, there are still other STEM domains, such as computer science, math, and physics, in which men are overrepresented [1].

In the physics community, researchers have started to focus on why certain groups of students, such as women, choose not to pursue physics or physics-intensive fields as a career [6-17]. This research has included an examination of motivational characteristics, such as students' epistemological beliefs about and attitudes toward physics learning [18-36]. Many motivational factors have been shown to affect students' engagement, persistence, and achievement in STEM learning [37-42]. For historically underrepresented student groups such as women, these motivational characteristics might be undermined due to lack of encouragement, negative stereotypes, and inadequate prior preparation, leading to withdrawal from STEM fields [43-52]. Hence, investigating the differences in motivational factors within different student demographics is critical to understanding and addressing diversity, equity, and inclusivity in the physics community.

Here, we describe a study of students' motivational characteristics in introductory physics courses taught at a large, public research university. In particular, we are interested in examining the development of science and engineering students' "physics identity" in the first year of college, a critical time during which the majority of them experience college-level physics for the first time. These calculus-based physics courses are particularly important to study because women are consistently numerical minorities (female students roughly make up 30% of the classes), which can, in part, contribute to the differences in the development of female and male students' physics identities (over and above all of the accumulated societal stereotypes individuals are exposed to growing up) and can impact their career choices. Specifically, we focus on the gendered patterns in students' physics identities and the ways in which motivational factors influence the development of students' physics identity as STEM majors in physics courses.

6.2 Background

6.2.1 Critical student attitudes and beliefs regarding physics

In a learning environment, a variety of motivational beliefs and attitudes, such as students' goal orientations [53-55], expectations for success, and beliefs and attitudes in the classroom, have been linked to academic outcomes [18-25,37,41,52,55-61]. Students with positive motivational beliefs and attitudes exhibit better use of self-regulated learning strategies, participate more in class activities, and show higher-level cognitive processing of the course material [61].

Identity in a particular STEM discipline, has been positioned as a particularly central driver of students' participation, career intentions, and professional choices in many disciplines [62]. For example, having a strong science identity was associated with a higher likelihood of minority students choosing science-related occupations [63]. In physics, high school experiences were linked to physics identity formation, which were then linked with later career intentions [34]. Several other studies extended these findings to identities in other related STEM disciplines, such as mathematics and engineering [64-66]. In one of these studies, physics identity and math identity were found to be strong predictors of choices to major in engineering [65].

However, forming a science topical identity, e.g., a physics identity, can be difficult for women and under-represented racial and ethnic minorities because they must negotiate conflicts between a science identity and the demographic identities of being a woman or a person of color [67-70]. In general, the challenges women face in developing physics identities have been examined through qualitative studies in middle/high school contexts [71], and during the time that students pursue undergraduate physics degrees [72], and doctoral degrees in astrophysics [73].

6.2.2 Motivational factors of students and physics identity

In Gee's foundational theoretical framework, identity is described as "the 'kind of person' one is seeking to be and enact in the here-and-now" [69]. Individuals can have multiple identities that are based upon personality, roles, interactions with others, and social group memberships, and these together define the individuals' core identities [74,75]. Under this broad definition, domain-specific (e.g., physics) identities form a subset. In particular, a domain-specific identity can be defined as a collection of distinctive characteristics shared by all members of a particular social group (e.g., examining students' STEM identities shared by the members of STEM community)

[75]. Domain-specific identities can explain why individuals may decide whether or not to become involved in and progress within the particular domain. For instance, understanding students' science identity may act as a lens in explaining their academic and professional career choices [34,67].

Another important distinction involves internal vs. external identity. Internal identity pertains to individuals' own perception of how they identify themselves within the domain. External identity, on the other hand, focuses on perceptions of how other individuals interpret one's identity, whether based on one's actions in a given domain or stereotypes. Since there are many possible others that may have different perceptions, research has identified three critical groups of others: family, friends, and instructor or teaching assistant (TA) [34]. Note that in both cases, it is the individual's perception (about oneself or others' perception of oneself), because the individual's perception is what drives personal choices. There is a basic drive to have a coherent identity in which internal and external identities agree, and thus each can influence the other [76]. However, at a given point in time, experiences may lead internal and external identities to be out of alignment with one another (e.g., an instructor provides ambiguous comments or comments based on very limited interactions that can impact a student's identity in a discipline).

Some studies have found sizeable gender differences in students' physics identities (i.e., male students in a course, on average, having stronger physics identities), mirroring the representation gap in the physics community [34,64]. The open question pertains to why the difference occurs. The shared working theory is that identity is built upon other motivational factors, and these other motivational factors are shaped by experiences in and out of class (e.g., interactions with peers, teaching assistants, or instructors in different circumstances and learning environments) [34,64-66]. Thus, the prior literature assumes: 1) that women experience stereotype messages, micro-aggressions, male-dominated classrooms, and male-only role models in physics; 2) these gendered experiences influence foundational motivational beliefs and attitudes towards physics; and 3) gender differences in these foundational motivational beliefs and attitudes towards physics produce gender differences in physics identity. *In this study, we expand the third hypothesis and investigate the mechanism connecting between gender and foundational motivational factors and whether this interaction can relate to students' physics identities* (see Figure 6-1). Although not tested directly, we assume such gender differences at this step can happen because of stereotypes regarding careers and career-related identities, such as physics identity.

What are the foundational motivational factors that explain the physics identity across gender? The widely-used Expectancy-Value Theory (EVT) proposed by Eccles et al. integrates three of the foundational motivational factors: expectations for success, subjective task value, and intrinsic interest [77]. The first portion of the model—*Expectancy*— involves individuals' beliefs about how well they will do in a task either in the near or distant future and borrows from Bandura's concept of self-efficacy [78]. Prior research regularly shows large gender differences in students' physics self-efficacy with male students having higher self-efficacy [18,20-22,31,33,35,36]. Several models have proposed self-efficacy as a foundation of identity [79].

The second part of the EVT model—*Value*—actually involves multiple components. The most commonly discussed value components are intrinsic value (enjoyment) and utility value (usefulness). When individuals perceive a task or domain as intrinsically valuable or extrinsically valuable to their goals, this can promote better engagement and motivation [34,67,77,80]. Intrinsic value in the original model is now often discussed as *Interest*, which is defined by positive emotions accompanied by curiosity and engagement in the particular content [81-83]. Relatedly, Hidi and Renninger investigated several phases of how to form and increase interest in specific domain. Their model, in particular, highlights that interest cannot be maintained unless it is supported by the environment [82]. For instance, teachers need to encourage students by providing them useful feedback and providing mentoring and support that can bolster students' interest in the subject being taught. Several models have been proposed regarding both utility value and interest being foundational to the development of individual's identity in a domain [77,81], although other models propose closer connections mainly to interest [40,81].

Sense of belonging is not part of EVT but is another factor that has been proposed as important to the development of an individual's identity in a domain [84,85]. It pertains to making students feel that they belong within a particular community [85]. The feeling of belongingness is related to being valued, respected and accepted in academic spaces [85]. Studies suggest that women in STEM fields might have lower levels of sense of belonging due to societal cues and masculine culture of STEM [86,87]. Furthermore, implicitly or explicitly biased attitudes of faculty (e.g., academic advisors, instructors or graduate teaching assistants) can also make women feel that they do not belong in the community [88]. Therefore, developing a strong sense of belonging in science courses can be a significant factor in whether and when students start building their identity as a scientist.

6.2.3 Theoretical model

In this study, we test a model that physics identity is developed from four key motivational factors (see Figure 6-1): self-efficacy, interest, value, and sense of belonging in a physics course. These factors are thought to affect both whether students view themselves as a physics person and whether they perceive others' recognition of their identity (these internal and external identities can interact with each other to form an individual's identity in a particular domain). There are numerous studies establish the role of physics identity in career intentions and choices in the STEM fields [34,64,65,72,73], but there are very few studies that have attempted to explore the connection between motivational factors and identity formation [79]. Moreover, no prior studies have examined whether the connections between the motivational factors and identity are gender-specific in calculus-based physics courses.

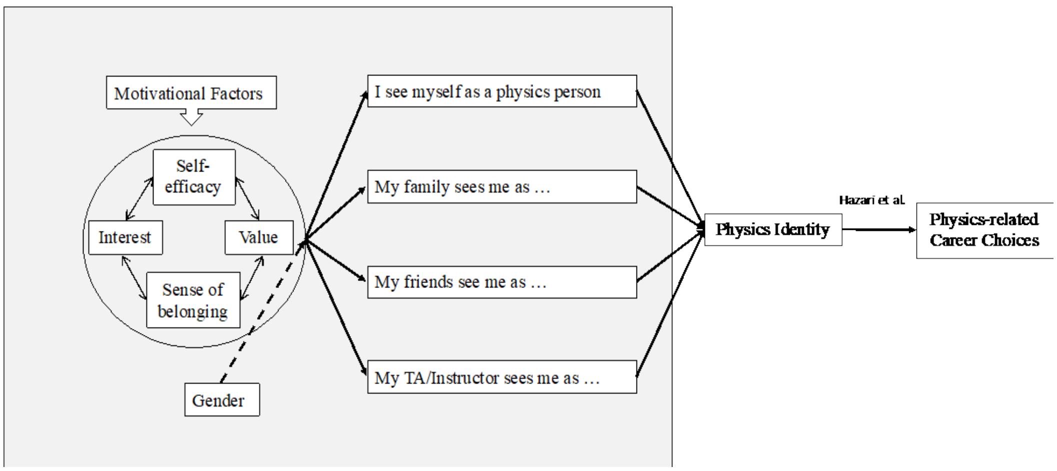

Figure 6-1 Schematic representation of gendered relationships of foundational motivational factors to the components of physics identity, which has been previously linked to career choices.

There are multiple ways in which identity in a discipline, e.g., physics, could be differentially related to motivational factors. For example, in one proposed model, identity could be a coherent and independent construct from the other motivational factors across both genders but there may be differential strengths of association between identity and the motivational factors (e.g., a stronger connection between physics identity and self-efficacy in women compared to men). This theoretical model can be tested using factor analysis and regression. For example, operationally, a factor analysis would produce a separate factor for identity, and then regression analysis would find differential relationships between the foundational motivational factors and identity (e.g., significant interaction terms between gender and a motivational component in the multiple regression).

In a second proposed model, identity could be so closely aligned with one of the motivational factors that it does not produce a separate factor, but which factor identity aligns with varies by gender (e.g., identity aligns with interest in men but with self-efficacy or competency beliefs in women). If this model is correct, operationally, a factor analysis would not produce a separate factor for identity but generate loadings under one of the motivational constructs proposed in our model such as Interest, Self-efficacy, Value or Sense of Belonging. Relatedly, the factor analyses done separately by gender would produce different alignments for how physics identity items load under different constructs. For instance, the identity items can factor under sense of belonging for female students while they load under interest for male students. In other words, we would find that identity items factor out under a particular construct for both gender groups or there might also differences across gender.

In a third proposed model, the different components of identity (internal versus external items) do not cohere with each other overall but may cohere for only one gender group (as proposed in the second model). Individuals seek to have coherent identities in which the way they perceive themselves is consistent with how others view them. However, in cases of stereotypes and under-representation, the under-represented groups can perceive mismatches between how they think of themselves and how others think of them. In particular, internal identity might be closely connected to interest but external identity, in the case of women, might be closely connected to self-efficacy or sense of belonging due to the stereotype about women's ability to excel in physics. The third model adds further dimension to our interpretation in which the identity component itself shows differences how they factor out and there might also be interaction effects when we investigate by gender. For instance, men's internal and external identities can factor under one of the motivational constructs such as interest but there might be differences between the two kinds of identity for women.

Calculus-based physics courses are fundamental to STEM curricula and are required to be taken by all engineering students. Since achievements in these gateway courses predict students' success later in the academic program [89], examining the ways in which freshman students build physics identities in these introductory level physics courses is a particularly important objective of our research. Our primary research question is: *What motivational factors relate to male and female students' internal and external identities?* In addressing this question, we aim to understand how the survey results relate to each of the three models.

6.3 Methodology

6.3.1 Development of motivational survey and instrument validity

This study used previously validated surveys to measure students' physics identity, self-efficacy, interest, and sense of belonging within students enrolled in calculus-based physics courses. All of the motivational constructs were embedded in the context of physics (e.g., self-efficacy in physics). The survey instruments were part of a larger survey that included other motivational constructs. The development and validation of these surveys are reported in prior work [18-22,60]. The surveys were based upon previous instruments [24,55-59] and validation of the revised instruments involved iterative use of Exploratory Factor Analysis (EFA), Confirmatory

Factor Analysis (CFA), and one-on-one student interviews both with introductory undergraduate and graduate students [18,60].

The *Physics Identity* survey evaluated the extent to which students see themselves as a physics person and their perception about whether others see them as a physics person [34]. The *Interest in Physics* survey measured students' enthusiasm in learning physical laws and their curiosity about cutting-edge research in physics [90]. The *Value of Physics* survey focused on students' views of whether learning physics will help them succeed in their major or later STEM career [90]. The *Physics Self-efficacy* survey measured students' confidence in their understanding and solving physics problems [91-94]. Finally, the *Sense of Belonging* survey measured whether students feel that they belong in the surveyed (introductory physics) classroom or not [95], and it focused on students' sense of being accepted and valued by their peers and instructors/TAs in the academic classroom setting [85].

The items in the survey were designed on a Likert scale of 1 (low endorsement) to 4 (high endorsement) (for Sense of Belonging, 1 to 5) [96]. The specific rating scales were purposely varied to provide a more valid overall measure of intensity by including frequency (e.g., Never, Once a month, Once a week, Every day) and intensity (NO!, no, yes, YES!; Strongly Disagree, Disagree, Agree, Strongly Agree). While having a lower score was an indication of negative endorsement (e.g., lower level of interest, self-efficacy, value, identity, and sense of belonging), higher scores corresponded to more positive beliefs (e.g., average self-efficacy score of 4 referred to the highest confidence level for a student). Some questions were reverse coded. For instance, one of the items in the sense of belonging was "Sometimes I worry that I do not belong in this physics class". A student responding to this question as "Mostly true" was given a score of 2. We only included 2-3 such questions because too many reverse coded questions could create ambiguity for respondents.

Table 6-1 shows all the constructs and the observed scale reliabilities in this dataset (Cronbach's α) for each motivational construct [97]. The reliabilities were good (>0.8) to excellent (> 0.9). We calculated mean scores for each motivational construct that showed roughly equal separation between levels on the Likert scales. For example, a student who answered "Yes!" to three of the self-efficacy questions and "no" to the other two self-efficacy questions would have an average self-efficacy score of (4+4+4+2+2)/5= 3.2.

Table 6-1 Listed for each motivational construct are number of items, sample items, Likert scale range, and observed Cronbach α. (R) corresponds to a reverse coded item.

Measure	# of items	Sample item	Likert Scale	Cronbachs' α
Physics Identity [34]	4	I see myself as a physics person My friends see me as a physics person	(1-4)	0.90
Interest in Physics [90]	5	I wonder about how physics works I am curious about the recent discoveries in physics	(1-4)	0.86
Physics Self-efficacy [91-94]	5	If I encounter a setback in a physics exam, I can overcome it I am able to help my classmates with physics in the laboratory or in recitation.	(1-4)	0.83
Valuing Physics [90]	4	Learning physics will help me succeed in my future career Learning physics will help me get a job that I want.	(1-4)	0.90
Sense of belonging in physics classrooms [95]	5	Sometimes I worry that I do not belong in this physics class (R) I feel like I can be myself in this class	(1-5)	0.84

6.3.2 Participants

The 559 participants completing the surveys were students enrolled in one of four different sections of introductory calculus-based physics. Each section was taught by a different instructor. This course is generally taken by engineering and physical science students in their first year of undergraduate studies.

Students' demographic information such as age, gender, ethnic/racial and academic major information were provided by the university using an honest broker process that linked survey responses with students' demographics in a de-identified way. Based on this university data, the survey participants were 33% female and 67% male students; one student did not have a reported gender status and was excluded from further analysis. In terms of ethnic/racial distribution, students were 77% White, 11% Asian, 4% Latinx / Hispanic, 3% Black, 4% Multi-racial and 1% Other. Regarding academic majors, 61% of the students were in an engineering track and 39% of students were in physical science majors. Also, 90% of the students were in their first year with an average age of 19.

6.3.3 Procedures

The motivational survey was administered in paper form during the last recitation of the semester by the graduate Teaching Assistants (TA). The survey took approximately 5-10 minutes for respondents to complete. Course instructors were encouraged to give students extra credit to participate in the survey. The survey response rate was 73% of students enrolled in the course. We have used the current dataset in a previous study to examine the attitudes that mediate the relation between gender and students' views of the extent to which they see themselves as a physics person [60]. In this study, we used the same dataset to carry out an in-depth analysis of the factors that influence the formation of female and male students' views of the extent to which they see themselves as a physics person as well as their perception about how others recognize them as a physics person.

6.3.4 Analysis

MANOVA Analysis of Gender Differences in Physics Identity. We first analyzed the survey data to compare the mean scores by gender for two components of physics identity. Multivariate analysis of variance (MANOVA) was used to test for statistical significance differences between female and male students on each identity component. MANOVA is appropriate when multiple correlated dependent measures are being examined in parallel. Based upon prior research, we selected the first item for physics identity – “I see myself as a physics person”— as a separate “physics person” construct [60]. Then we obtained an average score for the three remaining identity questions by labeling it *perceived recognition* by others (i.e., by family, friends and TA/Instructor), based in work by Hazari et al. [34]. We also examined each identity item in *the perceived recognition* construct separately since there might be variations in students’ responses depending on who is recognizing them as a physics person. For instance, students’ perception about being recognized by friends versus by TA/Instructor may vary.

Principle Component Analysis by Gender. To test for gender differences in the motivational factors underlying physics identity, we performed a Principal Component Analysis (PCA), which is essentially Exploratory Factor Analysis in our case. PCA is a widely-used method to extract underlying dimensions within large datasets of survey data. Separate PCAs were conducted for male student and female student data in order to understand whether male and female students’ group identities behaves in similar or different ways compared to other motivational constructs, both at the overall identity level and also for the different components of identity separately (see Figure 6-1). With this framework, we tested the extent to which there were four independent motivational factors (self-efficacy, sense of belonging, interest, and value) and the extent to which these were integrated into students’ physics identities.

We implemented PCA using a Promax rotation with Kaiser Normalization given that we expected the factors to be correlated with one another. We initially used the standard criterion of eigenvalues > 1 for determining the number of extracted factors. We also examined the scree plot to observe the relative change in modeled variance with growing number of generated factors. The scree plot leveled off as three factors (see Appendix D), and results with more than three factors did not produce more theoretically coherent factors nor different patterns of results for the core research questions than models with three factors. Therefore, we focus on the PCA analysis results with three output factors.

We then analyzed the component loadings for each survey item in terms of their connection to each of the three resulting factors. Higher loading values correspond to a closer affinity with the particular factor. Researchers usually suppress the factor loading below 0.3 or 0.4 since those contribute very little to variance in the factor (equals to the square of the factor loading). Therefore, we discuss results for 0.4 and above as primary, and between 0.4 – 0.3 as secondary.

6.4 Results and discussion

6.4.1 Gender differences in physics identity

Comparing female and male students' average physics identity scores for each item in the survey, we found a gender gap in students' perception of both being a physics person and being recognized by others as a physics person (see Figure 6-2). MANOVA with Wilks' test revealed that there is a statistically significant gender difference for all physics identity items in the survey (F (1,556) = 11.95, p <0.001). In other words, female students reported significantly lower identity scores than male students.

We also calculated effect sizes (Cohen's *d* values) of the gender gap in each physics identity item to provide a standardized measure (see Figure 6-2). Effect sizes range from 0.4 to ~ 0.6 suggesting moderate level effect size, which correspond to approximately 0.5 standard deviation difference between female and male students' physics identities [98]. The gender gap becomes especially important when we consider that the data were collected at the end of freshman students' first semester, which is often a critical transition point for many STEM majors in engineering and physical sciences deciding whether to continue within STEM overall and select particular majors. The gender differences in identity are similar in size to the ones in interest (d=0.63), self-efficacy (d=0.55) and belonging (d=0.51), and larger than the ones in value (d=0.26).

Mean scores across family or friends' recognition were very similar. However, the mean score's for TA/Instructor's recognition were lower than means for the other physics identity items in the survey. Using a paired t-test to contrast responses to Q18 against the mean of Q15-17, the difference was statistically significant for both female ($t = 2.63$, $p < .01$) and male ($t = 8.37$, $p <0.001$) students. This analysis gives a clear demonstration that students' views regarding how others see them as a physics person might vary depending on the level of recognizer's role in students' learning (e.g., whether the person is a family member or TA/instructor), which can further impact how a student see herself/himself as a physics person as a result of receiving (or not) this recognition.

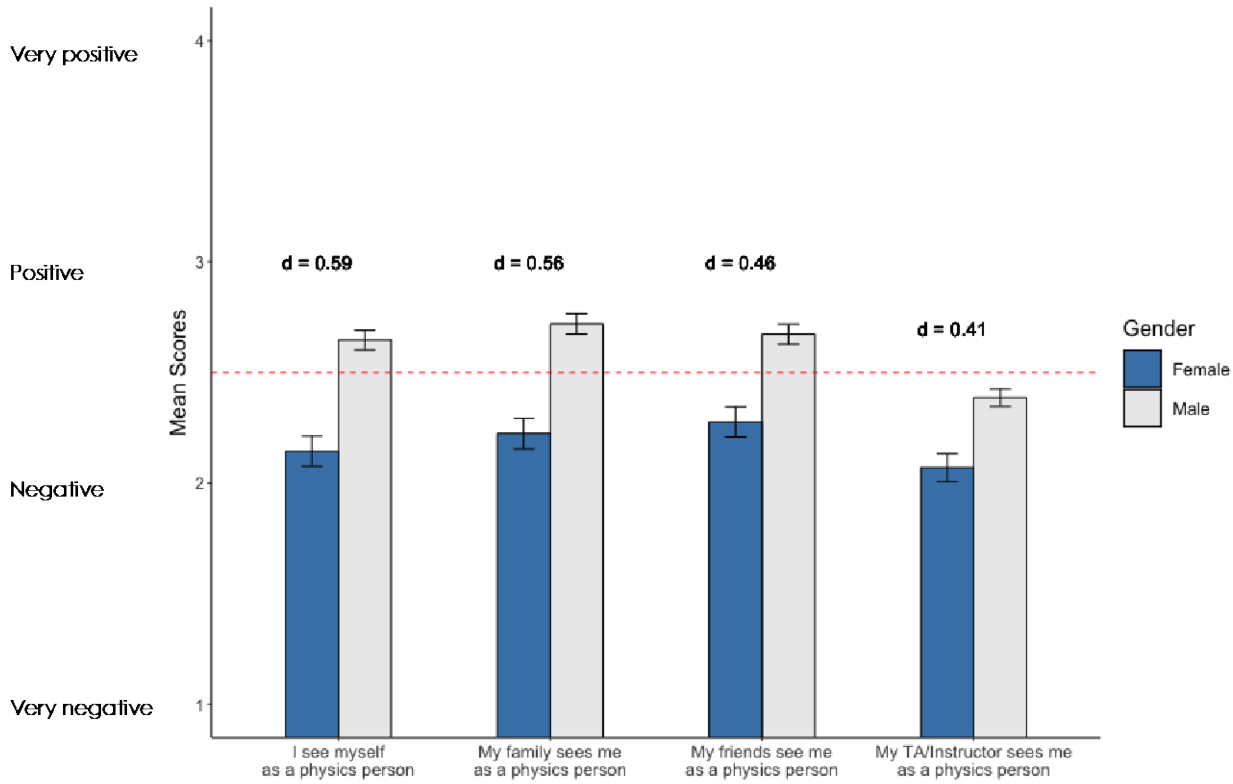

Figure 6-2 Mean scores (with standard error bars) for each physics identity item for female and male students. "d" denotes Cohen's *d* for the gender contrast. Red dotted line indicated the score of 2.5, which corresponds to neutral response for identity.

6.4.2 Principal component analysis by gender

The separate PCAs by gender resulted in several very important and intriguing findings (see the rotated factor matrix by gender in Table 6-2). In Table 6-2, the factor loading values between 0.3 and 0.4 are shown (which yields what we call secondary order results). However, it is also commonly accepted and practiced analysis to show and interpret only the values 0.4 and above (which we call primary order results and show with boldface font in the Table 6-2). General approach is to suppress the values under 0.4, therefore we highlight those in the Table 6-2, but community also often interprets the values between 0.4 and 0.3 so we include those loadings in our analysis as secondary results. For the primary order results, the estimated factor loadings or weights are similar for both gender groups in a number of ways. First, the items associated with "Interest" and "Value" dimensions consistently factored out as two separate constructs from one another. Although more intrinsic and extrinsic motivational drivers are conceptually distinct within Expectancy Value Theory, the two are often so highly correlated that it is hard to separate them within factor analysis [77]. It appears that within this population of engineering and physical science majors, interest in and value of physics are strongly independent components (each one has average item factor weighting values greater than 0.7). Thus, the first component in Table 6-2 is labeled as "Interest" and the third component is labeled as "Value".

Second, the component related to self-efficacy, the third component of Expectancy Value Theory, is a separate factor. Focusing on factor loadings above 0.4, the self-efficacy items consistently load on their own factor separate from interest and value, as predicted by Expectancy Value Theory. The factor weights are somewhat lower, especially for items that also loaded with weights between 0.3 and 0.4 on the interest scale.

Third, despite our initial hypothesis of having four separate motivational constructs (interest, value, self-efficacy, belonging) that might be separate or integrated with identity, the PCA analysis reveals that for both genders, sense of belonging and identity were not separate from the EVT constructs. Instead, sense of belonging was closely tied to self-efficacy, which was therefore labeled as "Self-efficacy/Belonging". Identity was closely tied (with one gendered-exception) to interest, which is therefore labeled "Interest/Identity".

There was one main and important difference between genders from the primary results: Q18 – "My TA/Instructor sees me as a physics person" loaded similarly with the other Identity items (i.e., with Interest) for male students (~0.56), it strongly factored with "Self-efficacy/Belonging" for female students (~0.68). In other words, female students' perception of being recognized as a "physics person" by their TA/Instructor is intimately related to their sense of belonging and self-efficacy in a physics classroom. By contrast, for male students, TA/Instructor recognition was closely aligned with their internal interest in physics.

Next, we discuss the secondary results (weaker effect than the primary results) given in Table 6-2, which includes factor loading values between 0.3 and 0.4. For male students, Q18 – perceived recognition of their TA/Instructor— shows a smaller connection with Self-efficacy/Belonging (~0.34) like female students as discussed in the primary results. In particular, students' perception of getting recognition from their course instructor or teaching assistant appears to be generally important aspect for building positive self-efficacy and sense of belonging for all students, but it is particularly strong for female students. Similarly, Q17 (My friends see me as a physics person) shows a gender difference in the secondary results. Q17 factors out under the "Interest" component for male students with a high loading (~ 0.82). For female students, the same item also shows a main contribution to the "Interest" component (~0.52), but for female students only, there is an additional loading under the Self-efficacy/Belonging component. In other words, female students' Self-efficacy/Belonging is positively related to recognition by friends as a physics person but not for male students. These results (both for Q17 and Q18) highlight that the extent to which students' physics identity integrates with other motivational factors may vary across gender groups.

Finally, Q1-Q5 measure students' interest and curiosity with discoveries in physics or current research in physics and shows a higher loading on the "Interest" factor for female students compared to male students. We also note that the Value component did not connect with any aspect of student's physics identity, even within the 0.3 to 0.4 factor loading range, for either gender. This suggests that motivational factors such as identity are inherently internal constructs, whereas utility Value that one may derive from physics as measured by the validated survey is a more external construct.

Table 6-2 Factor loadings from the Principle Component Analysis (PCA) results for female and male students. Loadings > 0.4 are in bold. Loadings below 0.3 are hidden.

		Female			Male		
	Questions \ Components	Interest /Identity	Self-efficacy /Belonging	Value	Interest /Identity	Self-efficacy /Belonging	Value
1.	I wonder about how physics works	**0.85**			**0.79**		
2.	In general, I find physics	**0.81**			**0.70**		
3.	I want to know everything I can about physics	**0.91**			**0.69**		
4.	I am curious about recent physics discoveries	**0.95**			**0.69**		
5.	I want to know about the current research that physicists are doing	**0.91**			**0.65**		
6.	I am able to help my classmates with physics in the laboratory or in recitation		**0.67**		0.36	**0.50**	
7.	I understand concepts I have studied in physics		**0.74**		0.35	**0.51**	
8.	If I wanted to, I could be good at physics research		**0.52**		0.38	**0.40**	
9.	If I study, I will do well on a physics test		**0.75**			**0.65**	
10.	If I encounter a setback in a physics exam, I can overcome it		**0.63**			**0.68**	
11.	Learning physics will help me in courses in my major.			**0.79**			**0.81**
12.	Learning physics will help me achieve admission into graduate school and/or medical school.			**0.84**			**0.78**
13.	Learning physics will help me get a job that I want.			**0.98**			**0.81**
14.	Learning physics will help me succeed in my future career.			**0.88**			**0.84**
15.	I see myself as a physics person.	**0.64**			**.68**		
16.	My family sees me as a physics person.	**0.46**			**.82**		
17.	My friends see me as a physics person.	**0.51**	0.33		**.82**		
18.	My physics instructor and/or TA sees me as a physics person.		**0.68**		**.56**	0.34	
19.	I feel like an outsider in this class.		**0.69**			**0.71**	
20.	I feel comfortable in this class.		**0.74**			**0.74**	
21.	I feel like I can be myself in this class.		**0.67**			**0.59**	
22.	Sometimes I worry that I do not belong in this physics class.		**0.89**			**0.88**	
23.	When I get a poor grade on a physics assignment or exam, I feel that maybe I don't belong in physics class		**0.98**			**0.80**	

6.5 General discussion

6.5.1 Summary

This study investigated male and female students' physics identities during an important transition point in the educational pathways for physical sciences and engineering: midway during introductory level calculus-based physics courses. Prior research has revealed the importance of physics identity on students' career intentions and choices in STEM, and therefore understanding why and how students' physics identity formation may differ across gender is critical to enhancing women's experiences in STEM. Our novel results indicate a moderate level of gender differences in students' internal and external physics identities, both in their internal coherence and their relations to other motivational factors that have been previously linked to identity formation [79]. Moreover, while some of the motivational constructs shown in Figure 6-1 such as self-efficacy, interest, and sense of belonging exhibited some gendered patterns in terms of how they factor out and relate to students' identities in physics, value as the other construct in the framework showed similar patterns across both genders.

In the research discussed here, we found that the components of Expectancy Value Theory (interest, self-efficacy, utility value) were independent of one another but sense of belonging in physics was so strongly aligned with physics self-efficacy that it did not factor separately for either gender. A similar result was found for physics identity ("I see myself as a physics person") in that it was so strongly aligned with interest and they did not factor out separately for either gender. This result is consistent with the second proposed model where physics identity was not found as a separate construct but closely related to other motivational constructs such as interest. Researchers rarely examine the factor structure of multiple motivational factors together. Here, we find that this matters much in that certain factors load – or are considered together and the results were consistent for both gender groups which has implications for what we think about these factors in this context.

Similar relations between the constructs of sense of belonging and self-efficacy were found in factor analysis in that they loaded as one latent variable instead of two distinct components. In the literature, while students' physics self-efficacy across gender has been examined in numerous research studies [18,31,33,35,36], little regard has been given to examining students' sense of belonging in physics classrooms. Our findings reported here suggest that students' sense of belonging in physics is strongly connected to their physics self-efficacy (in the exploratory factor analysis, the sense of belonging items and self-efficacy items on the survey always factored under the same latent variable), which is unfortunate because there are often large gender differences in self-efficacy and sense of belonging, much larger than any actual performance differences. While the causal direction is ambiguous based on the current data (e.g., self-efficacy causes belonging or belonging causes self-efficacy), prior research supports both directions [85,99]. Therefore, creating learning environments in which students feel comfortable to discuss their opinions and, in return, feel valued by peers and the instructor/TA is likely to be influential in increasing students' self-efficacy in physics.

Equally importantly, while there were general similarities across gender in how identity related to other motivational factors in exploratory factor analysis, there was one major difference. In particular, factor analysis by gender gave consistent results with the second model that female and male students' factor loadings for the identity items that measure student's perception of how others see them as a physics person showed differences in terms of the alignment with the 3 factors given in Table 6-2. Q18 –"My TA/Instructor sees me as a physics person" factored with

Identity/Interest for male students but with Self-efficacy/Belonging for female students. The primary results for Q17 –"My friends see me as a physics person" showed a similar pattern. Furthermore, the results are also consistent with our model 3 in which we proposed that there might be some differences in how students see themselves as a physics person (internal identity) and how they think others see them as a physics person (external identity). This mismatch between the internal and external identity was more salient among female students as discussed above. In other words, for female students, internal identity was aligned with the Interest/Identity component while the external identity such as recognition by peers or TA/Instructor's recognition showed moderate or strong correlation under the Self-efficacy/Belonging component, respectively. Again, the causal direction is unclear: are female students in calculus-based introductory physics course interpreting how others see them through a self-efficacy belief or are they more likely than male students to be given messages of exclusion based on stereotypes? There are reasons to think both directions are plausible, but additional data are required to tease these apart and estimate the extent to which each direction contributes to one of our central findings discussed here.

Given that achievements and experiences in gateway courses, such as physics, are an important part of students' STEM identities [100], supporting female students' physics identity formation and bolstering their positive attitudes towards physics may help sustain women's participation and advancement in STEM disciplines. In the next section, we discuss the possible sources and implications of physics identity differences across gender so that education researchers, physics departments, policy makers, and instructors can develop and implement more inclusive teaching approaches to address these gender discrepancies in physics classrooms.

6.5.2 Implications for physics instruction

The gender differences in students' physics identities can be attributed to several factors: unsupportive classroom experiences of female students in classes in which they are underrepresented, gender stereotypes, the masculine culture of physics, and the lack of female role models [88]. We found that for female students, their peers' and TA/instructors' perceptions, which contribute to their identity formation, relate to their sense of belonging and self-efficacy. Positive or negative classroom experiences and classroom interactions with peers and course instructors can enhance or reduce female students' motivation to engage, as well as their course performance, persistence in the major and their identity as STEM majors [100].

Stereotypes often undermine individuals' self-efficacy by increasing their anxiety (which can reduce the level of engagement and time on task) and can cause them to perform worse in test-taking situations than they actually can due to the limitations on working memory (part of the working memory is robbed by the anxiety when solving problems) [101,102]. Research suggests that women who feel stereotyped have lower science-related aspirations [103]. Students might rely on such stereotypes (whether they are positive or negative) especially in ambiguous situations. In the first semester of university, students are not familiar with college course sequences, and they need assurance for their performance as well as their choice of major [88]. In particular, for under-represented students who experience stereotype threat, belonging uncertainties in a particular setting may become even more heightened in ambiguous situations [103,104]. Our findings underline the important connection between students' self-efficacy and being identified by others as a physics person. Due to the negative stereotypes discussed, female students' self-efficacy may not only decrease and further negatively impact their identity in physics regardless of female students' actual physics skills.

We found that students' interest in physics was strongly connected to perception of oneself as a physics person, for both female and male students. This result highlights the importance of engaging both gender groups' interests into instructional strategies. For example, using relevant examples, incorporating classroom activities that tap into students' curiosity, and tailoring the curriculum towards students' values and learning goals can be stepping stones to enhance the way that female and male students identify themselves with physics. Students are more likely to develop positive physics identities if the topics they learn resonate with them. In this sense, our results show that using students' interest as a resource can be one way to improve students' learning and to support their physics identity.

One of the other highlights of our results is that getting positive recognition by the course instructor and graduate teaching assistant as someone who is good at physics is particularly important for women and relate to their sense of belonging and self-efficacy. While feelings of being recognized can boost female students' feelings of belonging in a physics classroom, lack of recognition from TAs and instructors can be especially detrimental to women's sense of belonging and self-efficacy. As an example, women, as a gender minority in physics courses, might experience constant fear of confirming the stereotypes about their gender group in freshman courses and are more likely to think that their instructors or peers doubt their skills in the physics classroom when they ask them how to solve a problem and the instructor/TA respond by saying that the problem is trivial, obvious or easy [105]. The ambiguity of what is meant by the instructor/TA activates the stereotype threat and female students are likely to interpret such a statement as being told they are not smart enough to do those types of problems and be in the discipline. Negative societal stereotypes and generalizations about who can succeed in physics and other STEM disciplines can lower women's sense of belonging and undermine their experiences, so that they may not form the necessary STEM identities as often as men do to excel. Therefore, gender-based generalizations have implications regarding the mechanism of identity formation in physics and the development of teaching strategies to support students' identities in physics classrooms.

Additionally, our results revealed that female students possess negative beliefs about their TA/instructors' recognition in physics courses. Although instructors might intend to treat all students equally, they might carry subtle biases about students' competence based upon their gender, race or ethnicity, and they can implicitly or explicitly manifest these biases when interacting with students. For example, these preconceptions regarding students' competence mostly favor male students over female students [106]. For instance, instructors more often praise male students for correct answers, while they overlook female students' displays of correct knowledge [107].

Moreover, when instructors praise students' performance as a sign of their intelligence rather than their effort, students can adopt fixed mindset views (i.e., intelligence is a fixed trait and innate intelligence that only men have is necessary to become successful in physics) [59]. A recent study finds that STEM faculty, who endorse fixed mindset beliefs (i.e., they believe some students have innate talent and others do not), have larger achievement gaps and inspire less students to succeed in their classes [108]. In general, being a genius or exceptionally smart is attributed to boys [109], and starting from early ages, boys and girls are externally exposed to these fixed intelligence views in our society and begin to internalize them [100]. One study found that women are underrepresented in fields, such as physics, that associate being "brilliant" with being successful [111]. Furthermore, norms of science curricula also promote the view that science is for high achievers or naturally gifted students [67], and women believe that they must achieve at exceptionally high levels in math and science to be successful in STEM professions [46]. In order to help all students learn physics, it is important that instructors do not convey fixed-mindset views but encourage beliefs that students who work hard using effective approaches can excel in physics.

Peer recognition is also an influential factor that moderates women's classroom experiences. In particular, our research presented here shows that female students' perceived recognition from friends was related to Self-efficacy/Belonging (in a similar way to TA/Instructor recognition although with a smaller factor loading), while there was no such connection for male students. This result highlights the role of peer interaction which may be more related to female students' self-efficacy beliefs and sense of belonging in physics classes.

Some instructional approaches, such as active engagement classrooms, incorporate collaborative group problem solving, which requires frequent peer interaction [16]. Well-designed collaborative classroom activities can increase students' engagement and learning but these activities need to be implemented in such a way that all students can equally benefit [112]. In particular, prior research suggests that evidence-based active engagement methods help all students but increase the gender gap in performance in that they do not help women as much as men (unless explicit efforts are made to mitigate the stereotype threat that is even more salient when people are working in groups as in an active-engagement course) [112]. One study found that female students who were assigned to more gender balanced groups during problem solving sessions demonstrated higher verbal participation and confidence as opposed to those who had been assigned to female minority groups [113]. In general, female students are often assigned to do menial tasks in cooperative activities (e.g., notetaking or data recording) and play less active roles, while male students are more likely assigned to leading roles, such as explaining materials to others [72,73,100]. In such group activities, female students might feel that their opinions are not valued, or they may feel that their contributions to group work are discounted, which can undermine their self-efficacy and sense of belonging. These differential learning experiences between female and male students can accumulate over time and lead to lack of gender parity in students' physics identities and performance.

Research-based classroom interventions are suggested to enhance women's sense of belonging and confidences as they focus on increasing students' persistence and achievements in physics courses [114-116]. For example, the Belonging-Mindset interventions target removing feelings of uncertainty created when students ask themselves: *Do people like me belong here*? [114,115]. These types of interventions acknowledge students' struggle in learning a new subject as normal and inform them that most students have similar difficulties. For this reason, it is important that instructors tailor these short interventions in their classes to dispel students' stereotypical beliefs early in the semester. Since our results presented here suggest that physics instructors/TAs play an important role in female students' sense of belonging and self-efficacy, instructors can use classroom interventions in order to help then recognize that they are not the only ones struggling in learning physics. Implementing these classroom interventions may take some class time but they can have lasting effects in supporting women's STEM identities and creating more diverse learning spaces [114].

6.5.3 Implications for STEM diversity

Students with strong physics identities can have better achievement, engagement and persistence in physics courses, which are foundational to their STEM curricula [34,89]. Positive experiences in the first-year college courses, such as physics, can foster students' interest and retention in STEM programs. Research suggests that the most common dropout points for women from physical science or engineering pathways are during the freshman and sophomore years [117]. Hewitt and Seymour extensively discussed the underlying reasons for first-year college students' decisions to leave STEM fields, a majority of whom were women and under-represented racial and ethnic minorities [87]. They found that the non-inclusive STEM environments, negative classroom experiences (e.g., peers' biased attitudes), and lack of support and encouragement from faculty are some of the primary reasons for switching out of STEM [88]. Because of these types of factors, women can lose their interest in pursuing STEM [118]. Likewise, these experiences can adversely impact female students' identities in physics and in other related STEM domains. Due to insufficient identity formation in physics courses, female students may hesitate to pursue engineering degrees that use advanced levels of physics.

Furthermore, the gender disparity in physics identity can translate into the representation gaps in postgraduate degrees in related STEM programs or vocational choices after graduation [41]. Lack of a sense of belonging can be detrimental in the long-term even for students who choose to continue in STEM fields. Female students, who often experience a lack of sense of belonging from the early years in STEM, may continue to feel less capable than their male counterparts as they proceed in their academic and professional careers. Over time, these feelings might turn into the "imposter phenomenon", where individuals feel that they are not capable enough to deserve their successes, think that they are just "lucky" to be in their positions, or constantly carry feelings that others may eventually perceive them as being a "fraud" [119]. Since women undervalue their skills and possess less strong identities in STEM [103], they may fail to promote their strengths especially in the competitive and male-dominated STEM occupations. Vesturlund et al. found that female students are less likely to compete against male students (as opposed to other women) despite their similar math skills and performance [120]. In our prior study, we also found similar results in that there were large self-efficacy gender gaps between similarly performing female and male students in physics courses [22,23].

Moreover, there are gender differences in students' educational and occupational aspirations depending on their goals and values. When female students make their professional career choices, they put more value on selecting fields that aim to help others and contribute to society. One approach to increasing female students' participation may include reforming STEM curricula to be relevant to both gender's interests and goals [121]. Moreover, based upon our research finding discussed here that shows that recognition by others plays a key role in female students' physics identity, we advocate making STEM occupations and workplaces more supportive for women so that they feel recognized, valued and empowered to succeed in physics.

Considering the possible consequences of identity gaps early in STEM careers, STEM culture in general must change in order to become more welcoming and less chilly for everyone at all levels; otherwise, the diversity issues in physics and other related STEM domains will perpetuate.

6.5.4 Future directions

In this study, we focused on the quantitative aspects of female students' internal and external physics identities in introductory level physics courses. However, examining female students' experiences in STEM disciplines with a qualitative study can expand our current understanding of how motivational factors impact female students' identities, performance and decisions in STEM fields.

As illustrated earlier, in ambiguous situations, women are likely to be victims of stereotype threat and negatively interpret what their physics TA/Instructor thinks about them and whether they see them as a physics person. Our research presented here suggests that TA or/and Instructor's views and communications can be extremely influential particularly for female students' sense of belonging, engagement and their physics identity. However, there might be some variations in students' responses to the survey questions if we had separated TA and instructor recognition as two separate questions. In particular, instructors teaching traditionally in big lecture halls have less opportunity to give students feedback, whereas TAs teach small recitation classes so that they can more frequently interact with each student and get to know them at personal level. Therefore, it would be worth conducting future investigation to test if the students' perception of TAs and instructors' recognition vary and play different roles in female students' sense of belonging and self-efficacy or whether they are similar.

Future research will investigate male and female students' identity in algebra-based physics classes. Unlike calculus-based courses, women are often the majority group in algebra-based courses, often making up 60% or more of the classroom (only algebra-based physics courses are required for students interested in health-related professions, which are more common career goals among women). Since students' goals and aspirations are likely to be different in these classes compared to the calculus-based courses, there might be differences in students' motivational beliefs and attitudes and their relation to their physics identities.

6.6 Chapter references

1. National Student Clearinghouse Research Center (2015). Science and Engineering Degree Attainment: 2004–2014 retrieved from https://nscresearchcenter.org/snapshotreport-degreeattainment15/

2. National Science Board Science and Engineering Indicators. National Science Foundation Report, retrieved from website: https://www.nsf.gov/statistics/2018/nsb20181/digest/sections/preface

3. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016.

4. *Women, Minorities, and Persons with Disabilities in Science and Engineering* (2017), Retrieved from https://nsf.gov/statistics/2017/nsf17310/digest/occupation/women.cfm

5. L. Tsui, Effective strategies to increase diversity in STEM fields: A review of the research literature, The Journal of Negro Education **79**, 555 (2007).

6. S. Eddy and S. Brownell, Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines PRPER **12**, 020106 (2016).

7. A. Traxler, X. Cid, J. Blue and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future PRPER **12**, 020114 (2016).

8. A. Traxler and E. Brewe, Equity investigation of attitudinal shifts in introductory physics, Phys. Rev. ST Phys. Educ. Res. **11**, 020132 (2015).

9. R. Ivie, S. White and R. Chu, Women's and men's career choices in astronomy and astrophysics, Phys. Rev. Phys. Educ. Res. **12**, 020109 (2016).

10. R. Ivie and K. Stowe, Women in Physics, 2000. AIP Report https://eric.ed.gov/?id=ED446559

11. P. V. Engelhardt, in *Getting started in PER*, Reviews in PER Vol. 2, edited by C. Henderson and K. A. Harper (American Association of Physics Teachers, College Park, MD, 2009).

12. R. Henderson, G. Stewart, J. Stewart, L. Michaluk and A. Traxler, Exploring the gender gap in the conceptual survey of electricity and magnetism, Phys. Rev. Phys. Educ. Res. **13**, 020114 (2017).

13. A. Madsen, S. McKagen and E. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? Phys. Rev. ST. Phys. Educ. Res. **9**, 020121 (2013).

14. E. Brewe, V. Sawtelle, L. Kramer, G. O'Brien, I. Rodriguez and P. Pamela, Toward equity through participation in Modeling Instruction in introductory university physics, Phys. Rev. ST. PER **6**, 010106 (2010).

15. J. Day, J. Stang, N. Holmes, D. Kumar and D. Bonn, Gender gaps and gendered action in a first-year physics laboratory, PRPER **12**, 020104 (2016).

16. M. Lorenzo, C. Crouch and E. Mazur, Reducing the gender gap in the physics classroom, Am. J. Phys. **74**, 118 (2006).

17. B. A. Adegoke, Impact of interactive engagement on reducing the gender gap in quantum physics learning outcomes among senior secondary school students, Physics Education **47,** 462 (2012).

18. 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).

19. Z. Y. Kalender, E. Marshman, T. Nokes-Malach, C. Schunn and C. Singh, Motivational characteristics of underrepresented ethnic and racial minority students in introductory physics courses, in the *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.046

20. T. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, in the *Proceedings of the 2017 Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

21. 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, *Proceedings of the 2018 Physics Education Research Conference* (2019) [Washington, DC, August 1-2, 2018], edited by A. Traxler, Y. Cao, and S. Wolf.

22. 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?, *Proceedings of the 2018 Physics Education Research Conference* (2019) [Washington, DC, August 1-2, 2018], edited by A. Traxler, Y. Cao, and S. Wolf.

23. 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? Phys. Rev. Phys. Educ. Res. **14**, 020123 (2018).

24. D. Hammer, Epistemological beliefs in introductory physics, Cognition Instruct. **12**, 151 (1994).

25. E. Redish, J. Saul and R. Steinberg, Student expectations in introductory physics, Am. J. Phys. **66**, 212 (1998).

26. K. Edmondson and J. Novak, The interplay of scientific epistemological views, learning strategies, and attitudes of college students, Journal of Research in Science Teaching **30**, 547 (1993).

27. S. Bates, R. Galloway, C. Lopston and K. Slaughter, How attitudes and beliefs about physics change from high school to faculty, PRPER **7**, 020114 (2011).

28. S. Glynn, P. Brickman, N. Armstrong and G. Taasoobshirazi, Science motivation questionnaire II: Validation with science majors and nonscience majors, Journal of Research in Science Teaching **48**, 1159 (2011).

29. L. McDermott, Oersted Medal Lecture 2001: Physics Education Research—The Key to Student Learning, Am. J. Phys. **69**, 1127 (2001).

30. G. Kortemeyer, Correlations between student discussion behavior, attitudes, and learning, Phys. Rev. ST. PER. **3**, 010101 (2007).

31. J. Nilssen and J. Shemwell, Gender, experience, and self-efficacy in introductory physics, PRPER **12**, 020105 (2016).

32. H. Fencl and K. Scheel, Research and teaching: Engaging students - An Examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course, J. Coll. Sci. Teaching **35**, 20 (2005).

33. C. Lindstrom and M. 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).

34. Z. Hazari, G. Potvin, R. Lock, F. Lung, G. Sonnert and P. Sadler, Factors that affect the physical science career interest of female students: Testing five common hypotheses, Phys. Rev. ST Phys. Educ. Res. **9**, 020115 (2013); Z. Hazari, G. Sonnert, P. Sadler and M. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, Journal of Research in Science Teaching **47**, 978 (2010); Z. Hazari, R. Tai and P. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, Science Education **91**, 847 (2007).

35. V. Sawtelle, E. Brewe and L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1096 (2012).

36. 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).

37. M. Wang and J. Degol, Motivational pathways to STEM Career Choices: Using Expectancy-Value perspective to understand individual and gender differences in STEM fields, Dev. Rev. **33**, 304 (2013).

38. R. Felder, G. Felder, M. Mauney, C. Hamrin and E. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Engineering Education **84**, 151 (1995).

39. E. Lichtenberger and C. George-Jackson, Predicting high school students' interest in majoring in a STEM field: Insight into high school students' postsecondary plans, Journal of Career and Technical Education **28**, 19 (2013).

40. A. Wigfield et al. Motivation to succeed. In Damon, W. and Eisenberg, N (Eds.) *Handbook of Child Psychology Vol. **3** Social Emotional and Personality Development*, New York: Wiley (2006).

41. J. Eccles, Understanding women's educational and occupational choices, Psychology of Women Quarterly **18**, 585 (1994).

42. J. Smith, C. Sansone and P. White, The stereotyped task engagement process: The role of interest and achievement motivation, J. Educ. Psy. **99**, 99 (2007).

43. E. Seymour, Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, Science Education **86**, 79 (2001).

44. A. Diekman, E. Clark. A. Johnston, E. Brown and M. Steinberg, Malleability in communal goals and beliefs influences attraction to STEM careers: Evidence for a goal congruity perspective, J. Pers. Soc. Psychol. **101**, 902 (2011).

45. S. Brainard and L. Carlin, A six-year longitudinal study of undergraduate women in engineering and science, J. Eng. Educ. **87**, 369 (1998).

46. S. Correll, Gender and the career choice process: The role of biased self-assessments, American Journal of Sociology **106**, 1691 (2001).

47. S. Correll, Constraints into preferences: Gender, status, and emerging career aspirations, American Sociological Review **69**, 93 (2004).

48. A. Strenta, R. Elliot, R. Adair, M. Matier and J. Scott, Choosing and leaving science in highly selective institutions, Res. High. Educ. **35**, 513 (1994).

49. E. Pascarella and P. Terenzini, *How College Affects Students: Findings and Insights From Twenty Years of Research*, San Francisco, CA: Jossey-Bass (1991).

50. S. Cheryan and V. Plaut, Explaining underrepresentation: A theory of precluded interest, Sex Roles **63**, 475 (2010).

51. C. Hill, C. Corbett and A. Rose, Why so few? Women in science, technology, engineering, and mathematics (2010). Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf.http://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2010.tb01066.x/epdf

52. S. Ahlqvist, B. London and L. Rosenthal, Unstable identity compatibility: How gender rejection sensitivity undermines the success of women in science, technology, engineering and mathematics fields, Psychological Science **24**, 1644 (2013).

53. R. M. Deci and E. L. Ryan, Intrinsic and extrinsic motivations: Classic definitions and new directions, Contemporary Educational Psychology **25**, 54 (2000).

54. P. R. Pintrich and D. Schunk, *Motivation in education: Theory, research and applications*, (Englewood Cliffs, NJ, Prentice Hall 1996).

55. A. Elliot and J. M. Harackiewicz, Approach and avoidance achievement goals and intrinsic motivation, Journal of Personality and Social Psychology **80**, 501 (1996).

56. W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, PRPER **2**, 010101 (2006).

57. A. Duckworth, C. Peterson, M. Matthews, and D. Kelly, Grit: Perseverance and passion for long-term goals, Personality Processes and Individual Differences **92**, 1087 (2007).

58. A. Duckworth and P. Quinn, Development and validation of the short grit scale (Grit-S), Journal of Personality Assessment **91**, 166 (2009).

59. C. Dweck, *Mindset: The New Psychology of Success* (Ballentine, New York, 2006); C. Dweck, *Self-theories: Their Role in Motivation, Personality, and Developmen*t (Psychology Press, Philadelphia, 1999).

60. Z. Y. Kalender, E. Marshman, C. Schunn, T. J. Nokes-Malach, and C. Singh, Why female STEM majors don't identify with physics: They don't think others see them that way, PRPER *(under review).*

61. P. Pintrich, A motivational science perspective on the role of student motivation in learning and teaching contexts, Journal of Educational Psychology **95**, 667 (2003).

62. S. Hurtado, N. L. Cabrera, M. H. Lin, L. Arellano and L. L. Espinosa, Diversifying science: Underrepresented student experiences in structured research programs, Res. High Educ. **50**, 189 (2009).

63. J. Stets, P. S. Brenner, P. J. Burke and R. T. Serpe, The science identity and entering a science occupation, Social Science Research **64**, 1 (2017).

64. R. M. Rock, Z. Hazari and G. Potvin, Physics career intentions: The effect of physics identity, math identity, and gender, *Proceedings of the 2012 Physics Education Research Conference* (2012).

65. A. Godwin, G. Potvin, Z. Hazari and R. Lock, Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice, Journal of Engineering Education **105**, 312 (2016).

66. G. Potvin and Z. Hazari, The development and measurement of identity across the physical sciences, *Proceedings of the 2013 Physics Education Research Conference* (2013) doi:10.1119/perc.2013.pr.058

67. H. B. Carlone and A. Johnson, Understanding the science experience of successful women of color: Science identity as an analytical lens, Journal of Research in Science Teaching **44**, 1187 (2007).

68. E. D. Tate and M. C. Linn, How does identity shape the experiences of women of color engineering students?, Journal of Science Education and Technology **14**, 483 (2005).

69. J. P. Gee, Identity as an analytical lens for research in education, Review of Research in Education **25**, 99 (2000).

70. R. Ceglie, Underrepresentation of women of color in the science pipeline: The construction of science identities, Journal of Women and Minorities in Science and Engineering **17**, 271 (2011).

71. L. Archer, J. Moote, B. Francis, J. DeWitt and L. Yeomans, The “exceptional” physics girl: A sociological analysis of multimethod data from young women aged 10-16 to explore gendered patterns of post-16 participation, American Educational Research Journal **54**, 88 (2017).

72. A.T. Danielsson, Exploring woman university physics students ‘doing gender’ and ‘doing physics’, Gender and Education **24**, 25 (2012).

73. A. Gonzalves, Exploring how gender figures the identity trajectories of two doctoral students in observational astrophysics, PRPER **14**, 010146 (2018).

74. J. E. Stets and P. J. Burke, A sociological approach to self and identity, Handbook of Self and Identity, Guilford Press, pp 128-152 (2003).

75. D. Oyserman, K. Elmore and G. Smith, in *Handbook of Self and Identity*, edited by M. R. Leary and J. P. Tangney, Guilford Publication, pp 69-104 (2011).

76. K. L. Tonso, Student engineers and engineer identity: Campus engineer identities as figured world, Cultural Studies of Science Education **1**, 273 (2006).

77. A. Wigfield and J. S. Eccles, Expectancy-Value theory of achievement motivation, Contemporary Educational Psychology **25**, 68 (2000).

78. A. Bandura, *Self-efficacy: The Exercise of Control* (W. H. Freedman, New York, NY, 1997).

79. A. M. Flowers III and R. Banda, Cultivating science identity through sources of self-efficacy, Journal of Multicultural Education **10**, 405 (2016).

80. P. Vincent-Ruz and C. Schunn, The nature of science identity and its role as the driver of student choices, International Journal of STEM Education **5**, 48 (2018).

81. J. M. Harackiewicz, E. K. Barron, M. J. Tauer and A. J. Elliot, Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation, Journal of Educational Psychology **94**, 562 (2002).

82. S. Hidi, Interest: A unique motivational variable, Educational Research Review **1**, 69 (2006).

83. S. Hidi and K.A. Renninger, The four-phase model of interest development, Journal of Educational Psychology **41**, 111 (2006).

84. R. Masika and J. Jones, Building student belonging and engagement: Insights into higher education students' experiences of participating and learning together, Teaching In Higher Education **21**, 138 (2016).

85. C. Goodenow, Classroom belonging among early adolescent students: Relationships to motivation and achievement, The Journal of Early Adolescence **13**, 21 (1993).

86. A. W. Astin, *What Matters In College: Four Critical Years Revisited*. San Francisco, CA: Jossey-Bass (1993).

87. N.M. Hewitt and E. Seymour, A long, discouraging climb, ASEE Prism **1**, 24 (1992).

88. E. Seymour and N.M. Hewitt, Talking About Leaving: Why Undergraduates Leave the Sciences. Boulder, CO: Westview Press (1997).

89. K. M. Whitcomb, Z. Y. Kalender, C. Schunn, T. J. Nokes-Malach and C. Singh, Laying a foundation for success in engineering coursework: A predictive model of the curricula, Journal of Engineering Education, under review.

90. http://www.activationlab.org/tools/

91. W.K. Adams, K.K. Perkins, N.S. Podolefsky, M. Dubson, N.D. Finkelstein and C.E. Wieman, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006) doi:10.1103/PhysRevSTPER.2.010101; B. Zwickl, T. Hirokawa, N. Finkelstein and H. Lewandowski, Phys. Rev. ST Phys. Educ. Res. **10**, 010120 (2014). doi:10.1103/PhysRevSTPER.10.010120.

92. Activation Lab. Tools: measures and data collection instruments. 2017.Available from: http://www.activationlab.org/tools/

93. J. Schell and B. Lukoff. Peer instruction self-efficacy instrument. Developed at Harvard University. Unpublished Instrument coetta. 2010

94. S. Glynn, P. Brickman, N. Armstrong and G. Taasoobshirazi, Journal of Research in Science Teaching **48**, 1159 (2011).

95. https://survey.perts.net/share/dlmooc

96. R. Likert, A technique for the measurement of attitudes, Archives of Psychology **140**, 1 (1932).

97. L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

98. J. Cohen, J. Cohen, *Statistical Power Analysis for the Behavioral Sciences.* L. Erlbaum Associates (1988).

99. J. S. Eccles, A. Wigfield and U. Schiefele, Motivation to succeed, in *Handbook of Child Psychology: Social, Emotional, and Personality Development*, edited by W. Damon and N. Eisenberg (Hoboken, NJ, US: John Wiley & Sons Inc., 1998).

100. R. Felder, G. Felder, M. Mauney, C. Hamrin and E. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, J. Engineering Education **84**, 151 (1995).

101. C. Steele and J. Aronson, Stereotype threat and the intellectual test performance of African Americans, Attitudes and Social Cognition **69**, 797 (1995).

102. A. Maries, N. I. Karim and C. Singh, Is agreeing with a gender stereotype correlated with the performance of female students in introductory physics? Phys. Rev. Phys. Educ. Res. **14**, 020119 (2018).

103. J. L. Cundiff, T. K. Vescio, E. Loken and L. Lo, Do gender-science stereotypes predict science identification and science aspirations among undergraduate science majors? Social Psychological Education **16**, 541 (2013).

104. G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women's performance in physics, International Journal of Science Education **35**, 3050 (2013).

105. S. Cheryan, S. Ziegler, A. Montoya and L. Jian, Why are some STEM fields more gender balanced than others? Psychological Bulletin (2016) doi: http://dx.doi.org/10.1037/bul0000052

106. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham and J. Handelsman, Science faculty's subtle gender biases favor male students, Proceedings of the National Academy of Sciences **109**, 16474 (2012).

107. S. Delamont, *Women's place in education.* Brookfield, MA: Avebury Publishers (1996).

108. E. A. Canning, K. Muenks, D. J. Green and M. C. Murphy, STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes, Science Advances **5**, eaau4734 (2019).

109. S. Upson and L. F. Friedman, Where are the female geniuses?, Scientific American Mind **23**, 63 (2012).

110. L. Bian, S. J. Leslie and A. Chimpian, Gender stereotypes about intellectual ability emerge early and influence children's interest, Science **355**, 389 (2017).

111. S. J. Leslie, A. Chimpian, M. Meyer and E. Freeland, Women are underrepresented in disciplines that emphasize brilliance as the key to success, Science **347**, 262 (2015).

112. N. I. Karim, A. Maries and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics, European Journal of Physics **39**, 025701 (2018).

113. N. Dasgupta, M. M. Scircle and M. Hunsinger, Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering, PNAS **112**, 4988 (2015).

114. D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They are magic, Review of Educational Research **81**, 267 (2011).

115. G.M. Walton, C. Logel, J. M. Peach, S. J. Spencer and M. P. Zanna, Two brief interventions to mitigate a "chilly climate" transform women's experience, relationship, and achievement in engineering, Journal of Educational Psychology **107**, 468 (2015).

116. A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock and G. L. Cohen, Reducing the gender achievement gap in college science: A classroom study of values affirmation, Science **330**, 1234 (2010).

117. I. Goodman, C. Cunningham and C. Lachapelle, The Women's Experience in College Engineering. Retrieved from https://files.eric.ed.gov/fulltext/ED507395.pdf (2002).

118. A. W. Astin, *What Matters In College: Four Critical Years Revisited*. San Francisco, CA: Jossey-Bass (1993).

119. P.R. Clance and S. A. Imes, The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention **15**, 241 (1978).

120. L. Vesterlund and M. Niederle, Do women shy away from competition? Do men compete too much?, Quarterly Journal of Economics **122**, 1067 (2007).

121. J. Brotman and F. M. Moore, Girls and science: A review of four themes in the science education literature, Journal of Research in Science Teaching **45**, 971 (2008).

6.7 Appendix D

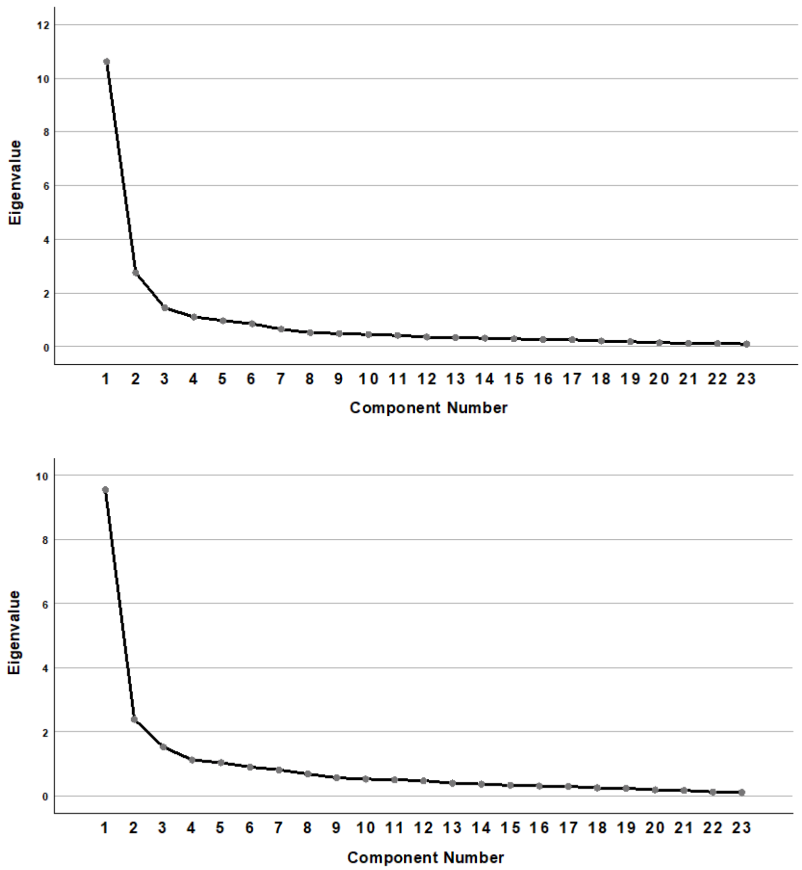

Figure 6-3 The Scree plots for the PCA analyses for female students (top) and male students (bottom).

7.0 Maybe others could but I surely can't: a framework for unpacking gendered mindset views in physics

7.1 Introduction

Over the last several decades, the higher education physics community in the US has been concerned with the low number of women in the field. Although there has been some increase in women's college enrollment in physics courses in the past several decades, the percentage of physics bachelor's degrees earned by women is still only 20% [1-4]. Moreover, several studies have documented performance gaps between female and male students in their course grades [5] and conceptual physics tests [6-9]. Several reasons for these representation and performance gaps in physics have been proposed including: prior knowledge differences [10], gender biases in test questions [11], stereotype threat and ensuing anxiety affecting females students' learning and performance [9,12], and motivational differences that can influence learning, test performance, and career choices [13-18].

With regard to motivational differences, the prior literature has identified a number of motivational factors that are central to effectively engaging students in learning [19-23]. Students with positive beliefs and attitudes about the learning domain (e.g., physics) actively participate in learning activities, more often seek help from others, have more interest in aptitude development in that domain despite facing setbacks, and frequently use self-regulated study strategies [22,24,25]. Within the broad field of motivational research, one attitudinal construct which has received a lot of attention and support as central to learning outcomes is students' mindset about intelligence [26]. Dweck and colleagues' beliefs of intelligence theory conceptualizes students as having one of two types of intelligence mindsets: a fixed mindset (i.e., intelligence is innate and unchangeable, you either have it or not) and a growth mindset (i.e., intelligence is malleable and can be developed through effort). Even though numerous studies have shown positive links between a growth mindset and academic achievement [26-28], including studies showing that brief interventions that change student's mindsets can have positive outcomes even years later [29], there has not been much focus on students' mindset views within the context of physics [13,30,31], and there are reasons to suspect that students might have different views about physics than about general academics or even other STEM disciplines, e.g., biology and chemistry, as described below.

We investigated students' intelligence mindset views about learning physics. Specifically, for students enrolled in introductory level calculus-based physics courses, we examined female and male students' mindset views about physics. In addition to societal stereotypes of who belongs and can succeed in physics, female students are underrepresented (roughly 30% of students) in these calculus-based physics courses. Therefore, the differences by gender in intelligence mindset views, which may arise more likely from environmental input, are particularly important to address in order to improve female student participation, retention and performance.

7.2 Background

7.2.1 Intelligence mindset theory

Carol Dweck, the pioneer researcher of intelligence mindset theory, has investigated the link between students' mindset and their learning behaviors and academic outcomes [26]. When individuals have a fixed mindset (also called an entity theory of intelligence), they see intelligence as an innate and unchangeable quality and one's potential in academic areas as being bounded by certain limits. By contrast, when individuals have a growth mindset (also called an incremental theory of intelligence), they believe that intelligence can be cultivated with practice and effort. In the original conception, like in many other binary motivational theories (e.g., mastery vs. performance goal orientations [32-34]), students were thought to vary along a single continuum, from strongly growth-mindset oriented to strongly fixed-mindset oriented. We first review the literature from that perspective, and then present a new conceptualization that is more multi-dimensional and includes dimensions that allow for differentiation between an individual's intelligence mindset pertaining to self vs. others.

7.2.2 Different interpretations of effort

The intelligence mindset theory suggests that students can make different types of associations with regard to effort needed to succeed in a specific task or course. In particular, depending on the mindset view (growth versus fixed), students might view effort and ability as positively related or negatively related [35,36]. Students with the growth mindset generally consider the time and effort spent at a task as an opportunity for growth and to stretch their ability. Since students with the growth mindset positively associate effort and skill development, they are also more likely to enjoy tackling difficult problems. The fixed mindset view, on the other hand, often inversely links effort and ability so that a high level of effort spent at a task is evidence for low intelligence or lack of natural talent.

The type of mindset can have a significant influence on students' engagement with an activity or task choice [26,30,37]. For instance, a fixed mindset can lead students to interpret making mistakes as a lack of talent and those students might disengage, reject opportunities to learn, and prematurely give up on challenging tasks [38]. In other words, students with a fixed mindset view are more likely to have a higher level of anxiety when engaging with learning and may be more reluctant to engage with challenging problems. Other students, who believe intellectual ability can grow with effort, may be more willing to engage in challenging tasks and be less afraid of making mistakes; instead they may view the mistakes as opportunities for learning and to do better in the activities in the future and on exams [39].

7.2.3 Student intelligence mindset and academic outcomes

Growth mindset views are found to be a significant predictor of students' course achievement even after controlling for differences in relevant prior knowledge and other academic aptitude measures [40-43]. For example, one study found that high school students' views about intelligence predicted their performance; those who viewed intelligence as a malleable quality sustained their interest and scored higher in exams than those who viewed intelligence as innate and unchangeable [43]. A recent meta-analysis including approximately 300 studies showed a moderate average effect of mindset on academic outcomes, with much larger average effects in the domain of math and for students who were from low socioeconomic status or were considered academically at risk in those contexts [44].

Moreover, students' goal orientations in the classroom (i.e., mastery goal vs performance goal orientation) have been related to the mindset view that they endorse. Mastery goal orientation refers to a focus on learning and developing a certain set of skills. In contrast, performance goal orientation is referred to as a focus on demonstrating competence via performance [32], and it involves behaviors focused on trying to outperform others or avoiding tasks that can show lack of ability [32-34]. When students have a growth mindset, they are more likely to adopt mastery goal orientations and less likely to focus on outperforming others [32]. Also, students with a growth mindset are less likely to have ability-based "helpless" attitudes [43].

In order to promote a growth mindset in academic learning, researchers have designed and implemented social-psychological classroom interventions [44-48]. The key element in these classroom interventions was acknowledging the normality and necessity of struggle to expand knowledge and skills. While these interventions aimed to help students adopt a growth mindset view, they also bolstered students' sense of belonging and self-confidence, especially for minority student groups [44]. In our institution, we have implemented a sense of belonging and mindset intervention in calculus-based physics courses, and initial results showed some promising outcomes in terms of eliminating the performance-based gender gap [49].

7.2.4 Instructor intelligence mindset and academic outcomes

Interestingly, instructors' mindsets have been found to influence student mindsets and achievement [33], particularly for students from the underrepresented groups [50]. A recent study found that the racial achievement gap was found to be twice as large in classrooms taught by the STEM faculty who endorsed a fixed mindset view than in classrooms taught by faculty who held and used a growth mindset in their teaching approaches [51]. Faculty with a fixed mindset view were found to often emphasized being smart or naturally gifted during the instruction while the faculty with a growth mindset view highlighted the significant contribution of practice and effort to students' classroom success [51].

In addition, unequal college preparation (e.g., due to differential out of class opportunities due to being from low socioeconomic background or parents not completing higher education, K-

12 education in schools with poor infrastructure, inadequate access to well-qualified teachers and pervasive discipline issues, lack of access to Advanced Placement courses or practice opportunities for SAT and ACT exams) create initial performance gaps [52] that are perpetuated if the instructors do not intervene with supportive teaching approaches. Instructors or college admission committees with a fixed mindset view often view these prior gaps due to opportunity differences as an indication of lack of intellectual ability, while those with a growth mindset recognize the variations in standardized test scores as gaps in pre-college preparation (and not lack of growth potential), which can be eliminated by using appropriate study strategies, promoting growth mindset attitudes, and receiving positive encouragement and support from mentors and instructors [31].

7.2.5 Intelligence mindset views in physics

There is a variation across academic areas in the extent to which people believe that foundational innate talent is required, and this variation has sometimes been linked to underrepresentation in some STEM fields [53]. Overall, there are pervasive cultural biases about gender-based intellectual ability; “genius” is a term commonly attributed to men [54] and children adopt these gender-based stereotypical intelligence beliefs at early ages [55]. For example, girls were found to be more likely to avoid activities that are thought to require being “really really smart” and more likely to designate the opposite gender as being “really really smart” [55]. In the context of mathematics, Dweck has found that fixed mindset views harm girls more than boys [27]. Likewise, women have been found to shy away from the disciplines that are thought to necessitate being brilliant or being “gifted [27].

Many people assume that “physics” in particular requires an innate talent or being “gifted” in order to be successful [53]. In fact, physics is one of the most extreme fields pertaining to stereotypes along this dimension [53]. The negative stereotypes and gendered biases about who can succeed in the discipline, along with the masculine culture of physics [56,57], can potentially produce a fixed mindset in the general physics community, which can mostly harm the stereotyped groups, such as women. One research study found that the fixed mindset is commonly adopted among graduate admission committee members in physics departments that tend to exclude women and racial minorities [57]. Therefore, the discipline of physics might suffer from low gender diversity at least partly due to fixed mindset views in the field which end up pushing women away from the field.

Although the literature broadly suggests that students' mindset view is an important factor that helps students succeed in reaching their academic goals—especially in terms of long-term learning outcomes [26], physics education research has rarely focused on students' mindset views. On the other hand, because of the general expectations of the need for "brilliance" in physics, it is likely to be especially important. Indeed, in a recent study of students' general and field-specific (i.e., physics) mindset views across gender in both algebra-based and calculus-based physics courses, we have found gender differences in physics-specific mindset but not in students' general academic mindset [13].

In thinking about student mindsets, it may be important to consider the nature of the student population. At the college level, some students, typically those enrolled in introductory calculus-based physics, are following a physical science or engineering major, having experienced past success in physics, chemistry, and mathematics coursework. Other students in college, typically those enrolled in introductory algebra-based physics courses, are following a life science track or are pursuing a health-related career that requires introductory physics, and usually either have avoided physics in the past or did not experience great success or did not enjoy the experience.

The latter group is often predominantly female, whereas the former group is often predominantly male [13].

Little et al. recently conducted a qualitative study to understand students' mindset views and responses to challenges in physics courses for life science majors [30]. In their research, students identified themselves as being "bad" at physics because it did not come to them naturally, but nonetheless sought help and worked hard to become better at it [30]. However, very little is known about calculus-based physics courses for physical science and engineering majors, who might have different mindset views towards learning physics. Therefore, more study is needed to understand the nature of physics mindset views during college level physics instruction, particularly in the courses in which women are underrepresented.

7.2.6 Unpacking mindset—which subdimensions are critical?

As noted previously, prior research on mindset has treated it as a single dimension with growth and fixed mindsets as the two endpoints. However, this is not the only logical possibility, and other motivational research areas have discovered that multidimensional perspectives better account for student variation than views that assume a single-dimension of opposing perspectives. Most saliently, in the closely-linked achievement goals research area, researchers discovered that students could have both mastery and performance goals at the same time [34]. For mindset, students could believe that both effort and ability matter, e.g., they could believe that some students are missing critical capacities, but even those students who have a foundational capacity still need to exert substantial effort to develop their physics capabilities.

Moreover, stemming from stereotypes about who is brilliant and can be successful in physics, it is possible that students make distinctions about themselves versus others. For example, a student from a demographic that is underrepresented in physics might believe that some people could develop abilities through effort but they themselves could not. Conversely, a student from a demographic over-represented in physics might believe that they are gifted in physics, but most others are not. Although prior measures of intelligence mindsets have included questions that refer to various combinations of abilities, effort, self, and others, prior researchers have not systematically analyzed these data to examine whether these dimensions are separable rather than just one overall growth versus fixed mindset dimension. Note that just because prior studies have reported acceptable levels of measure reliability (i.e., Cronbach's alpha), that alone is not evidence of unidimensionality [58].

In the mindset study presented here, we created different types of mindset survey items that assess students' beliefs about having natural talent (i.e., ability) or doing hard work (i.e., effort). Moreover, the questions are designed to determine students' mindset viewpoints about themselves (me/I) or the general population (others). The mindset questions were conceptually grouped into four categories (see Figure 7-1). As detailed in the Methods section, we used multi-dimensional scaling to determine whether the students also conceptually separated the items into those four categories based upon patterns of similarity/dissimilarity in responses to individual survey questions.

A major focus was on uncovering which dimensions of physics mindset of students represented by different quadrants in Figure 7-1 are most closely connected to course grades in introductory calculus-based physics. Further, we seek to understand whether gender differences in mindset are larger for a particular mindset dimension, and whether those mindset differences could account for gender differences in course grades. In particular, it may be the beliefs about the self that are most relevant to learning outcomes in physics since those would play a key role in determining how students react in response to challenges/difficulty in the course. Further, endorsement of a fixed mindset about the self is likely to be most indicative of non-productive reactions to difficulty out of all responses to survey questions probing different dimensions of mindset.

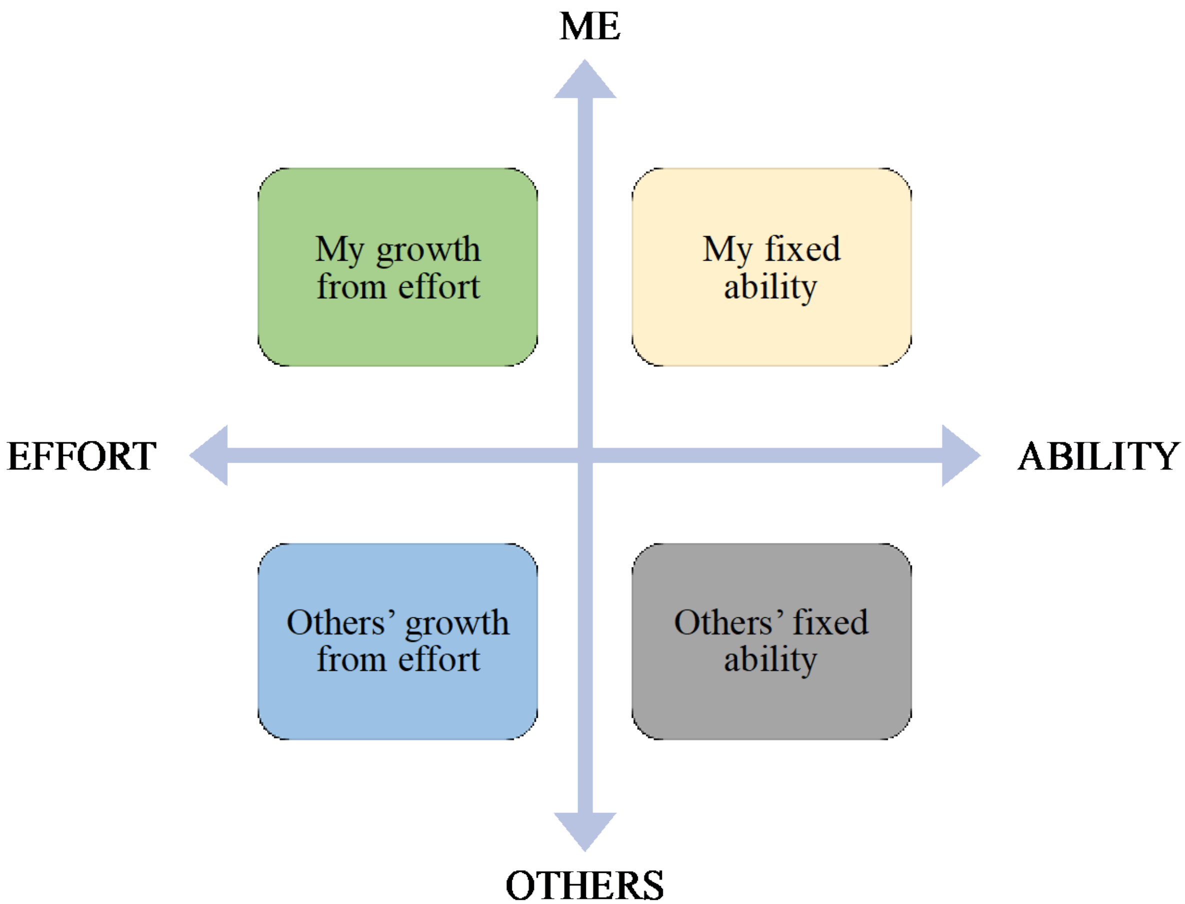

Figure 7-1 The representation of four mindset groups in two dimensions.

7.3 Methodology

7.3.1 Participants and course content

A total of 553 students across four sections of introductory calculus-based Physics 1 completed a mindset survey. Students who enrolled in these courses were mostly physical science or engineering majors. Students' demographic information was obtained via university records as de-identified but linkable data (i.e., demographic and survey data were linked via a research id created by an honest broker). Female students comprised 33% of the cohort, mirroring female enrollment rates across years in this course. In terms of ethnic/racial distribution, the cohort consists of 79% White, 10% Asian, 2% Black, 4% Hispanic, 4% Multi-Racial, and 1% Others (i.e., American Indian, Unknown etc.). The mean student age at the time of completing the survey was 18.7 (SD=1.4), since this course was typically taken by first year students and the majority of the students in our institution attend university immediately after high school.

There was some small variation in the pedagogy across the four different instructors, but they predominately used traditional lecture-based instruction. This course covers introductory level physics subjects such as kinematics, forces, energy and work, rotational motion, and gravitation.

7.3.2 Assessments

7.3.2.1 Development and administration of the physics intelligence mindset survey

A physics intelligence mindset survey was given as part of a larger attitude assessment tool [13,15] and adapted from validated surveys [26,59]. We administered the survey during the recitations. The Graduate Teaching Assistants handed out the survey at the last recitation of the semester (approximately 1-2 weeks before students take final exam) and students marked their responses on the scantron sheet. Completing the survey generally took 3-4 minutes for students.

The survey items, shown in Table 7-1, involved a four-point Likert scale of Strongly Disagree to Strongly Agree, recoded as 1 to 4, with questions 2, 4, 5, and 7 reverse coded. Higher scores indicate a malleable (growth) mindset whereas a low score corresponds to a more fixed mindset in physics learning. The survey was based upon the existing validated instruments [26] and then subjected to additional validation work for the current study context [13]. The mindset survey had seven questions. The Cronbach's α for measuring the internal consistency between the mindset questions was found to be reasonable with the value of 0.77 [60,61]. We also conducted interviews with 12 undergraduate and 3 graduate students to validate our survey so that students interpret the questions in the expected way.

Using the lavaan package in R [62], we performed a Confirmatory Factor Analysis (CFA) to test the item strengths of the physics mindset questions in the current study data. Commonly reported CFA fit parameters are: the Comparative Fit Index (CFI), which compares the fit of the proposed model to the null model; the Tucker-Lewis Index (TLI), which compares the fit of the proposed model to the null model but also taking into account the complexity of the proposed model (such as the correlations between the variables); Standardized Root Mean Square Residual (SRMR), which is the standardized difference between the observed correlation and the predicted correlation; and Root Mean Square Error of Approximation (RMSEA), which is the absolute fit of the model to the data taking into account the amount of data available. CFI > 0.95, TLI >0.95, SRMR < 0.08 and RMSEA ≤ 0.08 are considered acceptable fit parameters. The overall model fit parameters were all acceptable, with CFI = 0.98, TLI = 0.97, SRMR = 0.03, and RMSEA = 0.06 as were the factor loadings of each item (see Table 7-1).

Table 7-1 CFA results for physics intelligence mindset questions. Mindset questions in which "I" is not mentioned explicitly are classified as "others".

Group	Survey Item	R-square	Beta	P value
Others' growth	**Q1.** Anyone can become good at solving physics problems through hard work.	0.36	0.60	< 0.001
Others' ability	**Q2.** Only a few specially qualified people are capable of really understanding physics.	0.27	0.52	< 0.001
Others' growth	**Q3.** No matter who you are, you can change your intelligence in physics quite a lot.	0.39	0.62	< 0.001
Others' ability	**Q4.** To really excel in physics, a person needs to have a natural ability in physics	0.41	0.64	< 0.001
My ability	**Q5.** If I really have to struggle to solve physics problems, that means I'm just not a physics person.	0.32	0.57	< 0.001
My growth	**Q6.** If I spend a lot of time working on difficult physics problems, I can develop my intelligence in physics.	0.45	0.67	< 0.001
My ability	**Q7.** If I make mistakes on physics assignments and exams, I think that maybe I'm just not smart enough to excel in physics.	0.19	0.44	< 0.001

7.3.2.2 Prior academic performance

From university records, we also obtained pre-college test scores that are generally good predictors of academic performance in the first-year college STEM courses: SAT Math scores (200-800) and High School grade point average (HS GPA, 1-5 scale when including advance placement courses). In this data, students' SAT Math scores ranged from a minimum score of 420 to a maximum score of 800.

7.3.2.3 Course grade

Institutional data also provided us with each student's course grade at the end of the semester, which we used as the students' learning outcome for the current study. The final course grade was largely composed of students' midterm and final exam scores. The rest of the grade was made up of weekly homework, students' class participation, concept quizzes, attendance, and quizzes given during recitation. The policy regarding the letter grading and corresponding grade points at this university are given in Table 7-2.

Table 7-2 Course letter grades and corresponding grade points out of 4.

	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A/A+
Grade Point	0	0.75	1.00	1.25	1.75	2.00	2.25	2.75	3.00	3.25	3.75	4
Definitions	Failure					Minimum level to graduate			Adequate level to graduate			Superior attainment

7.3.3 Analysis

Grouping Mindset Questions. We hypothesized that there are four different groups for physics mindset items based on the referent (Me versus Others) and context (Ability versus Effort) of the question. For instance, item 6 (Q6)— "If I spend a lot of time working on difficult physics problems, I can develop my intelligence in physics" was placed in the "My growth" group, while Q1—"Anyone can become good at solving physics problems through hard work" was placed in the "Others' growth" mindset group. The mindset groups for each question are given in Table 7-1. While "My growth" has only 1 item, the other three groups include two items. During the validation process in the interviews, students also gave explanation for their answers to the mindset survey items that do not have “I”, as pertaining to other individual’s ability or effort. After we grouped the related mindset items, the average score was obtained for each of the four mindset groups. For instance, a student who had a score of 3 for Q1 and score of 4 for Q3 was given an average score of 3.5 for the “Others’ growth” mindset group.

To validate the separation of the survey into these four groups and the conceptual relationships between them, we analyzed these means using Multidimensional Scaling (MDS) [63], testing 1–3 dimensional solutions, overall and separate by gender to ensure consistent structure. MDS produces a map of constructs based on the pattern of intercorrelations among the measures, in which distance among items is inversely proportional to the correlation strength between items. We used the MASS (isoMDS) package in R to perform the MDS. MDS results are evaluated using a stress value that illustrates the goodness of the fit [63]. A stress value closer to zero indicates a better fit of the model, and common standards are: 0.200 poor, 0.100 fair, 0.050 good, 0.025 excellent and 0.00 perfect [63].

Gender Differences. We initially performed a MANOVA to test for significant mean differences across the four physics mindset groups between female vs. male students, where gender was the independent variable (IV) and the four physics mindset groups were the dependent variables (DVs). Furthermore, independent *t*-tests between two gender groups were conducted on students’ SAT Math, High School (HS) GPA, and Physics 1 grades. We used an alpha level (probability of rejecting the non-significant result) of 0.05 for all statistical tests.

Predicting Physics 1 Course Grade. We performed simple correlation and then several multiple regression analyses in order to uncover the best model for predicting students’ Physics 1 grades (dependent variable). The first regression model predicted students’ Physics 1 grade using only prior academic performance and demographic variables: SAT Math, HS GPA, and students' gender. In the second model, we added the "My ability" mindset group, the strongest correlate of grades. In the final model, we added all the mindset groups as predictors as a robustness test. Standardized regression coefficients (as an effect size estimate) and p-values (as a statistical significance measure) for independent variables in all three models were calculated.

7.4 Results

7.4.1 Multi-dimensional scaling

The two-dimensional model produced the best fit to the data based on the stress values. The two-dimensional model converged to a near perfect fit, with a stress value of ~ 0.00 (2×10^{-14}). By contrast, the one-dimensional model only converged to 0.04, indicating a good but not excellent fit. Since the two-dimension fit was essentially perfect, there was no advantage to adding a third dimension [62]. Analyzing the data separately by gender produced similar models. The results from the two-dimensional model for all students are presented here (see Figure 7-2). The four groups cleanly fell into their own quadrants as expected by referent (Y-axis) and context (X-axis).

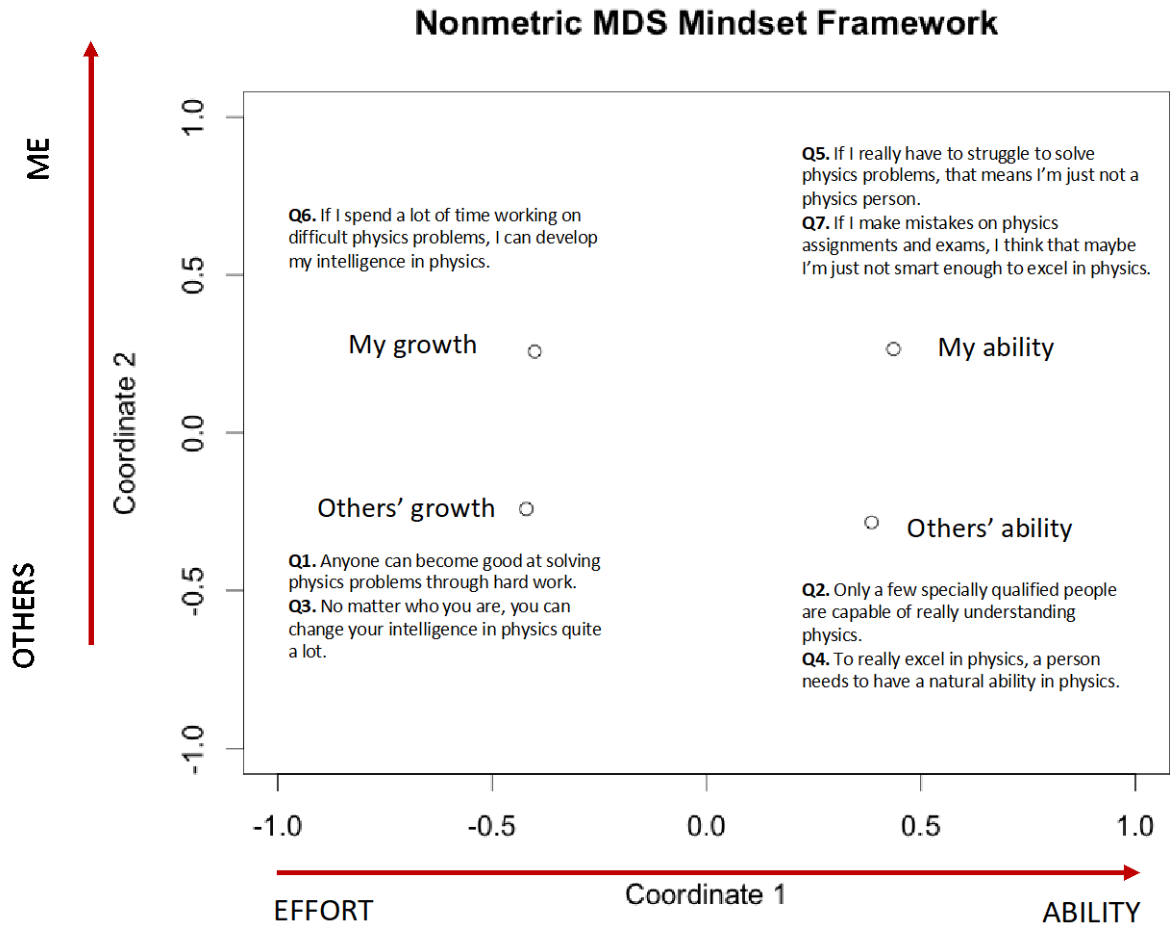

Figure 7-2 The two-dimensional MDS results for the four mindset question groups, along with the questions forming each group. A label to each dimension is added after the fact based on interpretation of the data.

7.4.2 Gender differences in mindsets, prior academics, and physics grade

In an initial MANOVA we examined gender as the independent variable (IV) and mindset views as the dependent variables (DV). It showed a significant multivariate effect of gender: Male students had significantly higher scores in overall mindset (F (1,552) = 7.47, $p < 0.001$; see Table 7-3). As a follow-up analysis, we performed t-tests between female and male students for each of the mindset groups (see Table 7-3). All of the mindset groups showed statistically significant gender differences favoring male students. The largest gender gap was for the "My ability" group, which indicates that female students are less likely to consider themselves as having an "ability" in physics compared to male students. For context, female students also scored lower on SAT Math but with a smaller effect size (d = 0.23), while they had a higher average overall high school GPA (d = 0.28). Both SAT Math and HS GPA are relevant predictors of physics performance and these constructs has a small but statistically significant effect on the Physics 1 grade with male students performing higher.

Table 7-3 The *t*-test comparisons between female and male students for mindset groups, pre-college academic scores and Physics 1 grade are given on the left side of the thick line. The correlations of each construct with each other are given on the right side of the thick line. A positive Cohen's *d* indicates that male students had higher scores than did female students. * = p < 0.05, ** = p < 0.01, *** = p < 0.001, ns = non-significant.

Measure	Means			Intercorrelations					
	F	M	Cohen's d	My growth	Others' growth	Others' ability	My ability	SAT Math	HS GPA
My growth	3.1	3.2	0.30^{**}	—	—	—	—	—	—
Others' growth	2.9	3.0	0.22^{*}	0.51^{***}	—	—	—	—	—
Others' ability	2.4	2.2	-0.23^{**}	-0.33^{***}	-0.45^{***}	—	—	—	—
My ability	2.4	2.1	-0.46^{***}	-0.37^{***}	-0.32^{***}	0.46^{***}	—	—	—
SAT Math	680	694	0.23^{**}	ns	ns	ns	ns	—	—
HS GPA	4.15	4.02	-0.28^{***}	ns	ns	ns	ns	0.24^{***}	—
Physics 1 Grade	2.45	2.60	0.17^{*}	0.21^{***}	0.17^{***}	-0.20^{***}	-0.35^{***}	0.29^{***}	0.30^{***}

7.4.3 Predicting course grade in physics 1 using mindset groups and prior academic performance

In Table 7-3 (right side of the thick line), we also show the Pearson intercorrelations between the mindset groups, SAT Math score, HS GPA and Physics Grade. The mindset group means were correlated with one another but not so highly as to be redundant measures. Interestingly, none of them were correlated with SAT Math or HS GPA, even though SAT Math and HS GPA were significantly correlated with Physics Grade, with nearly identical strength.

Physics Grade exhibited statistically significant correlations with each of the mindset groups. Three of the correlation values were small, ranging from 0.17 to 0.21. However, the “My ability” group had relatively higher correlation with Physics 1 grade, $r = -0.35$. This correlation value was also slightly higher than the correlation value between Physics 1 grade and students’ pre-college academic scores (SAT Math and HS GPA).

Multiple regression was used to tease apart the independent contributions of the various elements correlated with each other and grade. In Model 1, demographic information (gender) and prior academic skills (SAT Math and HS GPA) were entered as predictors of Physics 1 grade ($F(3,553) = 42.32$, $p < 0.001$; see Table 7-4). All three of the predictors were statistically significant, and HS GPA was the strongest predictor. In this model, male students were more likely to have higher grades in the college level Physics 1 courses than female students (i.e., even after control for students’ pre-college academic scores). Thus, there is some indication that another factor beyond prior academic performance is responsible for the gender difference in grades.

Table 7-4 Three regression models in which the independent variables are gender, SAT Math, HS GPA and mindset groups to explain the dependent variable—Physics 1 grade. B, SE and β correspond to unstandardized coefficient, standard error and standardized coefficient, respectively.

		Model 1			Model 2			Model 3		
Variable	Pearson Correlation	B	SE	β	B	SE	β	B	SE	β
Gender (M=1, F=0)		0.15	0.07	**0.08***	0.07	0.07	0.04	0.07	0.07	0.04
SAT Math	**0.29*****	0.002	0.00	**0.20*****	0.00	0.00	**0.19*****	0.002	0.00	**0.19**
HS GPA	**0.30*****	0.69	0.08	**0.32*****	0.54	0.09	**0.26*****	0.56	0.09	**0.27**
My ability	**-0.35*****				-0.41	0.05	**-0.32*****	-0.34	0.06	**-0.27**
My growth	**0.21*****							0.03	0.06	0.02
Others' growth	**0.17*****							0.06	0.06	0.05
Others' ability	**-0.20*****							-0.06	0.06	-0.05
R^2			0.18			0.25			0.26	

In Model 2, we entered "My ability" as the fourth predictor of Physics 1 grade and produced a better fit to the data (R^2 of 0.25 vs. 0.18; F (4, 437) = 36.90, $p < 0.001$). This predictor was actually the strongest predictor of grade, and the predictive power of HS GPA and SAT Math slightly decreased. Most importantly, the gender effect on grade became non-significant, suggesting that it could be attributed to differences in "My ability" mindsets.

In the final model – Model 3, we added the other three mindset groups (My growth, Others' growth, and Others' ability). The model fit was not better (R^2 of 0.26 vs. 0.25; F (7,429) = 21.55, $p < 0.001$), and the "My ability" mindset group was the only statistically significant predictor among all four mindset groups to explain students' physics course grade. Despite adding multiple predictors that had moderately high correlates with the "My ability" group, the standardized coefficient for the "My ability" only was reduced by a small amount; overall, the parameter estimates of Model 2 were robust.

7.5 General discussion

7.5.1 Summary

We find psychometric evidence that mindset in physics for calculus-based introductory courses can be classified into four sub-dimensions. One sub-dimension (endorsing My fixed ability) is especially important in this regard in that it has the largest gender difference with a moderate level effect size (d=0.46). This mindset group was also the main predictor of performance and fully explains the gender gap in course grade. The fact that female students recognize value in putting effort and working hard to be successful in physics but carry beliefs that natural ability is still required to attain higher-levels of achievement in physics has important implications for diversity issues in physics.

Past studies in the physics education literature show that female students have lower motivational characteristics (e.g., self-efficacy, interest, sense of belonging, identity, etc.) than male students in physics classes [13-18] and these gender differences persist or become larger even at matched performance levels at the end of the course [15,16]. The novel results of our study show a similar gender gap in a different motivational construct—mindset, and how it is related to gender. Female students are more likely to carry beliefs that innate talent is needed to excel in physics above and beyond successfully passing a course by making effort, and that they do not have that special talent. Our findings can inspire physics education researchers to design and implement teaching strategies that cultivate a growth mindset to improve all students' participation and success in learning physics.

Attaining success in physics is most commonly and stereotypically thought to require a natural ability or is associated with being "gifted". These ability-based beliefs about success in physics often disregard the substantial contribution of practice using effective study strategies, spending time on task productively and making effort, having interest in learning physics and using failures as a learning opportunity. Instead, these “brilliance” expectations communicate cues about the perceived rigidity of intelligence, which can only be possessed by a certain group of people. Historically well-known physicists have generally been depicted as being “genius” or some elite group of scientists with innate “abilities”. However, there is little mention of the amount of struggle these successful physicists had while learning new concepts or how many times they had failed before they eventually made a discovery in the field of physics. Without acknowledging the journey behind their accomplishments, the general population in the US regards physicists as naturally talented and “brilliant” men, who do not need to work hard to excel. Modern pop-culture and media also reinforce the stereotypical image of a physicist as being “male” and “brilliant”, such as in the famous TV series “The Big Bang Theory”. Thus, gender-based societal cues [64,65], the male-dominated culture of physics [56], and the strong emphasis on being “brilliant” [53] are factors that can cause women to have beliefs about a fixed mindset when it pertains to them but not necessarily about others —as we have shown in our data, and these factors may dissuade women from pursuing physics-related careers.

Our results offer an important piece of evidence for gendered differentials in students’ mindset where women are more likely to adopt a fixed mindset view, especially in the ability-related mindset categories. Having a fixed mindset has been found to yield negative academic outcomes especially for those who are underrepresented in a particular field and come to college with a less prior preparation and knowledge [39]. With that in mind, underrepresented students with a fixed mindset may attribute their struggle to the qualities that they think they cannot change, which makes them less willing to spend effort when they encounter difficult problems. For instance, female students, who do not enroll in AP physics classes as often as male students do, can readily associate their feelings of struggle at college level physics courses with a lack of ability rather than insufficient pre- college academic preparation. The results in this study also show that female students adopt more fixed mindset in the first-year college physics courses for the self- and ability-related questions and that they also underperform than male students. Since the prior literature did not investigate the relation between mindset and gender in the context of physics, our findings shed light on an additional possible reason for diversity issues in the physics courses and physics overall.

Additionally, female students in physics classes might confirm stereotypical beliefs against their group by espousing a fixed mindset. These two factors—fixed mindset and stereotype threat—can together undermine female students' classroom motivation, create additional anxiety for them, and adversely impact how they can perform in the class assessments. Maries et al. found that women in calculus-based introductory physics who agreed with a gender stereotype statement (according to my own personal beliefs, I expect men to generally perform better in physics than women) at the beginning of a physics course had statistically significantly lower performance than women who disagreed with this statement even though there was no different on pre-assessment at the beginning of the course [9]. In particular, their study suggests that confirming gender-stereotypes can be related to a fixed mindset view that can negatively influence female students' performance [9]. Relatedly, the findings in our study also point out that not adopting a "personal" fixed mindset view regarding oneself is positively linked to course performance. Since there have been many studies that show the importance of growth mindset in academic achievement [26-29,35,36,41- 43] but none of them quantitatively has shown growth mindset's influence on physics learning across gender, our findings can greatly contribute to physics education research on promoting diversity in physics.

7.5.2 Implications for teaching

Researchers in education and psychology have developed social-psychological classroom interventions to address students' concerns in classroom, enhance their learning experiences, and improve student outcomes. These classroom interventions have been implemented in various learning settings for different age groups to boost students' persistence and learning, especially for underrepresented groups [44-48]. One such intervention, named the Belonging and Mindset Intervention, focuses on promoting a growth mindset attitude in academic spaces [46-49]. The main focus of these short classroom activities is to reshape how students respond to challenges. By using these classroom interventions, instructors can help students perceive their struggle as an opportunity to develop skills and construct new knowledge, rather than an indication of low ability or self-failure. Classroom interventions can be useful particularly for freshman students since first-year-college experiences in pillar courses like physics may be a determinant factor for whether students decide to stay in or exit from a STEM program altogether [66]. Because first-year-college students undergo a challenging academic transition, students may feel overwhelmed with their struggle and attribute their difficulties as a proof of not belonging in the classroom. Belonging and mindset interventions send messages to all students that struggling and being overwhelmed with first-year-college courses, such as physics, is a very common experience among many students, and one can overcome difficulties and become an expert by working hard and working smart and embracing their failures as learning opportunities. This is especially a valuable message for the underrepresented students, because they can internalize their struggle as a sign of not fitting in, which can then result in students dropping the course or exiting from a STEM track [56,57].

Interventions such as the Belonging and Mindset intervention may not only shift students' mindset beliefs at individual level but also establish an inclusive and more welcoming classroom culture, so students feel more comfortable to share their ideas or ask for help from their classmates. Moreover, active-engagement classes, in which collaborative work and peer discussion is essential, can benefit from incorporating teaching strategies with a growth mindset view. Such classroom interventions create more productive and respectful classroom culture so students can fully engage at a learning task or class activity rather than having a feeling of anxiety for being disregarded or devalued by peers or the instructor. One investigation found that evidence-based active engagement methods help all students but increased the gender gap in performance in that they do not help women in calculus-based introductory physics courses as much as men. One hypothesis for why this may be the case is that evidence-based active-engagement courses often involve collaborative learning (students working with each other). Without explicit interventions by the instructors, in these collaborative situations, societal stereotypes about physics, e.g., who is capable of excelling in physics may become salient [8] and negatively impact engagement and learning.

When creating an inclusive learning environment that fosters growth mindset, instructors and their teaching approaches are a crucial part of that reform [51]. For instance, many students hesitate to answer promptly asked in-class questions because they worry about appearing unintelligent. In order to eliminate these concerns, which act as barriers in students' participation and learning, instructors need to assure students that making mistakes are part of the learning process and constructing a robust knowledge structure. Relatedly, instructor's recognition of students or lack thereof is also a vital part of the classroom culture. The way instructors endorse student work (e.g., praising intelligence versus praising hard work) can alter students' intelligence mindset view and their learning behaviors [67,68]. Since physics is a subject commonly associated with being naturally gifted, instructors, who praise intelligence or communicate a fixed mindset view, can further reinforce stereotypical inferences about physics and can demotivate minority groups. In particular, instructors' beliefs, that ability is fixed, can mainly disadvantage women and perpetuate inequalities in education. As we found in this study presented here, students' course grade is predicted by their ability-related mindset views, which had an almost equal regression coefficient as students' prior academic scores. When course instructors use and promote a growth mindset in their teaching methods, students can become more engaged despite difficulties, which can enhance their learning in physics courses.

To sum up, physics education researchers including those interested in pedagogical interventions to ensure that all students learn need to direct more attention on the mindset aspect of student learning in their efforts to enhance diversity in physics. Establishing teaching and learning standards with a growth mindset early in the semester, and sustaining it, can yield positive student outcomes such as improving underrepresented groups' retention, learning, and advancements in physics and related STEM fields.

7.6 Future directions

One of the major findings in this study is that students' mindset (growth versus fixed) may depend on the type of mindset question, such as whether the item is effort-related or ability-related and whether the question is about that person ("I") or general ("others"). Although students, in general, value the contribution of effort to be successful in physics, they might think that natural ability in physics remains a determinant of one's success and the extent to which the person excels in physics. This aspect can be especially critical for physics due to societal stereotypes and it can be more detrimental to minority groups' achievements. Since we find that the "My ability" mindset group had the largest effect on students' course grade, more studies are needed to unpack this specific dimension so that it can inform how to improve physics learning for all students and increase diversity in physics.

In this study, we only focused on students' mindset and learning across gender. Although our data show a direct relation between physics learning and specific mindset group in physics, it would be useful to explore why these gender differences in ability-related mindset group appear and how such a gap can impact students' other motivational characteristics, such as self-efficacy or sense of belonging. We also note that causality between mindset and performance has not yet been established in our correlational study, although other intervention work does suggest the relationship is casual. We need to do further work to show that it is the fixed ability about self that is a particularly important driver of performance gaps.

Furthermore, the degree to which students receive recognition from their instructors or graduate teaching assistants can also impact their mindset views. Therefore, future research can focus on the interaction between several other motivational constructs that are important for educational research and improving learning of diverse group of students. Also, collecting data at the beginning of the course about students' mindset can be useful to study how students revise their mindset after they undergo physics instruction. As a final note, testing the generalizability of our results to other populations (other teaching methods, other physics courses or other universities with different demographics) can also be useful to obtain a broader understanding of students' mindset.

7.7 Chapter references

1. National Student Clearinghouse Research Center (2015). Science and Engineering Degree Attainment: 2004–2014 retrieved from https://nscresearchcenter.org/snapshotreport-degreeattainment15/

2. National Science Board Science and Engineering Indicators. National Science Foundation Report, retrieved from website: https://www.nsf.gov/statistics/2018/nsb20181/digest/sections/preface

3. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016

4. *Women, Minorities, and Persons with Disabilities in Science and Engineering* (2017). Retrieved from https://nsf.gov/statistics/2017/nsf17310/digest/occupation/women.cfm

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. 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, *Proceedings of the 2018 Physics Education Research Conference* (2019) [Washington, DC, August 1-2, 2018], edited by A. Traxler, Y. Cao, and S. Wolf.

7. 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).

8. N. I. Karim, A. Maries and C. Singh, Do evidence-based active-engagement courses reduce the gender gap in introductory physics, European Journal of Physics **39**, 025701 (2018).

9. A. Maries, N. I. Karim and C. Singh, Is agreeing with a gender stereotype correlated with the performance of female students in introductory physics? Phys. Rev. Phys. Educ. **14**, 020119 (2018).

10. P. Labudde, W. Herzog, M. P. Neuenschwander, E. Violi and C. Gerber, Girls and physics: Teaching and learning strategies tested by classroom interventions in grade 11, Int. J. Sci. Educ. **22**, 143 (2000).

11. R. Henderson, P. Miller, J. Stewart, A. Traxler and R. Lindell, Item-level gender fairness in the Force and Motion Conceptual Evaluation and the Conceptual Survey of Electricity and Magnetism, Phys. Rev. Phys. Educ. **14**, 020103 (2018).

12. G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women's performance in physics, Int. Journal of Science Education **35**, 3050 (2013).

13. 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).

14. T. J. Nokes-Malach, E. Marshman, Z. Y. Kalender, C. Schunn and C. Singh, Investigation of male and female students' motivational characteristics throughout an introductory physics course sequence, *Physics Education Research Conference* (2018) doi:10.1119/perc.2017.pr.064

15. 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).

16. 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?, *Proceedings of the 2018 Physics Education Research Conference* (2018).

17. 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).

18. V. Sawtelle, E. Brewe, L. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, Journal of Research in Science Teaching **49**, 1096 (2012).

19. P. Pintrich and E. De Groot, Motivational and self-regulated learning components of classroom academic performance, Journal of Educational Psychology **82**, 33 (1990).

20. B. Zimmerman, Self-efficacy: An essential motive to learn, Contemporary Educational Psychology **25**, 82 (2000).

21. D. Schunk and F. Pajares, The development of academic self-efficacy. In A. Wigfield and J. Eccles (Eds.), Development of Achievement Motivation: A Volume In The Educational Psychology Series (pp. 15-31). San Diego, CA: Academic Press (2002).

22. T. Bouffard-Bouchard, S. Parent and S. Larivee, Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students, International Journal of Behavioral Development **14**, 153 (1991).

23. S.L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, Journal of Research in Science Teaching **45,** 955 (2008).

24. M. Wang and J. Degol, Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields, Dev. Rev. **33**, 304 (2013).

25. B. Zimmerman, D. H. Schunk and M. K. Dibenedetto, The role of self-efficacy and related beliefs in self-regulation of learning and performance, *Handbook of Competence and Motivation*. Edited by A. Elliot, C. S. Dweck and D. S. Yeager, The Guilford Press; Second Edition (2018).

26. C. Dweck, Mindset: The New Psychology of Success (Ballentine, New York, 2006); C. Dweck, Self-theories: Their Role in Motivation, Personality, and Development (Psychology Press, Philadelphia, 1999).

27. C. Dweck, Is math a gift? Beliefs that put females at risk. In: Ceci, SJ.; Williams, WM., Editors. *Why Aren't More Women in Science*? Top researchers debate the evidence. (Washington, DC: p. 47-55, 2007).

28. C. Dweck, *Mindsets and Math/Science Achievement.* New York: Carnegie Corporation of New York, Institute for Advanced Study, Commission on Mathematics and Science Education, 2008.

29. D. S. Yeager, G. M. Walton, S. T. Brady, E. N. Akcinar, D. Paunesku, L. Keane, D. Kamentz, et al., Teaching a lay theory before college narrows achievement gaps at scale, PNAS **113**, E3341 (2016).

30. A. Little, B. Humphrey, A. Green, A. Nair and V. Sawtelle, Exploring mindset's applicability to students' experiences with challenge in transformed college physics courses, arXiv:1807.11062 [physics.ed-ph]; A. Little, V. Sawtelle and B. Humphrey, (2016, November). Mindset in Context: Developing new methodologies to study mindset in interview data. *Proceedings of the 2016 Physics Education Research Conference*, Sacramento, CA (204-207). American Institute of Physics.

31. R. E. Scherr, M. Plisch, K. E. Gray, G. Potvin and T. Hodapp, Fixed and growth mindsets in physics graduate admissions, Phys. Rev. Phys. Educ. **13**, 020133 (2017).

32. A. J. Elliot and A. M. Church, A hierarchical model of approach and avoidance achievement motivation, Journal of Personality and Social Psychology **72**, 218 (1997).

33. P. A. O'Keefe, A. Ben-Eliyahu and L. Linnenbrink-Garcia, Shaping achievement goal orientations in a mastery-structured environment and concomitant changes in related contingencies of self-worth, Motivation Emotion **37**, 50 (2013).

34. A. J. Elliot and J. M. Harackiewicz, Approach and avoidance goals and intrinsic motivation: A mediation analysis, Journal of Personality and Social Psychology **70**, 461 (1996).

35. J. G. Nicholls, Achievement and motivation: Conceptions of ability, subjective experience, task choice, and performance, Psychological Review **91**, 328 (1984).

36. K. Muenks and D. B. Miele, Students' thinking about effort and ability: The role of developmental, contextual, and individual difference factors, Review of Educational Research **87**, 707 (2017).

37. D. S. Yeager, D. Paunesku, G. M. Walton, and C. S. Dweck, (2013). How can we instill productive mindsets at scale? A review of the evidence and an initial R&D agenda. In *white paper prepared for the White House meeting on "Excellence in Education: The Importance of Academic Mindsets."*

38. D. S. Yeager and C. S. Dweck, Mindsets that promote resilience: When students believe that personal characteristics can be developed, Educational Psychologist **47**, 302 (2012).

39. C. S. Dweck and E. L. Leggett, A social-cognitive approach to motivation and personality, Psychological Review **95**, 256 (1988).

40. V. L. Henderson and C. S. Dweck, (1990). Motivation and Achievement, In S. S. Feldman & G. R. Elliott (Eds.), *At the Threshold: The Developing Adolescent* (pp. 308-329). Cambridge, MA, US: Harvard University Press.

41. J. Aronson, C. B. Fried and C. Good, Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence, Journal of Experimental Social Psychology **38**, 113 (2002).

42. C. Good, J. Aronson and M. Inzlicht, Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat, Applied Development Psychology **24**, 645 (2003).

43. L. S. Blackwell, K. H. Trzesniewski and C. S. Dweck, Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention, Child Development **78**, 246 (2007).

44. V. F. Sisk, A. P. Burgoyne, J. Sun, J. L. Butler and B. N. Macnamara, To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analysis, Psychological Sciences **29**, 549 (2018).

45. D. S. Yeager and G. M. Walton, Social-psychological interventions in education: They are magic, Review of Educational Research **81**, 267 (2011).

46. G. M. Walton, C. Logel, J. M. Peach, S. J. Spencer and M. P. Zanna, Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationship, and achievement in engineering, Journal of Educational Psychology **107**, 468 (2015).

47. D. Paunesku, G. M. Walton, C. Romero, E. N. Smith, D. S. Yeager and C. S. Dweck, Mind-set interventions are a scaleable treatment for academic underachievement, Psychological Science **26**, 784 (2015).

48. A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock and G. L. Cohen, Reducing the gender achievement gap in college science: A classroom study of values affirmation, Science **330**, 1234 (2010).

49. K. Binning, N. Kaufmann, E. McGreevy, O. Fotuhi, Susie 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, http://d-scholarship.pitt.edu/36006/

50. A. Rattan, C. Good and C. S. Dweck, “It's ok—Not everyone can be good at math”: Instructors with an entity theory comfort (and demotivate) students, Journal of Experimental Social Psychology **48**, 731 (2012).

51. E. A. Canning, K. Muenks, D. J. Green and M. C. Murphy, STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes, Science Advances **5**, eaau4734 (2019).

52. E. Shaw, An SAT ® Validity Primer, College Board Research, https://files.eric.ed.gov/fulltext/ED558085.pdf

53. S. J. Leslie, A. Chimpian, M. Meyer and E. Freeland, Women are underrepresented in disciplines that emphasize brilliance as the key to success, Science **347**, 262 (2015).

54. S. Upson and L. F. Friedman, Where are the female geniuses? Scientific American Mind **23**, 63 (2012).

55. L. Bian, S. J. Leslie and A. Chimpian, Gender stereotypes about intellectual ability emerge early and influence children’s interest, Science **355**, 389 (2017).

56. N. M. Hewitt and E. Seymour, A long, discouraging climb, ASEE Prism **1**, 24 (1992).

57. E. Seymour and N.M. Hewitt, Talking about leaving: Why undergraduates leave the sciences. Boulder, CO: Westview Press (1997).

58. J. M. Cortina, What is coefficient alpha? An examination of theory and applications, Journal of Applied Psychology **78**, 98 (1993).

59. E. Redish, J. Saul and R. Steinberg, Am. J. Phys. **66**, 212 (1998). doi:10.1119/1.18847.

60. L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

61. J. Nunnally, *Psychometric Theory* 2nd Ed. New York: McGraw-Hill (1978).

62. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (2013), http://www.R-project.org/

63. J. B. Kruskal, Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika **29**, 1 (1964).

64. J. Eccles, Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices, Psychology of Women Quarterly **18**, 585 (1994).

65. J. Eccles, Where are all the women? Gender differences in participation in physical sciences and engineering, in S. J. Ceci and W. M. Williams (Eds.), *Why Aren't More Women in Science?: Top Researchers Debate the Evidence* pp. 199-210 Washington, DC, US: American Psychological Association (2007).

66. R. M. Felder, G. N. Felder, M. Mauney, C. E. Hamrin Jr. and E. J. Dietz, A longitudinal study of student performance and retention. III. Gender differences in student performance and attitude, Journal of Engineering Education **84**, 151 (1995).

67. C. M. Mueller and C. S. Dweck, Praise for intelligence can undermine children's motivation and performance, Journal of Personality and Social Psychology **75**, 33 (1998).

68. C. Dweck (2016), How praise became a consolation in prize, Interviewed by Christine Gross-Loh, *Eduction*, Available at https://www.theatlantic.com/education/archive/2016/12/how-praise-became-a-consolation-prize/510845/ (Accessed: 14 April 2019).

8.0 Motivational characteristics of ethnic/racial minority students in the introductory physics courses

8.1 Introduction

Research shows that students from underrepresented racial and ethnic backgrounds sometimes under-perform in physics courses and few major in physics-related fields [1,2]. Despite some attempts to increase diversity in physics courses and reduce the performance gaps by race/ethnicity in physics [3-5], the underlying causes of the low representation and relatively poor performance of racial/ethnic minorities are not well understood.

Much research in physics education has focused on improving learning outcomes through the development of pedagogies that build on students' prior knowledge and learning difficulties. Although many research-based instructional strategies have been effective in improving students' physics content knowledge and problem-solving skills, there has been little focus on how student motivation may interact with instruction and learning outcomes [6]. For instance, students' self-efficacy can affect their motivation to learn and, ultimately, learning outcomes [7]. Students' value and interest in a particular academic task can also affect students' persistence and learning [8]. Therefore, it is important to examine students' motivational characteristics when developing and implementing instructional strategies.

8.2 Background

We describe a study in which we assessed students' intrinsic interest and self-efficacy throughout algebra-based and calculus-based introductory physics course sequences. We report White, Asian, and underrepresented racial/ethnic minority students' initial motivational characteristics and how these change over time. The findings have implications for the development and implementation of effective pedagogies to help all students learn.

Certain racial and ethnic groups have been traditionally underrepresented in physics, including African Americans, Latinx/Hispanics, American Indians, and Pacific Islanders [1]. Furthermore, students from these underrepresented groups often underperform in physics courses [2]. Although prior work has focused on performance differences between underrepresented students and majority students, less work has examined the motivation of underrepresented racial/ethnic minority students in physics courses.

Expectancy value theory posits that students' expected success in performing a task as well as the value they associate with the task can influence their persistence and performance in academic settings [9]. Expectancy involves one's beliefs about his or her own level of ability and expectations of success, e.g., self-efficacy, which is the belief in one's capability to be successful in a particular task, subject area, or course [10]. Even after controlling for students' prior knowledge and skills, higher self-efficacy is positively correlated with grades in physics [10]. Value includes several components such as intrinsic value (enjoyment from doing a task, or interest), attainment value (how important one feels to do well on the task), and utility value (usefulness of task in achieving goals) [9].

Since there have been relatively few longitudinal studies focusing on underrepresented racial/ethnic minority students' self-efficacy and intrinsic interest in physics [11], we investigated students' motivational characteristics at three time points during a two-semester algebra-based and calculus-based introductory physics course sequence. The findings can be a stepping stone to developing and implementing learning tools in order to help all students succeed in physics courses. The investigation can also help determine the motivational factors for which there are initial differences by race/ethnicity or whether differences develop across courses.

8.3 Methodology

We selected set of six motivational constructs based on a review of prior research on learning and motivation. Two of the motivational constructs that related to expectancy value theory involved self-efficacy and fascination with physics (i.e., intrinsic valuing of physics). We also investigated other motivational constructs including value associated with physics (i.e., importance of the knowledge of physics), students' intelligence mindset (i.e., students beliefs about whether intelligence is a fixed trait that one is born with or is malleable and can be shaped by the environment), grit (i.e., the capacity to sustain both effort and interest even in the absence of positive feedback), and physics epistemology (i.e., beliefs about what constitutes physics knowledge and how that knowledge is acquired). A survey was constructed to cover the core aspects of each construct with attention to minimize total survey length. The survey consisted of 29 items adapted from prior surveys from the motivation literature [12-16]. Table 8-1 shows example survey items for each motivational construct. Here, we focus on students' self-efficacy and fascination with physics (i.e., intrinsic value associated with physics). We do not focus on students' responses to the "valuing physics" items since those questions focused on general views about the importance of physics knowledge but were not directly related to students' intrinsic value, attainment value, or utility value.

We administered the survey in a written format to students at a large research university in the U.S. Students in algebra-based and calculus-based classes were asked to respond to survey questions at the beginning of Physics 1 in the fall semester, the beginning of Physics 2 in the spring semester, and also at the end of Physics 2 in the spring semester. Physics 1 courses cover topics involving Newtonian mechanics and Physics 2 courses include topics related to electricity and magnetism. Most of the courses were in a traditional, lecture-based format. Students in calculus-based classes are mostly freshmen who are engineering or physical science majors. Students in algebra-based physics courses are mostly juniors who intend to pursue careers in health-related fields or biology. The survey was completed by most students in about 10-15 minutes.

For each survey item, students were given a score of 1-4. For the items related to fascination and self-efficacy, a high score means that a student is highly fascinated by physics and has a high level of self-efficacy. Students were given an average score for self-efficacy and fascination. For example, a student who answered "Yes!" to two of the fascination with physics questions and "no" to one of them would have an average fascination score of (4+4+2)/(3 questions) = 3.33.

At the beginning of the fall and spring courses, we analyzed the internal consistency of the subscales, i.e., fascination, value, self-efficacy, intelligence mindset, grit, and physics epistemology. Analysis showed that all Cronbach's alphas for the targeted constructs were above 0.60, which is considered fairly good [17]. The scales with five or more items all had a Cronbach's alpha of 0.70 or higher. No substantial increases in alpha for any of the scales could have been achieved by eliminating items. To establish categories because there are positive (for Asians) and negative (underrepresented students) stereotypes for the separability of the different subscales along with the validity of items as clear indicators of the scale to which they were assigned, we performed an exploratory factor analysis on the items in the survey based upon the early spring data. A principal components analysis was used, and the initial eigenvalues indicated that the first six components (all with eigenvalues greater than 1) explained a total of 49% of the variance (the 7th component explained only an additional 4% of the variance). The data supported the existence of six separable scales, and items loaded on the scales as intended.

Table 8-1 Motivational factors with the number of items, example survey items, and scale.

Factor	Example Survey Item	Scale
Fascination with Physics 3 items	I wonder about how nature works …	Never Once a month Once a week Every day
Valuing Physics 5 items	Knowing physics is important for being a good citizen.	No! no yes Yes!
Self-Efficacy 6 items	I am often able to help my classmates with physics in the laboratory or in recitation.	No! no yes Yes!
Intelligence Mindset 4 items	-You have a certain amount of intelligence, and you can't really do much to change it. -Anyone can become good at solving physics problems through hard work.	Strongly disagree Disagree Agree Strongly agree
Grit 3 items	-I often set a goal but later choose to pursue a different one. -I finish whatever I begin.	Not like me at all Not much like me Somewhat like me Mostly like me Very much like me
Physics Epistemology 8 items	I do not expect to understand physics equations in an intuitive sense; they must just be taken as givens.	Strongly disagree Disagree Agree Strongly agree

After linking students' responses to their demographic information, we separated the students into three groups: White, Asian, and underrepresented racial/ethnic minority students. Underrepresented racial/ethnic minority students are those who reported their race/ethnicity as Black, Hispanic, American Indian, Pacific Islander, or Multi-Racial including one of these races.

We examined these racial groups' performance together, and because there were too few underrepresented racial/ethnic minority students to further divide them in the in the analyses. Black, Hispanic, American Indian, Pacific Islander, and Multi-Racial students have been underrepresented in STEM fields, and in particular, physics [1]. Table 8-2 shows the number of students in each ethnic/racial group at the three time points at which the survey was given.

Table 8-2 The number of White (W), Asian (A), and underrepresented racial/ethnic minority (U) students who took the survey at the three time points.

	Beginning of Physics 1	Beginning of Physics 2	End of Physics 2
Calculus-based	N=364 W, 57 A, 42 U	N=219 W, 39 A, 21 U	N=178 W, 32 A, 12 U
Algebra-based	N=285 W, 82 A, 50 U	N=150 W, 52 A, 16 U	N=123 W, 40 A, 13 U

8.4 Results

See Figure 8-1 for students' average self-efficacy scores in physics courses. An ANOVA showed that at the beginning of calculus-based Physics 1, there was a statistically significant difference in students' self-efficacy scores in the three racial/ethnic groups ($F(2,460)=5.42$, $p=0.005$). White students initially had the highest self-efficacy score whereas underrepresented racial/ethnic minority students' self-efficacy is the lowest at the beginning of calculus-based Physics 1. We observed that underrepresented racial/ethnic minority students' self-efficacy score increased significantly more than White and Asian students between the beginning of calculus-based Physics 1 and 2 ($F(2, 277)=4.22$, $p=0.016$) and the differences in the self-efficacy of the three groups was eliminated. By the end of calculus-based Physics 2, we observed that all students' self-efficacy did not change significantly with respect to the beginning of Physics 2 and all three groups had approximately the same self-efficacy score (see upper left graph in Figure 8-1). In algebra-based physics courses, statistical analyses showed that there was not a significant difference in the self-efficacy of students in the three groups at the beginning of Physics 1. Also, students' self-efficacy stayed approximately the same throughout algebra-based Physics 1 and Physics 2.

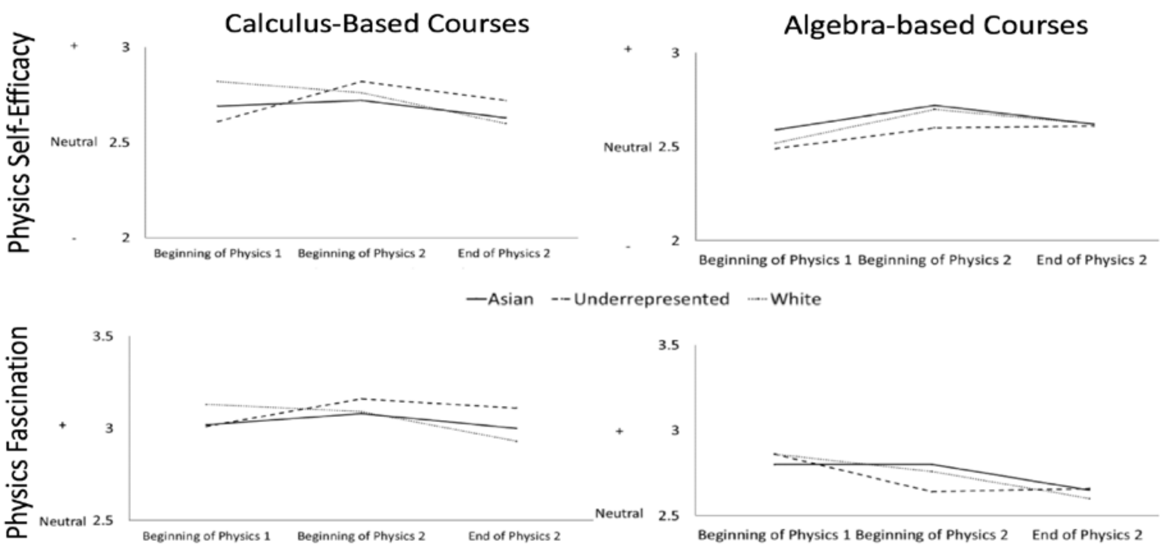

Figure 8-1 Physics Self-efficacy (top row) and Fascination (botton row) in calculus-based (left column) and algebra-based (right column) introductory physics sequences. "+" and "-" signs indicate positive and negative responses.

We now describe students' fascination with physics (i.e., their intrinsic interest in physics) throughout introductory physics course sequences. At the beginning of calculus-based Physics 1, statistical analyses showed that there was no significant difference in students' fascination scores in the three racial/ethnic groups. Also, students' fascination remained approximately the same throughout calculus-based Physics 1 and Physics 2. Students from underrepresented racial/ethnic minority groups reported the highest fascination with physics at the end of calculus-based Physics

2. In algebra-based physics courses, statistical analyses showed that there was not a significant difference in the fascination of students in the three racial/ethnic groups at the beginning of Physics 1. Students in algebra-based physics courses reported lower levels of fascination with physics than students in calculus-based physics courses, and all students' fascination with physics decreased slightly by the end of algebra-based Physics 2.

8.5 Discussion and summary

In calculus-based physics courses, underrepresented racial/ethnic minority students initially had lower self-efficacy than White and Asian students. Underrepresented racial/ethnic minority students who persisted throughout calculus-based physics courses showed an increase in self-efficacy after Physics 1 and the gap between underrepresented racial/ethnic minority students and majority students' self-efficacy was reduced. Underrepresented racial/ethnic minority students also reported approximately the same level of fascination with physics as White and Asian students throughout the calculus-based physics course sequence. On the other hand, students who persisted in the algebra-based physics courses reported approximately the same self-efficacy and fascination throughout the courses, regardless of their race/ethnicity. Furthermore, students in algebra-based physics courses reported lower fascination with physics than students in calculus-based physics courses, and their fascination decreased by the end of Physics 2.

In terms of Expectancy Value Theory, our results suggest that students who persist in taking physics courses for engineering and physical science fields tend to value physics, i.e., they have an intrinsic interest in physics, regardless of what racial/ethnic group they belong to. However, underrepresented racial/ethnic minority students in calculus-based physics courses tend to enter Physics 1 with lower expected success in the course (e.g., lower self-efficacy) than White and Asian students. It is possible that underrepresented racial/ethnic minority students initially have lower self-efficacy than White students because they feel that they do not belong or an awareness of the stereotype that minority students underperform in STEM fields. Our findings suggest that underrepresented racial/ethnic minority students who successfully progress in the calculus-based physics course sequence gain feedback about their ability in relation to other students (e.g., all students struggle in physics). This feedback informs them about their ability to succeed and increases their self-efficacy. It is also possible that underrepresented racial/ethnic minority students' intrinsic interest in physics and attainment value in physics (i.e., it is important to them to do well in physics to dis-confirm negative stereotypes about underrepresented racial/ethnic minority students in STEM fields) are important factors for persisting in physics courses.

In contrast, students who succeed in progressing through an algebra-based physics course sequence for biology and health fields are less intrinsically interested in physics than students in the calculus-based courses. Furthermore, their intrinsic value associated with physics decreases throughout an introductory physics course sequence, regardless of what racial/ethnic group they belong to. Our findings suggest that students in algebra-based physics courses have less intrinsic interest in physics perhaps because they do not see the relevance of physics to their major/career goals (they intend to pursue careers in health-related fields or biology). These students' motivation to persist in physics courses may be due, in part, to the utility value in that successfully completing the physics course sequence helps them achieve future goals since physics courses are required for their majors.

Instructors, researchers, and curriculum developers in physics can use these findings to develop and implement effective approaches and learning tools, in part, by taking into account students' motivational characteristics. Focusing on students' motivational characteristics, especially those of underrepresented racial/ethnic minority students, may prove fruitful in helping them succeed in physics courses and increase diversity in physics-related fields. The findings of this study suggest that it may be beneficial to provide explicit support to improve underrepresented racial/ethnic minority students' self-efficacy in physics courses for engineers and physical scientists to help more of them persist and succeed in those physics courses. For example, discussions early on in the physics course sequence that focus on effective study strategies, attributing success to hard work (as opposed to ability), and the fact that many students struggle in physics may improve underrepresented racial/ethnic minority students' self-efficacy and motivate them to persist in physics course sequences for engineers and scientists. In physics courses for students in biology and health-related fields, it may be useful to help students see the relevance of physics to health and biology related fields in order to motivate them to engage more deeply with course instruction. We plan to collect more data in the following semesters to draw stronger conclusions and investigate how the motivational characteristics discussed here relate to the type of instruction and learning over time in physics courses.

8.6 Chapter references

1. https://www.nsf.gov/statistics/2017/nsf17310/data.cfm

2. E. Brewe, V. Sawtelle, L. Kramer, G. O'Brien, I. Rodriquez and P. Pamela, Toward equity through participation in Modeling Instruction in introductory university physics, Phys. Rev. ST PER **6**, 010106 (2010).

3. P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, American Journal of Physics **60**, 637 (1992).

4. R. Steinberg and K. Donnelly, PER-Phys. Teach. **40**, 108 (2002).

5. A. Maries, C. Singh, Stereotype threat? Effects of inquiring about test takers' gender on conceptual test performance in physics, AIP Conference Proceeding **1697**, 120098 (2016); doi: http://dx.doi.org/10.1063/1.4937713

6. A. Wood, R. Galloway, R. Donnelly and J. Hardy, PRPER **12**, 010140 (2016).

7. R. Lent, S. Brown and K. Larkin, J. Counsel. Psych. **33**, 265 (1986).

8. P. Pintrich, J. of Educ. Psy. **95**, 667 (2003).

9. A. Wigfield and J. Eccles, *Contemporary Educational Psychology*, pp. 68-81 (2000).

10. A. Bandura and D. Schunk, J. Pers. Soc. Psy. **41**, 586 (1981); V. Sawtelle et al., J. Res. Sci. Teach. **49**, 1096 (2012).

11. E. Marshman, Z. Kalender, T. Nokes-Malach, C. Schunn and C. Singh., Can. J. Phys; doi: http://dx.doi.org/10.1139/cjp-2017-0185

12. E. Redish, J. Saul and R. Steinberg, Am. J. Phys. **66**, 212 (1998).

13. A. Duckworth and P. Quinn, J. Per. Assess. **91**, 166 (2009).

14. D. Molden and C. Dweck, Am. Psy. **61**, 192 (2006).

15. W. Adams, K. Perkins, N. Podolefsky, M. Dubson, N. Finkelstein and C. Wieman, Phys. Rev. ST PER **2**, 010101 (2006).

16. S. Glynn, P. Brickman, N. Armstron and G. Taasoobshirazi, J. Res. Sci. Teach. **48**, 1159 (2011).

17. L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

9.0 Future directions

The findings in this thesis have implications for improving teaching and learning of physics and for developing interventions that promote diversity in the physics community. The majority of the studies presented here focus on gender differences in introductory physics courses. The gender data obtained by university records included binary response options; however, gender is a sociocultural and multidimensional construct. In the future studies, gender measurement can be designed by incorporating multiple options, which can allow us to measure this construct in more nuanced ways.

Moreover, we mostly examined the calculus-based physics courses. Expanding upon the various analyses conducted in this thesis, algebra-based physics courses can also be investigated in a similar way to understand the interaction between gender and motivational characteristics. Apart from the differences in curriculum content and math-intensity, there are differences between the two types of physics courses in student demographics and classroom structure. First, calculus-based physics classes are taken by engineering and physical science majors, whereas health-sciences, biology, and pre-medical students enroll in algebra-based physics courses. Students' academic goals and level of interest in understanding physics may differ in the two kinds of physics classes. Equally importantly, algebra-based courses are generally taken in sophomore and junior years as opposed to introductory calculus-based physics courses, which are taken typically by physical science and engineering freshmen students. While the younger students in the calculus-based physics courses may still maintain high-school-based beliefs about physics, those who are in the algebra-based course have gone through several years of college and campus life experience and their attitudes might have shifted since high school. Finally, while women are a minority in calculus-based physics classes (approximately 30% of the cohort), this proportion is double in algebra-based courses. Each of these factors may change the core determinants of motivational factors and their role in learning outcomes for students in the algebra-based introductory physics courses.

In addition, the lack of race/ethnic diversity in physics is also a potentially important underlying issue which may be entangled with and contribute to the lack of gender diversity in physics. We have only conducted one study for ethnic/racial minority groups in this thesis. Due to the low number of ethnic/racial minority students, it is difficult to draw definite conclusions about their motivational characteristic differences in physics courses. Accumulating more longitudinal data in the introductory physics courses can increase our statistical power that can help us understand and address the representation gaps for race and ethnicity in physics.

Research into the motivational aspects of student learning has the potential to provide important insights regarding possible reasons for the under-representation of certain groups of students in physics-related STEM fields. Investigating the origins of motivational characteristics gaps among students by gender or race/ethnicity not only informs physics education researchers but also educational policy makers, curriculum developers, and instructors of physics in order to promote a more equitable, diverse, and inclusive learning environment for all students in the discipline of physics.