CS1 AND GENDER: UNDERSTANDING EFFECTS OF BACKGROUND AND SELF-EFFICACY ON ACHIEVEMENT AND INTEREST

By

Philip Sands

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Educational Psychology and Educational Technology - Doctor of Philosophy

2021

ABSTRACT

CS1 AND GENDER: UNDERSTANDING EFFECTS OF BACKGROUND AND SELF-EFFICACY ON ACHIEVEMENT AND INTEREST

By

Philip Sands

Over the past 20 years, the field of computer science has experienced a growth in student interest. Despite this increase in participation rates, longstanding gender gaps persist in computer science. Recent research has examined a wide variety of individual factors (e.g., self-efficacy, sense of belonging, etc.) that impact student interest and achievement in computer science; however, these factors are rarely considered in the context of existing learning theories. In this correlational study, I explored the relationship between prior knowledge of computer programming, self-efficacy, and the sources of self-efficacy as they differed by gender in a theoretical model of achievement and interest for students in first-year computer science (CS1) courses. This model was based on prior work from Bandura (1997) and others exploring self-efficacy and social cognitive theory in the context of mathematics and science fields. Using cross-sectional data from N=182 CS1 students at two universities, structural regressions were conducted between factors impacting CS1 students across the entire population and for men (N=108) and women (N=70) individually. This data was then used to address the following research questions. (1A) How do prior knowledge of computer programming, the sources of self-efficacy, and self-efficacy for computing predict CS1 achievement and student intentions to continue study in CS? (1B) How does self-efficacy mediate the relationship between student prior knowledge of computer programming and achievement in CS1? (1C) How are those relationships moderated by gender? (2) How does feedback in the form of student grades impact intention to continue in CS when considering gender as a moderating factor? For all students,

student self-efficacy for CS positively impacted CS1 achievement and post-CS1 interest. Aligning with past research, self-efficacy was derived largely from mastery experiences, with vicarious experiences and social persuasions also contributing to a moderate degree. Social persuasions had a negative effect on self-efficacy, which diverged from research in other fields. The relationship between prior knowledge of computer programming and CS1 achievement was not mediated by self-efficacy and had a small positive effect. For women, vicarious experiences played a stronger role in defining student self-efficacy in CS. Additionally, while the importance of self-efficacy on achievement was similar to that for men, self-efficacy and achievement both played a much stronger role in determining student interest in CS for women. All these findings are in need of further exploration as the analysis was underpowered due to a small, COVID-19 impacted sample size. Future work should focus on the role of feedback on student self-efficacy, the potential misalignment of CS1 feedback and social network feedback, and interventions that address student beliefs about CS abilities to increase opportunities for authentic mastery and vicarious experiences.

For Heidi, Robyn, Grace, Nicole, Lisa, Abby, Alexis, Rima, Samantha, Annie, Janka, Manasi, Rhea, Katarina, Shannon, Calgary, Meghan, Raj, Sharan, Tejaswi, Joyce, Sasha, Corinne, Tania, Kavya, Emily, Ashley, Kendall, Andrea, Pooja, Roopa, Victoria, Selin, Eehita, Yamini, Jackie, Leah, Shivani, Tori, Julianna, Angie, Jessie, Brittany, Megan, Marilyn, Rachel, and Hannah.

From GET-IT, to MAGIC, to a future when being a woman in CS won't come with so much extra baggage.

ACKNOWLEDGMENTS

"Don't become the thing you hated." - Dan Bejar

Before I offer my appreciation to those that helped me through my graduate experience, I want to acknowledge the incongruous relationship between my achievement and my personal beliefs. Based on my personal experiences in academia, I do not believe that academic research has a powerful impact on education, teachers, students, and learning environments. There are a significant number of education researchers, and particularly computer science education researchers, who are focused on their own career achievements and not the betterment of their subjects. I initially elected to follow this path because I naively believed that I might be part of a positive change in CS education, and that an advanced degree would give me access to opportunities that were otherwise closed to me. Having finished this work, I will press forward with a focus on improving the experiences of teachers and students and committing myself to these goals above all else. Hopefully, I will buck the trend, and this will not be vainglorious performance for the sake of ego and further self-deception.

Now that I have acknowledged my concerns about academia (and my complicit role in participating further), allow me to push forward to thank those who have supported me despite my misgivings. Completing this work has been a continuous struggle. I would not have been able to make it through in one piece without the support of my friends, colleagues, and family. I will risk trying to thank people by name, knowing full well that I will inevitably forget someone important. If you find yourself reading this and thinking, "Hey, what about that time I listened to you blather on about measurement error for 25 minutes? Don't I deserve a shout out?", then you probably deserve a shout out and I apologize for not giving you your due.

My family has always been supportive of education. I am grateful for my mother, Eileen, and her positivity while I fretted over sample size issues. I am thankful for my father, Gary, who offered both amusing stories of his own dissertation process, and reminders to me of the mercurial nature and at times superfluous role of faculty advisors. My siblings (Chris, Karen, and Diane) occasionally made efforts to check in on me to make sure I wasn't losing my mind. I'm also appreciative of those in my extended family for their love and support: Margaret, Eric, Rachel, Miriam, Elena, Victoria, Gwendolyn, and Verceli. Lastly, I would like to thank my in-laws, the McGovern clan, for being so generous over the years.

I have been well supported by several close friends during my graduate studies, including DJ and Emily Martin, Corey and Amy Williams, Dustin and Katie Harnish, Andrew Woolcock and Brian Hitselberger, Patrick Connolly, Jim Degenhardt, Chris and Jamie Seadeek, Sarah Pounders and Bryan Rayo, Nathan and Monica Copenhaver, Victory Soe and Gabbi Shafer, Lucas Woodcock and Nancy Kipka, Angelina Garner, and Lakisha David. Many of these folks have wonderful kids who have brought lots of joy to my life during my years in graduate school; I am grateful to be everyone's fake uncle! My own research has reinforced for me that role models matter, and so I hope that my efforts to persist will encourage these wonderful young people to continue to embrace education and take on challenges in their own lives. I can certainly point to a number of positive role models in my life, and many of them are included in the list of supportive people that I have named above. I am extremely lucky to have had the wonderful support and encouragement from such lovely people. Friendships come and go, but the ones who stick around and support you through your struggles are the ones you really appreciate. In that way I feel fortunate because these folks have been truly terrific to me.

My dissertation committee (Jen Schmidt, Colleen Lewis, Rich Enbody, & Aman Yadav) got more than they bargained for by taking me on as a student. They were patient when I needed extra time to work through some major changes during the planning phase and took the time to talk me through their ideas for final revisions. Those final revisions helped take a limited study from a questionable form to something I am quite proud to present. Having a committee that was understanding of my limitations as a researcher and who were willing to push me to overcome challenges as I faced them meant a tremendous amount to me. My dissertation goal was growth, knowing full well that I began the project in need of it. My committee supported this goal and encouraged every opportunity for me to develop as a scholar. Even if they hadn't, I feel that they deserve at least minor praise for agreeing to read through my work. Most of the committee attempted to read the entire thing! In particular, I want to thank Jen Schmidt for helping me to make advancements while engaging with a difficult personality, and for meticulously identifying all the holes in my study. If this ever reaches the publishing stage, it will largely be due to her noteworthy efforts to help me transform a developing study into a more complete one.

Being able to study remotely was a great affordance of the EPET program. I am thankful for the efforts of Cary Roseth, Matt Koehler, Lisa Linnenbrink-Garcia, Emily Bouck, Chris Greenhow, and Jack Smith. All taught me exciting things and pushed me to be a more complete scholar. More importantly, I would not have persisted through the EPET program without my dear lab mates, base group members, and peers. In particular, I would like to thank Katie Rich, Sukanya Moudgalya, Anna Sansone, Jessica Sprick, Bret Staudt-Willet, Alex Lishinski, Jon Good, Sarah Gretter, Chris Seals, and Kristy Robinson.

I began my work at Michigan State while still employed by Purdue University's Department of Computer Science. There I had great mentors and friends, including Tim Korb, Susanne Hambrusch, Sunil Prabhakar, Buster Dunsmore, Gustavo Rodriguez-Rivera, Ben Delaware, and Roopsha Samanta. All have pushed me to reach for greater heights and continue to mean a lot to me as I reflect on my career path. Additionally, I had excellent friends and collaborators in Purdue's College of Education which included Brenda Capobianco and Jim Lehman foremost among them. I have learned much from both Brenda and Jim and am thankful for their continued encouragement.

Over my career in education, I have been lucky to be inspired by so many great students. I want to thank them all for their love and support. My work in education continues to be dedicated to them. I would never be able to name them all, but I want to specifically thank Robyn Berkel, Sean Flannery, and Guna Kondapaneni for their hard work and friendship.

Music has played a tremendous part in my ability to persist on tasks that greatly reduce my enthusiasm. More than anyone else, I wrote this dissertation to the music of New Order (2011) and Kurt Vile (2013). I give much thanks to Bernard Sumner, Gillian Gilbert, Stephen Morris, Peter Hook, and Ian Curtis of New Order and Joy Division, and KV of Kurt Vile and the Violators. If I gave credit to all the musicians that kept me going, I would have to add extra pages to my dissertation, but I am grateful for all the great individuals and bands that have kept my spirits up as I labored away at my computer keyboard.

Saving the best for last, I am extremely grateful to my partner, Gina, who really *did* listen to me talk about measurement error (and literally every other thing that is in this dissertation). Her unflagging support and belief were crucial to me not quitting many times over, including the literal final days before I submitted the final draft. During the 20 months of work that I invested in this study, we would take long hikes with our dogs and talk about the work, the reasons for the work, and how the efforts might someday make the whole world marginally better. Every time I conversationally drifted downward, she found a way to elevate me. When I tried to reason with her that our lives would be so much easier if I just bought a food truck and sold burritos to college kids, she would convince me that I was being entirely irrational and should get back to work. I would be lost without her compassion, brilliance, and humanity, and am forever grateful that of all the people in the world that she could have partnered up with she chose to waste away her hours with me.

TABLE OF CONTENTS

LIST OF TABLES xiii

LIST OF FIGURES xv

CHAPTER 1: INTRODUCTION 1
1.1 Background 1

CHAPTER 2: RELEVANT LITERATURE 4
2.1 Literature and Theory 4
2.1.1 Gender and Computer Science 4
2.1.2 Self-efficacy in CS1 6
2.1.3 Sources of Self-efficacy 8
2.1.4 Student Background with CS 11
2.1.4.1 Prior Knowledge of Computer Programming 11
2.1.5 Other Factors Impacting CS1 Achievement and Post-CS1 Interest 12
2.2 Theoretical Framework 13
2.3 Goals of the Study 16

CHAPTER 3: METHODOLOGY 18
3.1 Design of the Study 18
3.2 Study Participants 20
3.3 Independent Measures 22
3.3.1 Self-efficacy 22
3.3.2 Sources of Self-efficacy 23
3.3.3 Prior Knowledge of Computer Programming 24
3.3.4 Factor Analysis of Measures 26
3.4 Dependent Measures 28
3.4.1 Academic Achievement in CS1 28
3.4.2 Post-CS1 Interest in Computer Science 28
3.5 Data Collection Procedure 29
3.6 Data Analysis 30
3.6.1 Data Preparation 30
3.6.2 Measurement Invariance 32
3.6.3 Missing Data 33
3.6.4 Structural Regression 34

CHAPTER 4: RESULTS 36
4.1 Analysis of the Sample 36
4.1.1 Descriptive Statistics 36
4.1.2 Correlational Statistics 37
4.1.3 Multivariate Assumptions 38
4.1.3.1 Outliers and Influential Data 38
4.1.3.2 Multicollinearity 39

4.1.3.3 Normality 40
4.2 Model for CS1 Achievement and Post-CS1 Interest 41
4.2.1 Model Specification 42
4.2.1.1 Conditional Independence of Latent Factors 43
4.2.2 Model Identification 43
4.2.3 Model Estimation and Fit for All Students 44
4.3 Model Considering Gender as a Moderator 46
4.3.1 Individual Path Differences Considering Gender 46
4.3.2 Model Differences by Gender 48

CHAPTER 5: DISCUSSION 50
5.1 Findings 50
5.1.1 Research Question 1A 50
5.1.1.1 Model Pathways Confirming Prior Research on CS1 51
5.1.1.2 Novel Findings from CS1 Model 55
5.1.2 Research Question 1B 58
5.1.3 Research Question 1C 60
5.1.3.1 Model Pathways Confirming Prior Research on Women in CS1 61
5.1.3.2 Novel Findings from CS1 Model for Women 62
5.1.4 Research Question 2 63
5.2 Implications 66
5.2.1 Implications for Research 66
5.2.2 Implications for Practice 70
5.3 Limitations and Delimitations 74
5.4 Conclusions 78

APPENDICES 80
APPENDIX A: TABLES 81
APPENDIX B: FIGURES 102
APPENDIX C: CODEBOOK FOR STUDY MEASUURES 117
APPENDIX D: MPLUS CODE 135
APPENDIX E: LITERATURE SUPPORTING EXPLORATION OF ADDITIONAL CONSTRUCTS 141
APPENDIX F: OPERATIONALIZATION AND PRELIMINARY ANALYSIS OF ADDITIONAL CONSTRUCTS 150

WORKS CITED 159

LIST OF TABLES

Table 3.1 *Demographic and school information for sample.* *81*

Table 3.2 *Dimensional reduction for sources of self-efficacy, self-efficacy, and cost measures.* *82*

Table 3.3 *Grade point scale and conversion used for grades data.* *84*

Table 4.1 *Means, standard deviations, and independent samples t statistics for study measures.*

.... *85*

Table 4.2 *Correlations between factors in the CS1 model for all students in the sample.* *86*

Table 4.3 *Correlations with dependent factors in the CS1 model and independent samples Fisher r-to-Z transformations by gender.* *87*

Table 4.4 *Correlations for indicators of latent factors in the CS1 model.* *88*

Table 4.5 *Tolerance and variance inflation factor statistics for CS1 achievement, post-CS1 interest, and self-efficacy.* *89*

Table 4.6 *Skewness and kurtosis values for all indicator items of latent factors in CS1 model.* *90*

Table 4.7 *Partial correlations for latent factors in CS1 model with given conditioning sets.* *91*

Table 4.8 *Standardized and unstandardized estimates for CS1 model for all students.* *92*

Table 4.9 *Direct and indirect effect estimates and 95% confidence intervals for prior knowledge of computer programming on CS1 achievement with self-efficacy as a mediator.* *93*

Table 4.10 *Changes in chi-square difference test for model given the constraint of paths to be equal by gender.* *93*

Table 4.11 *Model fit statistics for CS1 model with increasingly restrictive paths for estimation by gender.* *94*

Table 4.12 *Standardized and unstandardized estimates for CS1 model for women.* *94*

Table 4.13 *Standardized and unstandardized estimates for CS1 model for men.* *95*

Table F.1 *Means, standard deviations, and independent samples t statistics for unused task and cost measures.* *96*

Table F.2 *Correlations between factors in the CS1 model, utility value, and cost measures for all students in the sample.* 97

Table F.3 *Correlations between dependent factors and those measured but not included in the CS1 model with independent samples Fisher r-to-Z transformations by gender.* 98

Table F.4 *Means, standard deviations, and independent samples t statistics for prior experience factors.* 99

Table F.5 *Means, standard deviations, and one-way analyses of variance in AP CS A and programming language experiences on CS1 achievement by gender.* 100

Table F.6: *Correlations for measures of prior knowledge of computer programming and prior experience with computer science.* 101

LIST OF FIGURES

Figure 2.1 *Model of CS1 achievement and post-CS1 interest.* *102*

Figure 3.1 *Confirmatory factor analysis for motivational factors in CS1 model.* *103*

Figure 3.2 *Data collection timeline for survey instruments and pre-test.* *104*

Figure 4.1 *Mahalanobis distances of achievement predictors for all data in sample of CS1 students.* *104*

Figure 4.2 *Mahalanobis distances of interest predictors for all data in sample of CS1 students.* *105*

Figure 4.3 *Residual plot of prior knowledge of computer programming on CS1 achievement.* *106*

Figure 4.4 *Residual plot of self-efficacy on CS1 achievement.*.. *106*

Figure 4.5 *Residual plot of CS1 achievement on post-CS1 interest.* *107*

Figure 4.6 *Residual plot of self-efficacy on post-CS1 interest.*.. *107*

Figure 4.7 *Residual plot of prior knowledge of computer programming on self-efficacy.*. *108*

Figure 4.8 *Residual plot of mastery experience on self-efficacy.* *108*

Figure 4.9 *Residual plot of vicarious experience on self-efficacy.* *109*

Figure 4.10 *Residual plot of social persuasions on self-efficacy.*....................................... *109*

Figure 4.11 *Residual plot of physiological states on self-efficacy.*..................................... *110*

Figure 4.12 *Normal Q-Q plot of prior knowledge of computer programming on CS1 achievement.* *110*

Figure 4.13 *Normal Q-Q plot of self-efficacy on CS1 achievement.* *111*

Figure 4.14 *Normal Q-Q plot of CS1 achievement on post-CS1 interest.* *111*

Figure 4.15 *Normal Q-Q plot of self-efficacy on post-CS1 interest.* *112*

Figure 4.16 *Normal Q-Q plot of prior knowledge of computer programming on self-efficacy.* *112*

Figure 4.17 *Normal Q-Q plot of mastery experience on self-efficacy.* *113*

Figure 4.18 *Normal Q-Q plot of vicarious experience on self-efficacy.* 113

Figure 4.19 *Normal Q-Q plot of social persuasions on self-efficacy.* 114

Figure 4.20 *Normal Q-Q plot of physiological states on self-efficacy.* 114

Figure 4.21 *Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for all students.* 115

Figure 4.22 *Model of CS1 achievement and post-CS1 interest with path labels.* 115

Figure 4.23 *Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for women.* 116

Figure 4.24 *Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for all students.* 116

CHAPTER 1: INTRODUCTION

1.1 Background

The computer science (CS) community has struggled for decades to address the imbalance in participation by women across all educational and professional levels. The percentage of bachelor's degrees in CS awarded to women declined to 18.7% in 2016 compared with 37% in 1985 (NCSES, 2019; NCWIT, 2020). Employment numbers for women computer scientists are only slightly higher, with 25.4% of CS jobs being held by women in 2017 (NCSES, 2019). The rate of participation by women has shown some growth in recent years, with percentage of CS bachelor's degrees awarded increasing to 21% in 2020, but this remains a far cry from the high mark in the mid-1980s (NCWIT, 2020). Increasing the participation rates of women in CS has proven to be difficult and is made particularly challenging by the reinforcing effect of overtly masculine workplace culture that creates difficult working environments for women (Ashcraft et al., 2016; Cohoon & Aspray, 2006). Gender imbalance is a problem of significant importance, as it impacts the computing products created by businesses and leaves women out of the highest paying jobs that CS provides (Margolis & Fisher, 2003). Further, the exclusion of women from the computing field precludes technology companies from reaching their full potential for innovation, high quality products, and company culture (Ashcraft et al., 2016). These ideas can extend to the scientific research community as well. Imagining an environment in which roughly half of the potential scientific brainpower is not participating means fewer advances in computing technology and less frequent application of that technology to other fields. It is to the benefit of all those participating in computer science in some form to embrace gender equity.

The issue of increasing the number of women in computing involves both recruiting women to the field and retaining them once they begin studying computer science. There have been some recent recruiting successes, with 1.5% of college-aged women in 2017 declaring an interest in a computer science major versus 0.3% in 2009 (Pryor et al., 2009; Stolzenberg et al., 2019). Unfortunately, the attrition rate for women in CS programs has been found to be twice as high as the rate for male students (Biggers at al., 2008), with roughly 40% of women eventually dropping out of computing programs (Cohoon, 2001). University CS courses are often the first computing experiences for many students (Margolis & Fisher, 2003). Therefore, the study of student experiences with the first CS course (CS1) is of key interest to CS education researchers. Researchers have found a lower-than-average pass rate in the CS1 course across universities (Watson & Li, 2014), and have hypothesized that the awareness of the low pass rate leads to lower interest in CS for all students (Bennedsen & Caspersen, 2007). Many studies in CS education have focused on understanding factors that impact the decision of women to persist beyond the CS1 course, as it is frequently noted as a point of departure for women interested in a computing career path (Biggers et al., 2008; Pappas et al., 2016; Petersen et al., 2016).

Ultimately, the decision to continue in a computer science program of study is based on many complex factors and the interactions between them. While prior research has identified several individual factors that impact student interest in the study of CS (i.e., self-efficacy, sense of belonging), these factors are not often considered in the context of existing learning theories. Understanding how the most salient factors manifest for women entering university computer science programs is of critical interest to CS educators and researchers and may help focus efforts to retain women in computing careers by pointing to areas for interventions to be implemented. I conducted a correlational study to examine the antecedents of student self-efficacy and the role of self-efficacy and prior knowledge of computer science on post-CS1 interest and academic achievement in CS1 for women students. To explore these elements in this study, I addressed the following research questions.

RQ1A: How do prior knowledge of computer programming, the sources of self-efficacy, and self-efficacy for computing predict CS1 achievement and student intentions to continue study in CS?

RQ1B: How does self-efficacy mediate the relationship between student prior knowledge of computer programming and achievement in CS1?

RQ1C: How are those relationships moderated by gender?

RQ2: How does feedback in the form of student grades impact intention to continue in CS when considering gender as a moderating factor?

CHAPTER 2: RELEVANT LITERATURE

2.1 Literature and Theory

2.1.1 Gender and Computer Science

The percentage of women participating in CS jobs has been in decline since 1991, when 36% of total IT workers in the United States were women (Ashcraft et al., 2016). Women also leave IT jobs at over twice the rate of men, with 41% of women departing the IT career path (Ashcraft et al., 2016). At the university level, while the number of women getting advanced CS degrees has shown some minor gains, attainment of undergraduate computing degrees by women in that same period decreased by 50% (NCSES, 2019). Regarding the interest of girls in K-12 computing, there are also some concerning signs. Among students in grades 7-12, girls were half as likely to say that they were “very interested” in computer science as boys and less likely to express confidence that they could succeed in computing courses (Google & Gallup, 2016). Girls were also taking fewer computer science AP exams and fewer computer science courses than boys and were less likely to have taken a CS course by the time they began college (College Board, 2020; Google & Gallup, 2016). Research suggests that the experience of pursuing a computer science career path is different for women than for men, and that these aspects are impacting long-term interest for women in computing (Bernstein, 1991; Beyer, 2014; Beyer et al., 2003; Biggers et al., 2008; Cheryan et al., 2017; Margolis & Fisher, 2003; Master et al., 2016; Sax et al., 2017).

One line of research suggests that student interest development differs by gender in computer science specifically, and science, technology, engineering, and math (STEM) more generally. Sadler et al. (2012) found that participation in STEM career paths can be predicted by a combination of early high school STEM interest and gender, with women losing interest in STEM significantly over the course of their high school careers. Similarly, gender has been shown to be a predictor of involvement in IT and CS career pathways, with women expressing lower interest in CS (Zingaro, 2015), being less likely to choose an IT career path (Zarrett & Malanchuk, 2005), and being more likely to drop out from CS courses perceived as being more technical in nature (Sheard et al., 2008). In addition to its role in predicting interest, gender has also been shown to predict academic achievement in CS. Goold and Rimmer (2000) found that gender was a significant predictor of academic achievement in CS1 courses, although this effect did not persist in courses beyond CS1. Bergin and Reilly (2005) found similar results, with gender being a significant part of a model predicting CS performance along with math grades, science grades, and belonging in the course. Other studies have suggested that gender differences in achievement are either shrinking or can be attributed to prior experience with computing (Priess & Hyde, 2010; Wilcox & Lionelle, 2018; Wilson & Shrock, 2001). To better understand why interest and performance outcomes can be predicted by gender in CS, it is important to understand factors that impact self-beliefs, interests, and long-term career choices.

Due to concerns about negative participation trends for women in CS, the past several decades has seen an increase in research focused on how gender influences student performance in CS1 and student interest in computer science. The CS1 course is the introduction to computer programming and other core ideas in the computer science curriculum (ACM, 2013). For some students, this course is both the entry point and the departure point for them with regards to their CS careers (Quille & Bergin, 2019). Research focused on CS1 has highlighted certain factors as having significant impact for women students. Included in this group of factors are self-efficacy (Beyer, 2014; Cheryan et al., 2017; Kinnunen & Simon, 2011; Lishinski et al., 2016; Tellhed et al., 2017; Wiedenbeck et al., 2004), confidence (Beyer et al., 2003; Biggers et al., 2008; Jones & Burnett, 2008; Lewis et al., 2017; Sax et al., 2017; Tafliovich et al., 2013; Wilcox & Lionelle, 2018), prior experience with computer science (Bernstein, 1991; Biggers et al., 2008; Cheryan et al., 2017; Margolis & Fisher, 2003; Tafliovich et al., 2013; Wang et al., 2015; Wilcox & Lionelle, 2018; Zarrett & Malanchuk, 2005), and prior knowledge of computer programming (Hagan & Markham, 2000; Jones & Burnett, 2008; Margolis & Fisher, 2003; Petersen et al., 2016; Wilson & Shrock, 2001). In the sections that follow, the aforementioned factors will be examined in greater detail. Other factors that play in a role in CS1 achievement or post-CS1 interest but were excluded in this study are discussed in appendices E and F.

2.1.2 Self-efficacy in CS1

Self-efficacy and student confidence for computer science tasks have frequently been shown to have a significant positive impact on student achievement in CS1 courses (Beyer, 2014; Cheryan et al., 2017; Kinnunen & Simon, 2011; Lishinski et al., 2016; Tellhed et al., 2017; Wiedenbeck et al., 2004). Bandura (1986, p.391) defined self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance." Students with low self-efficacy for a task demonstrate avoidance patterns related to that task, as opposed to students with high self-efficacy who show motivated pursuit of their goals. Self-efficacy has been found to have a positive correlation with student performance in math and science domains (Schunk & Pajares, 2005; Valentine, DuBois, & Cooper, 2004). This holds in computer science, where computing self-efficacy positively impacts student achievement outcomes (Ramalingam et al., 2004; Watson et al., 2014; Wiedenbeck, 2005). Lishinski et al. (2016) explored the relationship between students' computing self-efficacy and their performance in a CS1 course across multiple time points. The authors highlighted the presence of a feedback loop where negative experiences (such as lower than expected project scores) impacted self-efficacy which in turn negatively impacted future performances. For women, this occurred earlier in the course than it did for men. Similarly, Kinnunen and Simon (2011) showed that fluctuations in self-efficacy occurred when students made comparison with peers in CS1 courses. These changes reflected perceptions students maintained about the abilities of others and how well other students were understanding the material in comparison to their own understanding. Ultimately, changes to their own self-efficacy were reflected in student achievement. In addition to influencing achievement, self-efficacy has also been shown to be a predictor of career choice. In a variety of fields, students with high self-efficacy are more likely to pursue a career in that field (Lent, Brown, & Hackett, 2002). As in other fields, self-efficacy predicts interest in CS, particularly for young women (Sax et al., 2017; Weisgram & Bigler, 2006). Beyer (2014) found that women who had higher CS self-efficacy were more likely to take CS courses, and this effect was greater when their initial CS experience was positive. Research has shown existing gender gaps in CS with regards to self-efficacy, which impacts participation rates (Cheryan et al., 2017; Tellhed et al., 2017; Zarrett & Malanchuk, 2005). Wiedenbeck et al. (2004) offered a lack of prior programming experience as a potential explanation for low self-efficacy in CS1. Prior research has shown that women self-report lower self-efficacy for computer programming than men (Frieze & Quesenberry, 2015; Rubio et al., 2015), and that this did not align with their actual ability (Beyer, 2014).

In addition to studies of self-efficacy in CS1, some authors have looked at student confidence in computer science courses. Prior research has suggested that women have lower confidence than men in CS specifically (Beyer et al., 2003; Sankar et al., 2015; Wilcox & Lionelle, 2018) and STEM subjects more generally (Lewis et al., 2017). Lower levels of confidence have an impact on the rate by which students depart computer science programs. Reviewing four decades of computer science program data, Sax et al. (2017) found that academic self-confidence positively predicted intentions for continuing in computer science for CS1 students. Levels of student computing confidence have been connected to prior computing experiences; students with limited pre-university exposure to computing ideas expressed lower confidence for CS (Biggers et al., 2008; Jones & Burnett, 2008) and less interest in studying CS (Tafliovich et al., 2013).

The role of self-efficacy and student confidence on student achievement and interest have been well-established in the existing literature, but self-efficacy itself must be cultivated before it can have any positive impact for students. The following section addresses the factors that serve as precursors to the development of student self-efficacy.

2.1.3 Sources of Self-efficacy

Bandura (1997) proposed several factors that would contribute to the development of student self-efficacy for a task. These include mastery experiences, vicarious experiences, social persuasions, and physiological states. Enactive mastery experiences are personal experiences with a task that provide performance-based information to the individual; this type of information is viewed as the most valuable experience for positively increasing student self-efficacy (Palmer, 2006; Usher & Pajares, 2008). Vicarious experiences are those in which a model is observed having success at the task, particularly one who has perceived similarity to the observer. In academic settings, students can show increases in self-efficacy due to the performances of peers, particularly in situations where the student has limited personal experience within the domain (Bandura, 1997). Social persuasions take the form of support and encouragement from people that have important influence on one’s personal beliefs. It has been shown that social persuasions have a limited impact and are largely dependent on the individual's perception of that influential person as being knowledgeable in the domain in question (Bong & Skaalvik, 2003). Lastly, physiological states describe information from a student's natural physical systems, including elements of mood, emotion, and other physiological responses to engagement with the task. Physiological states have a limited but non-trivial impact on self-efficacy beliefs (Pajares, 1997).

The individual sources of self-efficacy have been studied in a limited way in computer science. Lin (2016) found that mastery experiences significantly predicted learning self-efficacy, but not computer self-efficacy or programming self-efficacy for a group of Taiwanese undergraduate students who had all taken at least one CS course. Further, the author found that social persuasion accounted for the highest amount of variance for learning self-efficacy and programming self-efficacy, while vicarious experience accounted for the most variance for computer self-efficacy. Lastly, the author noted that physiological states only played a predictor for the three self-efficacy measures for men and not women in the sample. More broadly in STEM subjects, there is evidence that the sources of self-efficacy are important for establishing self-efficacy, particularly for women and other underrepresented student groups (Hutchison et al., 2006; Zeldin et al., 2008; Zeldin & Pajares, 2000). Zeldin et al. (2008) interviewed successful men in STEM fields to identify formative experiences leading them to STEM careers, and then compared the outcomes to the results of an earlier study (Zeldin & Pajares, 2000) in which ten women had been interviewed. For the men in the study, the authors found that the emphasis was on successful mastery experiences in their field that led to greater confidence as they pursued their career goals. They noted to a lesser degree the impact of family, peers or teachers as motivation to engage with STEM either as models or sources of encouragement. On the other hand, the successful women in STEM from the Zeldin and Pajares (2000) study were much more reliant on social persuasions and vicarious experiences to build their self-efficacy in STEM. Hutchison et al. (2006) found that students in an introductory engineering course initially described influential experiences that influenced their confidence in the course as mastery experiences, but these turned to social comparisons when the course reached the midpoint of the semester. The authors noted that vicarious experiences were more salient for women in STEM courses with regards to their self-efficacy development. Sawtelle, Brewe, Goertzen, and Kramer (2012) also found that girls were more dependent on vicarious experiences and social persuasions for developing self-efficacy in a high school physics course. Usher and Pajares (2006) found that social persuasions were most important for academic self-efficacy (a measure addressing self-efficacy within multiple subject areas including computers) among 6^{th}-grade girls, while mastery and vicarious experiences explained most of the variance for boys of the same age. While their study looked at younger students, they did note similar effect sizes as found in Zeldin and Pajares' (2000) work with adult women. Other research has focused on differences between men and women and how information is interpreted in STEM courses. Men interpret average marks as evidence of success, whereas women with higher scores have interpreted the outcome as unsatisfactory (Zimmerman, 2000). As women are more likely to rely on the judgments of others in developing their STEM self-efficacy, this self-critical disposition can lead to lower overall self-efficacy despite feedback from course assessments telling the women that they are capable of success (Gorson & O'Rourke, 2020).

In addition to formal studies of the sources of self-efficacy, there have been several studies of CS1 that address factors that measure elements similar to the sources of self-efficacy.

These include social supports, perceived similarity, sense of belonging, and stereotype threat. More information about these factors can be found in appendix E.

2.1.4 Student Background with CS

It has been hypothesized that an important contributor to achievement in the CS1 course and overall interest in the study of CS are the computing experiences and knowledge of computing developed by students before beginning their undergraduate studies (Bernstein, 1991; Biggers et al., 2008; Cheryan et al., 2017; Hagan & Markham, 2000; Jones & Burnett, 2008; Margolis & Fisher, 2003; Petersen et al., 2016; Tafliovich et al., 2013; Wang et al., 2015; Wilcox & Lionelle, 2018; Wilson & Shrock, 2001; Zarrett & Malanchuk, 2005). Research has not provided clarity on best practices for measuring student background in computer science. There is evidence of measurement issues with existing assessments for CS1 (Xie et al., 2019), as well as concerns over the accuracy of self-reports of CS experience (Dochy et al., 1999). Due to research suggesting that women devalue their own CS experiences (Ashcraft et al., 2012; Beyer, 2014), and the potential for overlapping measurement between prior experience and mastery experience measures (Britner & Pajares, 2006; Schunk & Pajares, 2005), the focus of this study will be on an assessment measure of student prior knowledge of computer science and computer programming. Literature supporting the role of prior knowledge of computer programming is included in the section that follows. Further information about prior CS experiences can be found in appendices E and F.

2.1.4.1 Prior Knowledge of Computer Programming. In CS1, it has been shown that it is important to have prior knowledge of a programming language to be able to persist and attain high marks in the course (Hagan & Markham, 2000; Jones & Burnett, 2008). To this effect, Wilson and Shrock (2001) found that prior programming knowledge predicted midterm student grades in a CS1 course. Wilcox and Lionelle (2018) also saw a positive effect of between 6% and 10% on student grades based on prior programming knowledge, although the impact of prior programming knowledge in courses beyond CS1 diminished. Women often have less experience with computer science (Bernstein, 1991; Cheryan et al., 2017), although those women who do have prior programming knowledge have been shown to outperform male peers (Wilcox & Lionelle, 2018). The impact of prior programming knowledge on student achievement in CS1 and interest in CS can potentially be explained by considering self-efficacy beliefs. Students with prior programming knowledge are more likely to have high self-efficacy in CS1 (Wiedenbeck et al., 2004). As discussed earlier, self-efficacy is highly correlated with performance in CS, and students adjust self-efficacy based on feedback from their course experiences (Lishinski et al., 2016). Prior research on student self-concept has shown that some students hold a misconception that because CS1 is an introductory course, all students will begin the course having had similar exposure to computer programming (Petersen et al., 2016; Sands & Capobianco, 2020). This misconception impacts their self-beliefs during the course and can lead to student dropout (Tafliovich et al., 2013). There is no evidence of a direct effect of prior knowledge of computer programming on post-CS1 interest in taking more computer science courses.

2.1.5 Other Factors Impacting CS1 Achievement and Post-CS1 Interest

Beyond self-efficacy, the sources of self-efficacy, and prior knowledge of computer programming, there are several other factors that have been shown to impact CS1 achievement and post-CS1 interest in computer science. These include social support for CS (Beyer et al., 2003; Margolis & Fisher, 2003; Petersen et al., 2016; Wang et al., 2015; Zarrett & Malanchuk, 2005), sense of belonging (Cheryan et al., 2017; Lewis et al., 2017; Lewis et al., 2019; Sankar et al., 2015; Tellhed et al., 2017), utility value (Gaspard et al., 2017), and cost (Gaspard et al., 2017; Petersen et al., 2016). While there is evidence that these factors impact students in the CS1 course, their measure has potential overlaps with factors already being included in this study. More information about these factors can found in appendices E and F.

2.2 Theoretical Framework

Based on the important role of self-efficacy in predicting CS1 achievement and post-CS1 interest, Bandura's (1997) social cognitive theory (SCT) was used to guide several important choices in this study. Self-efficacy plays an important role in the conceptualization of SCT, particularly when considering agency, where student self-efficacy perceptions influence student choice of action and future behaviors. Based on these perceptions, students elect activities where they feel they can be successful and avoid activities where they do not feel confident of success. In these ways, SCT suggests a proactive approach to learning rather than a reactive one (Schunk et al., 2014). In developing a model for CS1 achievement and post-CS1 interest (figure 2.1), I utilized Bandura's conceptualization of self-efficacy and the antecedents to self-efficacy to describe the experiences of students in introductory computer science courses with regards to CS1 achievement and post-CS1 interest in studying computer science. This model positions self-efficacy in relation to other factors deemed important to understanding the persistence and achievement of students in CS1 as reviewed previously. A description of the elements of social cognitive theory and self-efficacy most critical to this study follows.

The social cognitive theory view of human adaptation and change involves a series of processes that are cognitive, vicarious, self-regulatory, and self-reflective in nature (Schunk & Pajares, 2005). Bandura (1997) described learners as participating in reciprocal interactions between personal factors (such as cognition and affect), behaviors, and their environment. Learners gather information from learning situations and use that to alter their future behaviors to successfully achieve their goals. In the context of this study, it was anticipated that CS1 students would use a variety of information sources within the context of the course and the broader CS environment to either engage more deeply with CS, or to avoid further study of computer science. The sources of information contributing to student self-assessments could include course feedback in the form of assessment grades, comparisons with peers, and implicit and explicit messaging from the course instructor and teaching staff. For students to be successful in their learning, Bandura (1997) described a need for an "exercise of control". If student interactions with the environment suggest that this control is not possible, it may negatively impact their ability to engage with CS and lead to their eventual departure from the major. Based on this view of CS1 students and their behaviors and interactions, the focus of the CS1 model was on both self-efficacy beliefs and the sources students use to generate these beliefs.

Self-efficacy is described as an individual's beliefs about their ability to be successful on a specific task (Bandura, 1986). In the context of this study, the specific task for which self-efficacy was measured was student ability to program a computer to solve problems. According to SCT, self-efficacy beliefs are developed based on a reciprocal process during which personal, behavioral, and environmental information is interpreted by the individual in the support of their ability to succeed on that task. Research on self-efficacy beliefs have shown that a learner's belief in their ability to succeed is more important than their ability with regards to their motivation to engage with the task (Schunk & Pajares, 2005). Prior research in CS and other STEM fields suggests that self-efficacy drives both interest and student achievement outcomes (Barker, et al., 2009; Beyer, 2014; Kinnunen & Simon, 2011; Lishinski et al., 2016; Ramalingam, et al., 2004; Schunk & Pajares, 2005; Tafliovich et al., 2013; Valentine et al., 2004). While self-efficacy has been shown to have a positive impact on achievement, it is important to note that high self-efficacy for a task will not overcome a lack of ability, low value for the task, or negative expected outcomes (Schunk, 1995). This suggests that the CS1 model should include some measures of ability and value in order to capture expected variation in CS1 achievement and post-CS1 interest in future study of CS. Bandura (1997) proposed four sources of information that influence self-efficacy: enactive mastery experiences, vicarious experiences, social persuasions, and physiological states. It has been shown that mastery experiences, the interpreted results of previous performances, have the most influence on self-efficacy (Bandura, 1997; Schunk & Pajares, 2005). Due to the strength of this effect, occasional deviations from pattens of success or failure do not have a significant impact on student self-efficacy (Schunk & Pajares, 2005). In situations where learners have limited prior experience or doubts about their abilities to succeed on a task, vicarious learning experiences become an important source of self-efficacy. This is often done through models who serve as a diagnostic of the learner’s ability. These models are chosen based on perceived similarity to the learner in both background characteristics and due to alignment between the goals of the learner and the model (Bandura, 1997). Models and important social relations can have a further impact through meaningful feedback and support (Bandura, 1997). Physiological and emotional states have been shown to have impact on self-efficacy (Pajares, 1997), with students showing high self-efficacy interpreting an affective response to engagement with a task as motivation to overcome a challenge, and students with low self-efficacy interpreting the same information as evidence that a negative outcome is unavoidable (Bandura, 1997). It has been shown that there is variation in the influence of the sources of self-efficacy for math and science tasks dependent on gender (Britner & Pajares, 2006; Lent et al., 1996; Sax et al., 2017; Usher & Pajares, 2006; Zeldin et al., 2008). For women in STEM subjects, it is believed that vicarious experiences and social persuasions play a much greater role, whereas men often cite mastery experiences as being influential (Britner & Pajares, 2006; Lent et al., 1996; Usher, 2009).

Taking the components of SCT and self-efficacy in consideration, the CS1 model was structured to capture student perceptions about their own background, the context of the CS1 course, and the interactions between these elements as they impacted student self-efficacy, CS1 achievement, and post-CS1 interest in further study of CS. This model reflects current beliefs about the role of specific factors on students in CS1 courses while incorporating broader beliefs about the mechanisms underlying self-efficacy. In addition to the use of SCT and self-efficacy, this model echoes prior work in motivation in mathematics that relates prior knowledge, self-efficacy, and student interest (Lent et al., 1994). Further decisions regarding the construction of the model will be addressed based on limitations and constraints of the study.

2.3 Goals of the Study

The main goal of this study was to test a model including factors established by prior research to be significant to CS1 achievement and post-CS1 interest based on gender. These factors include: prior experience with computer programming, self-efficacy, and the sources of self-efficacy (Beyer et al., 2003; Beyer, 2014; Biggers et al., 2008; Cheryan et al., 2017; Hagan & Markham, 2000; Jones & Burnett, 2008; Kinnunen & Simon, 2011; Lewis et al., 2017; Lishinski et al., 2016; Margolis & Fisher, 2003; Petersen et al., 2016; Sax et al., 2017; Tafliovich et al., 2013; Tellhed et al., 2017; Wiedenbeck et al., 2004; Wilcox & Lionelle, 2018; Wilson & Shrock, 2001). Differentiating this approach from previous attempts to generate a model explaining outcomes in the CS1 course is the consideration of self-efficacy and the sources of self-efficacy as the focus of student motivation and background as they relate to CS. Additionally, this research aims to understand the role that these factors play for students in CS1 when considering gender. Results from this study could inform the development of interventions and practices aimed at encouraging women in computing.

Research questions addressed in this study include: (1A) How do prior knowledge of computer programming, the sources of self-efficacy, and self-efficacy for computing predict CS1 achievement and student intentions to continue study in CS? (1B) How does self-efficacy mediate the relationship between student prior knowledge of computer programming and achievement in CS1? (1C) How are those relationships moderated by gender? (2) How does feedback in the form of student grades impact intention to continue in CS when considering gender as a moderating factor?

CHAPTER 3: METHODOLOGY

3.1 Design of the Study

The main goal of this study was to understand the relationships between factors related to CS1 achievement and post-CS1 interest as moderated by gender. The factors selected for inclusion were prior knowledge of computer programming, mastery experiences, vicarious experiences, social persuasions, physiological states, and self-efficacy. From these factors, self-efficacy has been found to have a significant relationship with CS1 achievement and post-CS1 interest. Despite the prevalence of self-efficacy in prior computer science education research, there is little work focused on the antecedent causes to self-efficacy development. Using self-efficacy as a core component of the model provided greater clarity for establishing the relationships between the dependent and independent factors used in the study.

The development of the CS1 model in this study was based on prior work focused on self-efficacy in STEM domains (Beyer, 2014; Britner & Pajares, 2006; Cheryan et al., 2017; Kinnunen & Simon, 2011; Lishinski et al., 2016; Lopez et al., 1997; Pajares, 1997; Tellhed et al., 2017; Usher & Pajares, 2006; Wiedenbeck et al., 2004; Zeldin et al., 2008). As Bandura (1986, 1997) hypothesized in his social cognitive theory, a student derives self-efficacy beliefs from a variety of sources which include mastery experiences, vicarious experiences, social persuasions, and physiological states. Alongside these sources, the model included prior knowledge of computer programming as a predictor of CS1 achievement and tested whether this effect is mediated by self-efficacy similar to work by Lopez et al. (1997). Expectation of a mediating effect by self-efficacy in the relationship between prior knowledge of programming and CS1 achievement comes from prior research on self-efficacy in computer science (Petersen et al., 206; Sands & Capobianco, 2020; Wiedenbeck et al., 2004). Lastly, the model includes two dependent variables, CS1 achievement and post-CS1 interest in computer science. Relationships between prior knowledge of programming and self-efficacy were proposed with CS1 achievement as evidence suggests that these factors play a role for students in the CS1 course (Kinnunen & Simon, 2011; Lishinski et al., 2016; Ramalingam et al., 2004; Watson et al., 2014; Wiedenbeck, 2005; Wilcox & Lionelle, 2018). Additionally, it has been suggested that self-efficacy in many fields including CS impacts student interest and career commitment (Beyer, 2014; Sax et al., 2017; Weisgram & Bigler, 2006). One final relationship was included between CS1 achievement and post-CS1 interest to understand more clearly how course feedback impacts student interest. The model as described can be found in figure 2.1.

The choice to use gender as a moderator stems from prior research suggesting that the formative experiences of women and men differ as they relate to computer science (Margolis & Fisher, 2003; Sadler et al., 2011; Sheard et al., 2008; Zarrett & Malanchuk, 2005; Zingaro, 2015). Prior work has focused on differences for women in their early experiences with computing, support given for pursuit of computer science careers, and exposure to computing before college coursework in CS (Bernstein, 1991; Beyer, 2014; Beyer et al., 2003; Cheryan et al., 2017; Margolis & Fisher, 2003; Wang et al., 2015; Wiedenbeck, 2004). These differences in turn could impact the development of self-efficacy, achievement in CS1, and interest in computing over the long-term. This study was designed to test hypotheses regarding specific differences in the relationships between independent and dependent factors for women so that appropriate interventions can be designed to address these factors in future iterations of CS1 courses.

Quantitative data in this observational study was collected from four cross-sectional surveys of CS1 students during two semesters of introductory computer science across two universities. The motivational factors and the student prior knowledge factor were collected using validated instruments derived from prior work. Each of the measures and instruments are discussed in greater detail below. Some of the instruments had not been used in prior studies of computer science and were tested using confirmatory factor analysis prior to analysis. The analysis plan for this data is focused on correlational research. Specifically, the relationships between factors were studied in a structural regression using structural equation modeling. Groups were used to understand the role of potential moderators in the study. The data analysis is discussed in more detail in the sections that follow.

3.2 Study Participants

The population of interest in this study was first-year undergraduate computer science students. Students were recruited from CS1 courses at two universities on a voluntary basis, with the offer of course extra credit for their participation. To obtain the sample, a notice went out in each class from the instructor regarding the study. Following this notice, I was granted access to the course message boards for both courses where future communication regarding the study would occur. Before each portion of the data collection, I submitted a message board post with information about the study and access to the consent form for students to review before agreeing to participate. Students must have completed all four components of the study to be included in the final sample used in this study. All data collection occurred via completion of online surveys and an online pre-test using Qualtrics software. Data collection procedures were approved by both university institutional review boards.

The sample of the study consisted of 182 total students, of which 108 (59.3%) identified as men, 70 (38.5%) identified as women, 3 (1.6%) as non-binary, and 1 (0.5%) did not wish to identify by gender. When considering the CS1 model by gender, the non-binary and non-identifying students were excluded due to their small proportion of the overall sample. The race and ethnicity of the sample included 13 (7.1%) Hispanic or Latinx students, 90 (49.5%) Asian students, 1 (0.5) Black or African American student, 60 (33.0%) White students, 10 (5.5%) students reporting two or more races, and 8 (4.4%) students not wishing to report race or ethnicity. From these students, 117 (64.3%) attended university A and 65 (35.7%) attended university B. Students in the sample also reported information regarding their major. The study was intended for computer science majors, but upon reviewing the data it was determined that data science majors, students electing a CS minor, and students that had yet to declare their major should also be included. There were 143 (78.6%) computer science majors, 30 (16.5%) data science majors, and 9 (4.9%) students mentioning a CS minor or having no declared major. Participants by gender distributed across universities equally but there was some minor variation in student major. More data science majors identified as women (18) than men (11). I do not believe this had an impact on the outcomes but diverged from the expectation. A group of six sampled students were removed due to extreme values and were not included in the summary data above (see section 3.1.3.1 for more details). All demographic and school information can be reviewed in table 3.1.

The sample came from a population of 4 total CS1 course offerings at 2 different universities. At university A, the total population of the Fall semester course was N=570 students, of which the sample (N=102) represented 17.9% of the student population. The population of this course was reported as N=450 men and N=118 women. For the Spring semester, the university A student population was N=610 students with a sample (N=16) that represented 2.6% of those students. The population of this course during Spring semester was reported as N=478 men and N=129 women. At university B, the CS1 course had N=522 students, of which the sample (N=56) represented 10.7% of the student population. In this course, there were N=407 men and N=115 women. The Spring semester offering at university B contained N=501 students, of which the sample (N=11) represented 2.2% of the population. During the Spring there were N=398 men and N=103 women taking the course. At both universities, the Spring semester featured a higher percentage of non-majors in the course, which resulted in a lower rate of eligible participants for this study. Overall, the course population across universities was N=2203 students, and the sample represented 14.8% of the enrolled students.

3.3 Independent Measures

The independent measures collected for use in this study included self-efficacy, mastery experiences, vicarious experiences, social persuasions, physiological states, and prior knowledge of computer programming. Details for each of these measures are included in the sections that follow. For details about measures that were collected but removed from use in the study including the justification for their exclusion see appendices E and F.

3.3.1 Self-efficacy

To measure the self-efficacy for tasks related to introductory computer science, I used the Patterns of Adaptive Learning Scale (PALS) instrument (Midgley et al., 2000). Many different instruments have been used to collect information about student self-efficacy in studies that focus on student motivation (e.g., General Self-efficacy Scale, Collegiate Academic Self-efficacy Scale, Motivated Strategies for Learning Questionnaire). The PALS was selected because it has shown good reliability while using a limited number of survey items to capture student self-efficacy beliefs. The PALS instrument has six Likert-scale items (1 – Strongly disagree to 5 – Strongly agree) with a Cronbach's alpha reliability measure of α=.78.

Changes were made to the items to re-frame the questions to address computer science but were otherwise left in their original form. The observed reliability for the self-efficacy scale using the sample of students in this study was α=.92. Items as they were re-written can be found in appendix C.

3.3.2 Sources of Self-efficacy

To measure the sources of self-efficacy as proposed by Bandura (1997), I modified the Middle School Mathematics Sources of Self Efficacy Scale (Usher and Pajares, 2009) to address computer science. The original scale uses 6 items for each source of self-efficacy (mastery experiences, vicarious experiences, social persuasions, and physiological states) measured using a 5-point Likert scale. One item from the mastery experience statements is reverse-coded and was manipulated after data collection to align properly with the other items. Each of the sub-scales had good internal reliability as measured by Cronbach's alpha: mastery experience (α=.88), vicarious experience (α=.84), social persuasions (α=.88), physiological states (α=.87). Changes made to these items re-framed the questions to address computer science rather than mathematics but were otherwise left as proposed by the scale's original authors. The observed reliability for the scales in this study are as follows: mastery experience α=.89, vicarious experience α=.78, social persuasions α=.90, and physiological states α=.94. Items as they were re-written can be found in appendix C.

3.3.3 Prior Knowledge of Computer Programming

To measure students' prior computer programming knowledge, I used a validated assessment from prior work exploring measures of programming ability. The assessment selected was Parker et al.'s (2016) Secondary Computer Science 1 (SCS1) assessment. This assessment covers the following topics: programming basics, conditionals, definite/for loops, indefinite/while loops, logical operators, arrays, recursion, function parameters, and function return values. In covering these topics, the SCS1 attempts to measure ability in computer science relevant to the CS1 course in a way that does not bias for original language of study. The original assessment had 27 multiple choice items for which the authors reported a Cronbach's alpha of α=.59. This version of the test was intended to be taken over a single 60-minute period. Further work on the SCS1 by Xie et al. (2019) used item response theory to analyze the assessment using a new sample of 507 undergraduate computer science students. The authors identified 4 items that do not accurately capture prior knowledge of computer programming and these items have been excised from the assessment (leaving 23 items remaining). With the changes made by Xie et al. (2019), the reliability in their sample was α=.723.

To implement the pre-test in this study, I elected to make several changes. I was unable to enforce a time limit for the volunteer participants given the online distribution of the assessment. Due to my concerns that the length of the pre-test would lead to a lower response rate, I elected to reduce the number of questions from 23 to 13 and suggested the students take the pre-test in one 30-minute period. I reviewed the SCS1 items and identified questions representing core concepts in the CS1 course and eliminated redundant items. For example, the original pre-test featured 6 items covering the behavior of functions and passing parameters between the main program and a subprogram. I selected one item that addressed this topic most clearly using my best judgment in conjunction with data provided by Xie et al. (2019) regarding the difference to the Cronbach's alpha level if that item were to be removed from the exam. The final pre-test given to students in this study had 13 items from the original SCS1 assessment. Using the aforementioned information about impact on Cronbach's alpha due to item removal from Xie et al. (2019) the reliability of the assessment with the reduced number of questions was calculated as α=.671. This number is lower than the minimum acceptable rate and further analysis of the pre-test was required.

After collecting student data, I completed a reliability analysis on the pre-test using the following methods. I first ran a confirmatory factor analysis to verify the unidimensionality of the items referring to computer programming knowledge. The model fit for all 13 items was $\chi^2(65) = 76.773$ ($p=0.151$) which means that we would not reject the null hypothesis that the data fit the model perfectly. The comparative fit index (CFI) was 0.847 which is less than the desired level (0.9). However, the root mean square error of approximation (RMSEA) was 0.024 which is below the 0.1 threshold. Several questions loaded poorly on the programming knowledge latent factor, and through an iterative process I generated several additional models to try and improve the overall fit. Removing three questions (9, 17, and 23), and allowing two error covariances to freely estimate (questions 1 and 14, and questions 14 and 26) resulted in significantly better fit. The model fit for the 10-item pre-test was $\chi^2(34) = 36.771$ ($p=0.342$), the CFI was 0.970 and the RMSEA was 0.019 which both suggest good fit. Based on these model fit statistics, I believed that the 10 questions reasonably capture the latent factor of computer programming knowledge. This model used N=239 student participants (including both computer science and non-CS students from the voluntary participants for the study) and was estimated using maximum likelihood with IBM Amos software. Based on the limited number of questions in the pre-test, Cronbach's alpha is not a good measure of internal reliability. To proceed, I computed the Kuder-Richardson Formula 20 statistic (KR-20) which can be used to find inter-item consistency (Kuder & Richardson, 1937). Computing this value for the 10-item pre-test resulted in a value of KR-20=0.88 which suggests good internal reliability.

The 13-item version of the SCS1 instrument was included in the initial survey of students during week 1 of the CS1 course, and the score on this pre-assessment removing questions 9, 17, and 23, was used to represent student prior knowledge of computer programming. Items for this assessment do not appear in the appendices as the assessment authors have asked that they not be shared without permission.

3.3.4 Factor Analysis of Measures

For the study of CS1 students, the intention was to study them in relation to one another. This necessitated the use of confirmatory factor analysis for all the study factors included in one model, and analysis of discriminant and convergent validity for the measurement model. Before reporting on the results of the CFA, I will discuss the preliminary work completed to establish the needed sample size to obtain adequate power.

For the CFA, a power analysis was conducted using a Monte Carlo approach (Muthén & Muthén, 2002). This method was selected over the Satorra and Saris (1985) method due to concerns about the normality of the underlying data. For this approach, model parameters (factor loadings and error variances for each observed factor, and correlations between latent factors) serve as inputs. These values were in most cases based on established values from prior research. Using 2500 samples and a starting estimated sample size of N=100, average values and standard deviations were computed for each parameter. This information was used to estimate the sample size needed to reach a power of .8 as recommended by Cohen (1988). Additional important factors that were analyzed to ensure that the sample size calculations were appropriate were parameter and standard error biases below 10% or below 5% for parameters of focus in the power analysis, and coverage between .91 and .98 (Muthén & Muthén, 2002). The results of the Monte Carlo simulations suggested a sample of N=200 was adequate for performing the CFA with 66 free parameters to attain power of at least .814 for all parameters of interest in the study. Mplus code for the Monte Carlo simulation appears in appendix D.

During the process of conducting a CFA of the measurement model it was discovered that the vicarious experience items showed low convergent validity. The average variance explained (AVE=0.518) suggests that the indicators for vicarious experience did not correlate well with one another inside the latent factor. Several of these items had been shown to have poor factor loadings suggesting that they may not be measuring what was intended. To further explore potential explanations for the poor loadings and poor convergent validity, I reviewed a dimensional reduction using the indicators for the sources of self-efficacy, self-efficacy, and cost using promax rotation (note: details of the measurement issues related to cost and subsequent removal of these items can be found in appendix E). Removing items with poor loadings or with troublesome cross-loadings reduced the measures for vicarious experience to just two of the original 6 items.

To test the measurement model, I ran a confirmatory factor analysis and evaluated it using suggested fit indices. For absolute fit, the chi-square and SRMR indices were examined. The chi-square test statistic is sensitive to large sample sizes, so other statistics were used to verify findings, or to have greater confidence given concerns over sample size related inflation of the chi-square statistic. For parsimonious fit, the RMSEA index was used. For comparative fit, CFI was used. A confirmatory factor analysis including the sources of self-efficacy and self-efficacy with the adjusted indicators provided adequate fit: $\chi^2(285) = 407.763$, $p<.001$; CFI=.953; RMSEA=.049; SRMR=.051. This model can be found in figure 3.1. Mplus code for this analysis and corresponding output can be found in appendix D. It should be noted that mastery experiences and self-efficacy were highly correlated and thus the cross-loading between these items was ignored in the analysis of this data.

3.4 Dependent Measures

The two dependent measures in the study were the student grade in the CS1 course and the student's self-reported intention to continue studying computer science beyond CS1. These measures were used to capture elements of CS1 achievement and post-CS1 interest respectively. Details for each of these measures are included in the sections that follow.

3.4.1 Academic Achievement in CS1

Academic achievement in CS1 was measured using overall student grade in the CS1 course based on a standard 4-point GPA scale. The grade point scale used is shown in table 3.3. For each course, the total student score was calculated differently. At university A, the overall grade was computed based on 45% examination score, 45% computer programming projects, and 10% short programming exercises. The first midterm made up 10% of the overall grade, the second midterm 15%, and the course final 20%. At university B, the overall grade was computed based on 30% examination score, 50% project score, 15% homework score, and 5% quiz score. No specific grade distributions were provided for the exams in the course at university B.

3.4.2 Post-CS1 Interest in Computer Science

To measure a student's interest in continuing in computer science after the CS1 course, I used a five-point ordinal scale of my own creation. Existing measures such as the STEM Career Interest Scale (STEM-CIS) and the STEM Career Interest Questionnaire (STEM-CIQ) were considered for use, but neither adequately captured immediate student intentions with regards to their coursework or were deemed to be age-inappropriate (Sadler et al., 2012; Tyler-Wood et al., 2010). The interest scale that I created was designed to capture the degree to which the student felt confident that they would take more courses in computer science. The wordings of the stems for this item were chosen to allow students to specify their interest in taking further courses and a subjective degree to which they felt that they would be likely to follow through on that choice. Items on this scale can be found in appendix C.

3.5 Data Collection Procedure

Participants for this study were recruited during the first week of their CS1 class at both universities. Course instructors mentioned the opportunity to volunteer for the study during early class sessions and asked teaching assistants to do the same within student lab sections. Instructors told students that volunteer participants would receive extra credit if they completed all parts of the study. Links to the consent form were made available through the course discussion forums and were also provided to the course instructors to be included in their course materials. Throughout the term, I frequented the discussion forum to send participation reminders and to respond to participant questions.

Data was collected via Qualtrics survey at three distinct points in the CS1 courses as shown in figure 3.2. The first survey included the study consent form and demographic information. This survey was made available to students through the course discussion forum and on the course content management system. When students completed the first survey and consent form, they were emailed a link to the pre-test of computer programming knowledge. The second survey was automatically sent to students at approximately week 8 of the CS1 courses. This was estimated to be a point in the semester after students had completed and received feedback on one exam and several computer programming projects. The second survey included items addressing student self-efficacy and the sources of self-efficacy. The final survey was delivered at the end of the course, and collected information about future intentions for study of CS. At this point, data was linked across surveys and student volunteers that participated in all parts of the course were noted to course instructors so that they may receive credit for completing all parts of the study.

3.6 Data Analysis

To answer the research questions, the following approach to analysis was employed. In addition to reporting descriptive, correlational, and inferential statistics for the independent measures by gender, I used structural equation modeling (SEM) to conduct multi-group structural regression on the proposed model with gender as a moderator. The sections that follow discuss the steps taken to prepare the data for analysis, the preliminary analysis used to examine the underlying assumptions regarding the data set, and the process of conducting SEM to address the research questions of the study.

3.6.1 Data Preparation

After the data collection process was completed, data from the consent form, three surveys, and pre-test of computer programming knowledge were downloaded from Qualtrics. The student grade data was requested from the course instructors and delivered via comma separated values (CSV) file format. Prior to merging the imported files, duplicate records were removed. These duplicate records were often cases in which a student had submitted information using two different email accounts, or places where they had mistakenly submitted an incomplete entry before completing the survey on a second attempt. In all cases, a determination was made regarding which record provided the most information for analysis and other records for that participant were deleted. Once each data file contained only a single record for each participant, the files were merged using a combination of individual identifiers that were unique to the individual. As a final step before merging individual files, the pre-test data was “scored” so that test scores could be treated as a single input in the final data analysis. This scoring was done based on the outcomes provided by the SCS1 pre-test designers.

At each stage of the merger, records were removed where individuals had completed the first survey but not surveys that followed. Over both semesters, the initial survey and consent form had 556 records. After merging with the other files, the number of records was reduced to 357 complete records. Once all files were joined, including the student grade reports from both universities, final steps were taken to complete the data records for analysis. Student records were associated with their corresponding university to allow for the school attended to serve as a check for measurement invariance in future portions of the analysis. Then, all identifying information was removed and replaced with four-digit numbers used as unique identifiers. The final step of reducing the data involved eliminating records for students that were not computer science majors, data science majors, students with a computer science minor, and undeclared students. Doing this, the resulting data set had 188 complete records.

To complete the preparation of the data for analysis, I recoded each variable to allow for quantitative analysis and easy reference using numerical codes. Most of the items requiring recoding were Likert-style questions that used the same scale. Gender was recoded to simplify multigroup analysis for the structural regression model. Race and ethnicity were recoded into one variable using federal guidelines for handling multiple racial and ethnic identities. Grades were then translated to a 4-point GPA scale for use in analysis where independent variables would be used to predict academic achievement for students. The final step was to decide on how to handle missing data values. An initial analysis of missing values suggests that there were no distinct patterns of missing values among participants. All missing values from survey items were replaced with the value “-999” which was outside the range of expected values.

3.6.2 Measurement Invariance

To test for measurement invariance based on student gender, I conducted three tests to establish configural, metric, and scalar invariance. The configural invariance test shows whether the factor structure is adequate when tested without constraints across the gender groups. The metric invariance test shows whether both groups have comparable factor loadings for the measurement model. The scalar invariance test shows whether indicator intercepts are equivalent across groups. If the model fit suggests adequate fit across the man and woman gender groups, future analysis based on this measurement model can be done with confidence that it captures what is intended to be captured for men and women equivalently.

To begin testing for measurement invariance, I first checked the configural invariance of the model using one group for students identifying as men and one group for students identifying as women. Running a model with freely estimated parameters for men and women produced the resulting CFA model fit statistics: $\chi^2(570) = 816.353$, $p<.001$; CFI=.914; RMSEA=.05; SRMR=.0675. The fit of this model is adequate for two groups with freely estimating parameters (Kline, 2016). Next, I restricted the factor loadings to be equal to test for the metric invariance across groups. The model fit with these constraints was $\chi^2(596) = 840.731$, $p<.001$; CFI=.915; RMSEA=.048; SRMR=.0855. A chi-square difference test between models ($\chi^2(26) = 24.378$, $p=.554$) suggests no significant difference and thus I assume that the groups are not different at the model level. Finally, I constrained the structural covariances to test for scalar invariance. The model fit with these constraints was $\chi^2(584) = 833.35$, $p<.001$; CFI=.913; RMSEA=.049;

SRMR=.0768. A chi-square difference test between models ($\chi^2(14) = 16.997$, p=.256) suggests no significant difference and thus I assumed equal indicator intercepts. Thus, I proceeded with my analysis assuming measurement invariance across gender groups given the data collected in my study.

3.6.3 Missing Data

A usual concern when collecting and analyzing data is missing data and the patterns of missing data that would suggest a systematic reason that participants may not have answered certain items in a survey tool. For this study, the participants were asked to respond to 70 items related to the latent factors described in previous sections. In the Fall semester, across both universities, there were 162 respondents who completed all four components of the study. In the Spring semester there were only 26 respondents who completed all four components. I will review missing data patterns for each survey individually and look specifically across items related to the same latent factor to ensure that missing data patterns are not related to specific factors of interest.

In the first survey, there were nine items collecting data about past experiences with computer science and computer programming, and one item asking about the student’s current major. There were other questions asking for demographics information, but these items had options for non-response and did not show any signs of missing data. Full participants in this study answered the items on the first survey completely; no patterns of missing data were present.

The pre-test of student knowledge of computer programming contained 13 questions. Due to the time limit for the pre-test (participants were expected to complete the pre-test in 30 minutes) I chose to treat missing values as skipped questions due to a lack of computer programming knowledge. Thus, for this part of the data collection I did not evaluate missing data.

The second survey contained forty-six statements regarding student beliefs and motivations with regards to computer science. Across all participants there were four items that had any missing data. No individual student participant had more than one missing item and looking across all data there was no pattern to this missing data.

The third survey, given at the end of the semester, had a single item regarding post-CS1 interest in continuing in computer science courses. There was no missing data for this survey.

The limited amount of data in the study that is incomplete was missing completely at random (MCAR). With no patterns in the missing data, this allowed for the use of maximum likelihood estimation in the structural equation modeling process.

3.6.4 Structural Regression

The aim of the multi-group structural regression using SEM was to determine if the model was different for women in specific ways than it was for men in the introductory computer science course. The steps for data analysis followed the approach described by Kline (2016) including specification, identification, data collection, estimation, evaluation, and the optional model re-specification phase.

To have significant power for the analysis via structural regression the desired sample was computed using a series of Monte Carlo simulations using Mplus software. Prior to using the simulation approach, several estimates were made based on existing theory and suggested rules-of-thumb. One statistic that can be calculated to provide an estimate for path analysis is Hoelter's critical N which gives the minimum number of data points required for critical alpha levels (Hoelter, 1983). For the path analysis in this study, the value of Hoelter's statistic was given as

N=133 for the .05 level and N=141 for the .01 level. One frequently cited rule-of-thumb suggests a ratio of 10 to 1 for each model parameter in the study. Based on the 83 free parameters that were in the identified model, this suggests a sample size of N=830 students. This was not outside the realm of possibility given the collection of multiple semesters data but given the effects of the global COVID-19 pandemic it was deemed impractical. Using the Monte Carlo approach to estimate sample size given a minimum power of .8 for the structural paths in the revised model with 83 free parameters, a sample of N=450 was found to be sufficient. In this case the power for all parameters of interest in the study was a minimum of .844. Mplus code for the Monte Carlo simulation appears in appendix D.

To evaluate the structural regression model, I used indices of absolute fit (chi-square, SRMR, RMSEA), and comparative fit (CFI and TLI). These measures will be considered in combination to ensure adequate fit. Due to the sensitivity of chi-square to the sample size, I also used SRMR (< .08) and RMSEA < .08 to check for absolute fit (Hu & Bentler, 1999). For comparative fit, CFI ≥ .95 and TLI ≥ 0.9 will be considered adequate. All paths will be examined for significance and decisions will then be made about respecification. In addition, based on gender as a moderating factor, paths will be tested again in comparison with the base model using the chi-square difference test with the Satorra-Bentler correction.

CHAPTER 4: RESULTS

4.1 Analysis of the Sample

4.1.1 Descriptive Statistics

For the factors in the study, I will report information both for the total sample and by gender. Means and standard deviations for all measures appear in table 4.1. Additional statistics for other measures not used in the CS1 model can be found in appendix F.

The mean self-efficacy response for the total sample was 4.031 (SD=0.77), for men was 4.188 (SD=0.667), and for women was 3.793 (SD=0.85). The difference between women and men was significant for self-efficacy (t=3.465, df=176, p=.001) at the 1%-level.

The four sources of self-efficacy include mastery experiences, vicarious experiences, social persuasions, and physiological states. The mean mastery experience response for all participants was 3.612 (SD=0.82), for men was 3.80 (SD=0.688), and for women was 3.305 (SD=0.919). The average mastery experience was significantly higher for men than for women at the 1%-level (t=4.071, df=176, p=0). The mean vicarious experience response for the total sample was 3.665 (SD=0.836), for men was 3.759 (SD=0.769), and for women was 3.536 (SD=0.91). The difference between women and men was not significant for vicarious experiences (t=1.761, df=176, p=.08) at the 5%-level. The mean social persuasions response for the total sample was 3.116 (SD=0.839), for men was 3.228 (SD=0.802), and for women was 2.943 (SD=0.875). The difference between women and men was significant for social persuasions (t=2.238, df=176, p=.026) at the 5%-level. The mean physiological states response for the total sample was 2.414 (SD=0.98), for men was 2.30 (SD=0.909), and for women was

2.633 (SD=1.057). The difference between women and men was significant for physiological states (t=-2.255, df=176, p=.001) at the 5%-level.

Student prior knowledge of computer programming was measured using a 10-question pre-test. The mean pre-test score for the sample was 4.429 (SD=2.78). For women in the sample the mean pre-test score was 3.914 (SD=269) which was not significantly lower than for men in the sample (M=4.713; SD=2.819; t=1.881, df=176, p=.062) at the 5%-level.

The mean post-CS1 interest response for the total sample was 4.69 (SD=0.644), for men was 4.80 (SD=0.525), and for women was 4.53 (SD=0.775). The difference between women and men was significant for post-CS1 interest (t=2.748, df=176; p=.007) at the 1%-level.

CS1 achievement was measured by course GPA using a standard 4-point scoring system (0.0-E, 4.0-A). The mean GPA score for the total sample was 3.465 (SD=0.999), for men was 3.56 (SD=0.906), and for women was 3.301 (SD=1.136). The difference between women and men was not significant for CS1 achievement (t=1.682, df=176, p=0.094) at the 5% level.

4.1.2 Correlational Statistics

Pearson's correlations were calculated between all dependent and independent factors in the study. In addition, separate statistics were calculated for students identifying as women and men in the study. Values for the total sample can be found in table 4.2.

Several strong relationships stood out in the analysis of the correlations between dependent and independent variables. For all students, CS1 achievement had a significant positive relationship with mastery experience (r=.485, p<.001), post-CS1 interest (r=.468, p<.001), self-efficacy (r=.369, p<.001), prior knowledge of computer programming (r=.330, p<.001), and social persuasions (r=.247, p=.001). Strong and moderate positive relationships with post-CS1 interest included mastery experience (r=.478, p<.001), self-efficacy (r=.473, p<.001), vicarious experience (r=.368, p<.001), prior knowledge of computer programming (r=.261, p<.001), and social persuasions (r=.310, p<.001).

Of these effects, there were specific instances where the relationships differed between women and men significantly. For the correlations between CS1 achievement and post-CS1 interest there was a sizable difference between men (r=.236) and women (r=.646) which resulted in a statistically significant difference (Z=-3.38, p=.001) at the 1%-level. When looking at post-CS1 interest there were significant differences between men and women regarding self-efficacy (r_M=.191, r_W=.646, Z=-3.68, p<.001) and mastery experiences (r_M=.237, r_W=.615, Z=-3.04, p=.002) at the 1%-level. All differences by gender can be seen in table 4.3.

4.1.3 Multivariate Assumptions

The process of conducting structural equation modeling is essentially simultaneous equation modeling similar to multiple regression. To proceed with the analysis of the data using structural equation modeling and structural regression, several multivariate assumptions must be met for the data in the study. These include the absence of multicollinearity of observed data and univariate and multivariate normality for dependent variables. In addition to checking these basic assumptions, I identified outliers and other influential data points that may exacerbate these issues.

4.1.3.1 Outliers and Influential Data. Some data points will exert undue influence on the outcomes of the study if they are far outside of the rest of the data. These can include outlying data points and those exhibiting high leverage. There are several measures of influence and leverage, but in this study, I chose to use the Mahalanobis distance statistic to determine if there were influential points. The distance values were calculated for the two dependent variables and examined to determine which points should be removed from the study for further calculations.

To use the Mahalanobis distance statistic, it is recommended that the computed values be compared with the chi-square test statistic at the .1%-level with degrees of freedom equal to the number of observed variables (Kline, 2016). For CS1 achievement, the Mahalanobis distance was calculated using the four sources of self-efficacy, self-efficacy, and prior knowledge of computer programming. The critical value for a chi-square distribution at the .1%-level with 6 df is 22.458. Based on this calculation, six data points appeared as outliers and were flagged to be examined more closely. Post-CS1 interest is predicted using the same variables as achievement, but also includes achievement as a predictor as well. The critical value for a chi-square distribution at the .1% level with 7 df is 24.322. Six data points appeared as outliers and were flagged to be examined more closely. These six data points were extreme ones for each of the dependent variables and have undue influence on the rest of the data. They were removed from further analysis. Plots of Mahalanobis distances for each dependent variable can be found in figures 4.1 and 4.2.

4.1.3.2 Multicollinearity. The assumption of linearity states that all relationships between dependent and independent factors are, in fact, linear in nature. The violation of this assumption eliminates the possibility of conducting any regression analysis of paths in the models being explored in this study. To ensure linearity, I plotted residuals for each linear relationship proposed in the model against the predicted residual values and checked for patterns suggesting a non-linear relationship. These plots can be seen in figures 4.3 - 4.11. Based on these plots, the assumption holds that all paths being explored in this study are linear. All plots show random scatter of residuals about zero.

When observed data are highly related to one another this suggests the potential for multicollinearity. To determine if observed data showed strong overlap, I reviewed all correlations between the indicators of the latent factors. Note that these are not the correlations of latent factors, but correlations between the indicator items that are used to represent the latent factors. Correlations between indicators higher than $r=.85$ were flagged as demonstrating strong overlap. No pair of indicators was in excess of $r=.772$, thus it was determined that there were no issues with multicollinearity between indicator variables. All correlations for indicators in this study can be found in table 4.4.

Additionally, it is important to make sure that predictive relationships between dependent and independent factors in the structural model do not demonstrate the effects of multicollinearity. This effect is measured by tolerance ($<.1$) and variance inflation factor (>10) metrics. For both CS1 achievement and post-CS1 interest the tolerance and variance inflation factor are within expected ranges. These statistics can be seen in table 4.5.

4.1.3.3 Normality. The assumption of normality is important when using the maximum likelihood estimation method in structural equation modeling. For this assumption to be met, univariate normality must exist for all endogenous factors, bivariate normality must exist in the joint distribution of all pairs of factors, and residuals for pairs of factors must be linear and show homoscedasticity (Kline, 2016). While MLE is robust against non-normality, it is still optimal for data to show normality to maintain confidence in outcomes of structural regression. Testing for univariate normality involves the skewness and kurtosis for all data distributions of observed variables. While there are no absolute guarantees regarding these statistics, a conservative rule suggests that skewness index values above 2 and kurtosis index values above 7 to indicate moderate non-normality. For the independent data in this study, there were no absolute skewness values greater than 1.39 and no absolute kurtosis values greater than 2.13. It is safe to assume that all univariate data distributions do not deviate from normality enough to impact future analysis. All skewness and kurtosis values can be seen in table 4.6.

To test for bivariate normality, I used Mardia's coefficient of multivariate kurtosis. The Z-score associated with this coefficient is concerning at levels above 6 (Yuan et al., 2005). For the data in this study, the value of the coefficient is 18.475 which does suggest deviation from bivariate normality. To address this, I will use a robust maximum likelihood estimator in the structural regression model analysis as is suggested for cases with extreme non-normality (Yuan et al., 2000).

For several measures in the study, the residuals showed some deviation from normality. These included the linear relationships between prior knowledge of computer programming and CS1 achievement, self-efficacy and post-CS1 interest, self-efficacy and CS1 achievement, and CS1 achievement and post-CS1 interest. Distributions of residuals in Q-Q plots can be seen in figures 4.12 - 4.20. In particular, the CS1 achievement and post-CS1 interest measures appear in several of these relationships. Due to the underlying skew of the data in these variables, the residuals show some deviation from the expected pattern. This will be noted and addressed using the robust maximum likelihood estimator as suggested above.

4.2 Model for CS1 Achievement and Post-CS1 Interest

In this study, I hypothesized that prior knowledge of computer programming impacts CS1 achievement as moderated by student gender. Additionally, I was interested in understanding whether this relationship was mediated by student self-efficacy for computer programming tasks. Other research has suggested additional factors that impact the student experience in CS1, and as part of the model that explains CS1 achievement and post-CS1 interest

I have included the sources of self-efficacy as measures to capture these effects. The model that was explored in this study with proposed relationships appears in figure 2.1. In the section that follows, I will discuss the specification of this model, the degree to which the model has been identified, the estimation of parameters using SEM software, and evaluation of the path coefficients. Respecification of the model will be discussed with regards to gender specific versions of subsequent path models.

4.2.1 Model Specification

Hypothesized relationships in this model include factor loadings between indicators and latent factors, directional relationships between latent factors, covariances between endogenous factors, and variances for all error terms associated with indicators and latent factors. In this model there are 83 free parameters (8 path coefficients, 16 factor loadings, 6 covariances between endogenous factors, 3 covariances between error terms, with the remaining terms as error variances) specified for estimation. For each latent factor, the first indicator loading will be fixed to 1 and all other loadings will be allowed to freely estimate. Covariances between error terms are generally not specified but given the possibility for error between similar indicators several covariances were allowed to freely estimate in this model.

The directional paths in this model between latent factors all derive from prior theory regarding student motivation in CS and STEM as described in the literature review. Additional paths were considered only if the model estimation and conditional independence testing suggested re-specification. As an example, modification indices indicated that the model would be improved significantly if a relationship was modeled between mastery experiences and student achievement. A relationship between these factors would make some sense, as self-efficacy and mastery experiences covary highly and self-efficacy has a direct relationship with

CS1 achievement. In the context of this study this relationship was not explored because prior theory suggested that mastery experiences impact self-efficacy directly and self-efficacy strongly relates to student achievement. The model was not adjusted to include an additional path despite the suggestion that this would have improved the fit of the model to the data.

4.2.1.1 Conditional Independence of Latent Factors. Misspecification can occur for path models when pathways that may explain variance in the model have been omitted. To address the potential for misspecification it is suggested that tests of conditional independence be performed (Kline, 2016). For this study, pairs of endogenous factors were tested that as specified in the model should not be related. The partial correlation is considered in these cases to be a correlation residual. Absolute values of the correlation residual above 0.1 can be considered a sign of misspecification. Factor pairs, conditioning sets, and partial correlations can be found in table 4.7. Based on this data, it appears that there are two places where misspecification could be occurring. Mastery experience may not be independent of achievement and vicarious experience may not be independent of interest in CS. These potential missing paths are noted here but were not included as global fit tests suggested that the model could be fit to the data adequately without their inclusion.

4.2.2 Model Identification

The model specified for this study is identified. Using the two-step modeling approach (Anderson & Gerbing, 1988), I first converted the fully latent structural regression model as a CFA measurement model. This measurement model has five latent factors, each with two or more indicators and thus is identified (Kline, 2016). For the structural model, identification follows on the basis that his is a recursive model (Kline, 2016). Given the satisfaction of these two conditions, the model is identified.

For further evidence that the model meets the requirements for identification, I provide evidence based on the number of observed variables and the number of model parameters. There are p=29 observed variables and m=60 model parameters. Calculating the degrees of freedom proceeds as follows.

$$df = p(p+1)/2 - m$$

$$df = 23(23+1)/2 - 60$$

$$df = 216$$

Since the number of degrees of freedom for this model is positive the model is over-identified.

4.2.3 Model Estimation and Fit for All Students

The model proposed in this study was first estimated and evaluated for fit for all students in the CS1 course regardless of gender. Mplus software was used to estimate the fit of the model to the data. The fit was determined by use of absolute fit metrics ($\chi2$, RMSEA, SRMR) and one comparative fit metric (CFI). For the entire student population, the model shows excellent fit: $\chi2(216) = 313.617$, $p<.001$; RMSEA=.05; SRMR=.047; CFI=.952. The Mplus code for the specification of this model for all students can be found in appendix D.

Most of the relationships in the model were found to be significant in the estimation of fit of this model to the data. For these relationships, no changes were made during further analysis. The relationship between prior knowledge of computer programming and self-efficacy was viewed considering direct and indirect effects and was explored in greater detail. All unstandardized and standardized parameter estimates for the structural regressions can be found in table 4.8.

Evidence from the estimation of the indirect and direct effects of the relationship between prior knowledge of computer programming and student achievement via self-efficacy suggest that self-efficacy did not mediate this relationship for all students. Specific indirect effects were not significant in the estimation of this model ($B=-.012$, $p=.099$). Estimates of direct and indirect effects and 95% confidence intervals can be seen in table 4.9. To test the relationship further, the direct path between prior knowledge of computer programming and student achievement was constrained to zero and the model was estimated again. The resulting fits were then compared using the chi-square difference test. The Satorra-Bentler correction was first calculated due to the use of the robust maximum likelihood estimator in both cases. This correction factor was then used in calculating the difference between the nested and original model.

$$cd = (217*1.0999-216*1.102) / (217-216) = 0.6463$$

$$TRd = (322.602*1.0999-313.617*1.102) / 0.6463 = 14.272$$

$$\chi 2(1) = 14.272, p<.001$$

Based on this outcome, I rejected the equal-fit hypothesis for the model without the direct path between prior knowledge of computer programming and achievement. The model does not fit equally well without the direct path of prior knowledge of computer programming to achievement.

The estimation of this model shows the relative impact of each of the latent factors on academic achievement in the CS1 course and post-CS1 interest. Path coefficients for mastery experiences to self-efficacy ($B=.770$, $p<.001$), vicarious experiences to self-efficacy ($B=.285$, $p=.004$), social persuasions to self-efficacy ($B=-.269$, $p=.048$), prior knowledge of computer programming to CS1 achievement ($B=.078$, $p<.001$), self-efficacy to CS1 achievement ($B=.484$, $p<.001$), self-efficacy to post-CS1 interest ($B=.355$, $p<.001$), and CS1 achievement to post-CS1 interest ($B=.194$, $p=.007$) were all significant. The model with all parameter estimates and significant path coefficients can be seen in figure 4.21.

4.3 Model Considering Gender as a Moderator

To test gender group differences, the base model described previously was run allowing for free estimation for all parameters for a group of individuals identifying as men and a group of those identifying as women. If there were no gender differences, then this model would have had similar fit to the base model; if not then it would be expected that at least one path varies significantly based on gender. Each regression path was constrained in isolation and the resulting model was then compared to the freely estimated model with the chi-square difference test using the Satorra-Bentler correction. After identifying paths that significantly differed by gender, a final model was run allowing these paths to freely estimate and corresponding fit statistics were produced. Results appear in the following sections.

4.3.1 Individual Path Differences Considering Gender

All paths in the structural regression model were allowed to freely estimate and the resulting model was run to establish a baseline for comparisons across gender. The freely estimated model has adequate fit across gender groups: $\chi2(448) = 640.218$, $p<.001$; RMSEA=.069; SRMR=.064; CFI=.907. The Satorra-Bentler correction factor for this model was 1.0092. Based on the fit of this model, it was assumed that there were gender differences for paths in the model.

Each individual path in the model was tested on its own and results compared using the chi-square difference test with the Satorra-Bentler correction. Paths in the model were labeled and can be seen in figure 4.22. An example of the process of testing a pathway has been provided below.

To test for gender differences on the "A" path between mastery experiences and self-efficacy, the model was fit again with one change. The "A" path was constrained to estimate for both genders as if there were no differences. This produced model fit statistics that could then be used to calculate the difference in chi-square values. The model with the constrained "A" path had the following fit to the data: $\chi2(449) = 640.145$, p=0; RMSEA=.069; SRMR=.064; CFI=.908. The Satorra-Bentler correction factor for this model was 1.0093. A chi-square difference test produced the following results between this model and the freely estimated model.

$$cd = (449*1.0093-448*1.0092) / (449-448) = 1.0541$$

$$TRd = (643.145*1.0093-640.218*1.0092) / 1.0541 = 2.8633$$

$$\chi2(1) = 2.8633, p=0.091$$

Based on this outcome, I failed to reject the equal-fit hypothesis for the model with and without the constrained path between mastery experiences and self-efficacy. Initial estimates suggest that these models fit equally well and there is no significant difference by gender for the estimates for this pathway; however, this path was re-evaluated in the context of iterative constraints as shown below. The ultimate decision for each path was considered using this two-step process.

This constraint evaluation was repeated for each of the nine pathways in the model. For the remaining pathways, the results of the tests appear in table 4.10. Using this data, the model was tested further by restricting multiple pathways in an iterative fashion until the level of misfit suggested that the constraints were not appropriate for the fit of the model to the data. The restricted pathways and model fit statistics for each test appear in table 4.11. Based on this process, the following paths differ by gender and should be allowed to freely estimate: mastery experiences to self-efficacy, vicarious experiences to self-efficacy, self-efficacy to post-CS1 interest, CS1 achievement to post-CS1 interest. Note that some of these paths (notably, the path between mastery experiences and self-efficacy demonstrated above) did not show significant fit differences when isolated but had a notable impact on fit when attempts were made to constrain them in the context of the model that had other constraints.

4.3.2 Model Differences by Gender

The final analysis step for testing CS1 models by gender was to consider the significance of the remaining paths. From the original set of pathways, the following were determined to differ by gender: mastery experiences to self-efficacy, vicarious experiences to self-efficacy, self-efficacy to post-CS1 interest, CS1 achievement to post-CS1 interest. The focus of this final stage was on these pathways and the potential for improving the model by trimming non-significant pathways from this group.

The overall fit for the model for women was poor: $\chi 2(216) = 349.490$, $p<.001$; RMSEA=.094; SRMR=.064; CFI=.872. It is believed that fit would improve with an increase in the size of the sample. It should be noted that there are fewer women sampled in the study (N=70) than the 83 free parameters in the model, which suggests that it was extremely underpowered. In this model there were two pathways with statistically non-significant estimates. These were the path between social persuasions and self-efficacy (B=-0.516; p=0.123) and the path between prior knowledge of computer programming and self-efficacy (B=-0.037; p=0.076). Concerns about the size of the sample should be considered with regards to the significance of individual paths as high variability impacts the likelihood that these paths will produce test statistics with significant values. Based on the sizable effect of social persuasions on self-efficacy for women, this path was not trimmed although the lack of significance was noted. The relationship between self-efficacy and prior knowledge of computer programming has been addressed previously. No other modifications to the model were suggested.

The estimation of the model for women shows the relative impact of each of the factors on academic achievement in the CS1 course and post-CS1 interest. Path coefficients for mastery experiences to self-efficacy ($B=.772$, $p<.001$), vicarious experiences to self-efficacy ($B=.455$, $p=0.048$), prior knowledge of computer programming to CS1 achievement ($B=.099$, $p=.008$), self-efficacy to CS1 achievement ($B=.488$, $p=.01$), self-efficacy to post-CS1 interest ($B=.466$, $p<.001$), and CS1 achievement to post-CS1 interest ($B=.296$, $p=.001$) were all significant. The final model with significant and non-significant path estimates can be seen in figure 4.23. All path estimates and corresponding statistics can be seen in table 4.12.

The model estimation for men had excellent fit: $\chi2(216) = 271.834$, $p=0.01$; RMSEA=.049; SRMR=.062; CFI=.945. Despite the fit, there were several pathways that did not have significant estimates. These included the paths from vicarious experience to self-efficacy ($B=0.15$; $p=0.181$), social persuasions to self-efficacy ($B=-0.151$; $p=0.211$), prior knowledge of computer programming to self-efficacy ($B=-0.019$; $p=0.295$), self-efficacy to interest ($B=0.152$; $p=0.210$), and achievement to interest ($B=0.102$; $p=0.297$). I tested the model fit when trimming the paths for the latter three from this set. While the constraining of each of these paths to zero led to failure to reject the equal-fit hypothesis, the fit of the model was not markedly better without these paths. Changes based on empirical results is not recommended due to the possibility that non-significant path estimates are due to chance when considering the random sample (Kline, 2016). Path coefficients for mastery experiences to self-efficacy ($B=.784$, $p<.001$), prior knowledge of computer programming to CS1 achievement ($B=.070$, $p=.001$), and self-efficacy to CS1 achievement ($B=.508$, $p=.005$) were all significant. The final model with significant and non-significant path estimates can be seen in figure 4.24. All path estimates and corresponding statistics can be seen in table 4.13.

CHAPTER 5: DISCUSSION

5.1 Findings

The major goal of this study was to examine how self-efficacy, the sources of self-efficacy, and prior knowledge of computer programming impacted students in first-year computer science experiences by gender. These factors were included in a model based on prior work focusing on self-efficacy stemming from Bandura's (1997) social cognitive theory. Using structural regression models, these factors were examined in relation to one another, and individual paths were tested for significance across groups. As anticipated, the effects of some of these background factors were stronger for women than they were for men. Due to the COVID-19 pandemic, some of these results will be considered in the context of a different CS1 environment than would be expected in normal circumstances. The COVID-19 period included more frequent remote learning and other pedagogical constraints.

In the following sections I review each of the research questions and discuss the corresponding major outcomes. I begin by summarizing major results related to the research question. Following this, I situate outcomes in the context of current perspectives from related research. I close the discussion of each research question by reviewing outcomes that diverged from my expectations and offer potential explanations for these outcomes. After reviewing the research questions, I then address additional findings from the data collected in the study.

5.1.1 Research Question 1A

The first research question centered on the role that prior knowledge of computer programming, the sources of self-efficacy, and self-efficacy for computing had on both CS1 achievement and post-CS1 interest for all students. Prior research has suggested that these factors play a role in either achievement or interest for students in their computer science courses and programs (Beyer 2014; Cheryan et al., 2017; Lishinski et al., 2016; Margolis & Fisher, 2003; Wiedenbeck et al., 2004; Wilcox & Lionelle, 2018). For this research question, individual factors were examined using descriptive and inferential statistics. Relationships between these factors were then tested using structural regressions in the context of structural equation modeling. As mentioned previously, this model of self-efficacy is derived from prior work focused on social cognitive theory and the role of self-efficacy for learning in STEM settings (Bandura, 1997; Lopez et al., 1997; Schunk & Pajares, 2005; Usher, 2009; Usher & Pajares, 2009).

The model of CS1 achievement and post-CS1 interest fit to the data very well for all students. The largest of the significant path coefficients was the one between mastery experiences and self-efficacy, followed by those between self-efficacy and CS1 achievement, and between self-efficacy and post-CS1 interest. Other significant paths included those between vicarious experiences and self-efficacy, between social persuasions and self-efficacy, and between prior knowledge of computer programming and CS1 achievement. The only non-significant path in the model was the one between prior knowledge of computer programming and self-efficacy (see section 5.1.2). While it is expected that several equivalent models could be produced with similar fit, the strong fit of this model gives confidence that it has captured many of the latent factors in the context of the CS1 classroom as originally specified.

5.1.1.1 Model Pathways Confirming Prior Research on CS1. Overall, analysis of the model suggested that self-efficacy was the most influential factor for positively impacting academic achievement in the CS1 course. Students reported strong self-efficacy beliefs in the context of their computer science work, and self-efficacy correlated highly with their CS1 achievement. This aligns with prior research in CS and research in other STEM subjects about the importance of self-efficacy on student outcomes (Beyer 2014; Cheryan et al., 2017; Gaspard et al., 2015; Kinnunen & Simon, 2011; Lishinski et al., 2016; Wiedenbeck et al., 2004; Williams & George-Jackson, 2014; Tellhed et al., 2017). The path analysis suggested the relative strength of this relationship, as self-efficacy had six times the effect on academic achievement as prior knowledge of computer programming did. While this outcome was expected, it is important to recognize the strength of the relationship that self-efficacy has on CS1 achievement regardless of the other factors being considered in the model. As Bandura (1997) elaborated in his social cognitive theory, self-efficacy perceptions play an important part in student choice of action and future behaviors. Students will pursue activities that they believe they will be successful in and will avoid activities that they feel will lead to failure. In addition, there is evidence that high levels of self-efficacy increase the likelihood that a student will persist on difficult tasks (Schunk & Pajares, 2005). In computer science, Kinnunen and Simon (2011) pointed specifically to the importance of avoiding repeated failures at the beginning of the learning process on programming-related tasks for student self-efficacy. Students that experienced failure were disinclined to continue with challenging programming tasks in the CS1 course. The current study did not look at the progression of self-efficacy throughout the course but given the strength of the connection between self-efficacy and achievement it confirms the importance of this factor in the CS1 course.

The impact of self-efficacy on post-CS1 interest for all students was also an important outcome from this study. In particular, the model suggested that self-efficacy had a strong positive relationship with post-CS1 interest, which was almost twice as strong as the impact of student grade in the CS1 course on post-CS1 interest. Prior research on interest in CS has found that self-efficacy has an impact on whether students express an interest in continuing in CS beyond the current course (Beyer, 2014; Cheryan et al., 2017; Sax et al., 2017). The role and relative strength of the relationship between self-efficacy and interest was not surprising as other studies have reported similar effects for both CS specifically and STEM courses more generally (Beyer, 2014; Marra et al., 2009; Sax et al., 2017; Williams & George-Jackson, 2014).

Regarding the sources of self-efficacy and their influence on CS1 achievement and post-CS1 interest via self-efficacy, there were several notable findings. First, the role of mastery experiences in predicting student self-efficacy was substantial with relation to prior knowledge of computer programming, vicarious experiences, social persuasions, and physiological states. Relatively speaking, the influence of mastery experiences on self-efficacy for all students in the CS1 course was three times as great as for both vicarious experiences and social persuasions (which had a reverse relationship and will be addressed below), and almost 38 times as strong an impact as prior knowledge of computer programming. In previous work, many have noted the power of mastery experiences to predict student self-efficacy, but the effect in this study was greater in computer science than previously reported in other areas (Britner & Pajares, 2006; Usher & Pajares, 2008). While this finding does align with prior research, there is one caution with regards to the measurement of mastery experience in this study. Items from the mastery experience scale and the self-efficacy scale were quite similar and differences between them may not have been clearly detected by study participants. The measurement of mastery experiences should focus on how an individual interprets the results of their previous performances (Bandura, 1997; Schunk & Pajares, 2005). The wording of the mastery experience items should therefore emphasize that these performances have already occurred and refer to the student's self-assessment. In this study, the wording of the items may have been overlooked by participants leading to a strong overlap in the measurement between mastery experiences and self-efficacy.

For vicarious experiences, there was a modest effect on self-efficacy for all CS1 students. By comparison with mastery experiences, this effect of peers and instructors on student self-belief was not as substantial but still played some role in how students conceptualized their ability to succeed on CS1 tasks. The measure of vicarious experiences was reduced to just two items, both of which related to the ability of the student to see themselves solving CS problems similarly to others. Items that were dropped from the model included items discussing the influence of other adults participating in CS, competition with CS1 peers, and competition with oneself. Furthermore, due to the COVID-19 pandemic, students in the CS1 courses were likely to have reduced exposure to students with perceived similarity, and consequently less opportunity to use these students as diagnostic models for their own performance. Prior work related to the influence of peers on self-efficacy support this finding and its relative effect in CS (Rosson et al., 2011; Sax et al., 2018; Ying et al., 2021), but questions remain about the strength of this result due to the unusual course conditions for students in this study and the measurement of vicarious experience utilizing only two indicators.

In fitting the model to the data, this study showed that prior knowledge of computer programming had a significant but weak positive effect on CS1 achievement. It was expected that prior knowledge of computer programming would have a positive impact on achievement as students who had previously programmed would see similar ideas throughout the CS1 course. Research has consistently pointed to prior experiences with CS as a predictor of future performance in the CS1 course (Hagan & Markham, 2000; Jones & Burnett, 2008; Wilson & Shrock, 2001). While the prior knowledge factor did impact academic achievement to some degree, it had a smaller effect than originally anticipated. This is more in line with work by Wilcox and Lionelle (2018) who saw a small positive effect on CS1 achievement based on previous academic records of student CS performance.

5.1.1.2 Novel Findings from CS1 Model. While many of the findings for all students aligned with expectations based on prior research, there were a few exceptions. These included the negative relationship between social persuasions and self-efficacy and the impact of physiological states and cost measures on model fit,. Each of these are reviewed below with respect to existing research and with potential explanations for these unexpected outcomes.

The negative relationship between social persuasions and student self-efficacy was a particularly interesting finding in the context of the CS1 model given that prior work has suggested that students who receive social support from others are more likely to see increases in their self-belief on domain-specific tasks (Lin, 2015; Sawtelle, Brewe, & Kramer, 2012; Singh et al., 2007). In this study, the data suggested that there was a moderate negative effect stemming from the positive support from teachers, family, and peers for students to pursue CS studies and careers. There are several potential explanations for this outcome. One explanation is that the relationship is a result of students feeling external pressure from family or teachers to pursue CS as an undergraduate major, which occurred for students with an undeclared major in recent work by Lehman (2019). It is possible that when students get negative feedback in the course in the form of assignment or exam grades, it runs counter to positive social support and in turn leads to a negative overall impact on self-belief. As Bong and Skaalvik (2003) noted, the view of the persuader as being a valid judge of ability is important to the strength of a social support. In the circumstance described above, the messaging does not align, and students put more precedence on the source that they perceive as having a greater amount of validity in the context of computer science. Another related explanation for the negative relationship between social persuasions and self-efficacy is that students might view the positive support from family sources with a wary eye suspecting that it is disingenuous. This approach to encouragement essentially communicates to the student that they cannot, in fact, achieve at the level they need to be successful; otherwise, the social support would not be occurring. This view aligns with prior work that has examined the interpretation of self-efficacy information, in which learners put more value in messages coming from valid sources and those that aligned with their other beliefs about their ability to succeed (Bong & Skaalvik, 2003; Schunk & Pajares, 2005; Usher, 2009). For both potential explanations, the focus was on adult persuaders, but peers also serve as potential sources of support with regards to the development of self-efficacy. As discussed previously, the COVID-19 pandemic had an impact on the frequency with which students interacted with peers. As a result, the findings from this study may differ with other studies due to the lack of peer support during these iterations of the CS1 course. Due to the uncertainty about why this outcome occurred in this study, there is a need to further explore the role of social supports in CS, particularly after students have elected to study CS at the undergraduate level. If the effects of supports are confirmed to be negative in normal classroom settings, future work could explore ways to alleviate this pressure by building self-efficacy in other ways.

Another component of the sources of self-efficacy, physiological states, did not have a significant effect on students' self-efficacy in the CS1 course. There was evidence of a strong negative relationship between physiological states and post-CS1 interest in the correlational data. Attempts to include physiological states in the model resulted in increased misfit. Additionally, the path coefficient between physiological states and self-efficacy was found to be non-significant during attempts to include it. As a result, this factor was removed and was not considered in the context of CS1 achievement and post-CS1 interest in computer science. While self-efficacy theory suggests that the impacts of physiological states on student self-efficacy are limited in comparison to the other sources of self-efficacy, it is still viewed as an important part of this relationship (Bandura, 1997; Pajares, 1997; Gaspard et al., 2015). The results of this study did not support this finding for the CS1 course. One reason for the lack of significant findings could be due to the perception of the affective response to the CS1 course. Prior research has shown that students with high levels of self-efficacy often interpret their physiological reactions to their schoolwork as a source of positive motivation for overcoming the challenge, while students with lower self-efficacy interpret the affect as a sign that they will not be successful on the task (Usher, 2009). In this study, students could be interpreting their physiological responses in ways that counter the overall effect, which led to non-significant findings. The extreme variation in student responses to physiological states items supports this hypothesis and could be a cause for the lack of statistically significant results. Further work should be done to identify reasons for inconsistent reactions to the CS environment and the role that existing self-efficacy levels plays in interpreting messaging from physiological reactions to CS1 tasks.

Another set of measures that were not included in the final model were indicators for utility value and cost. For the utility value measure, there was no significant relationship between value and CS1 achievement. There was a positive relationship between value and post-CS1 interest, but this did not differ by gender and due to the need to limit model factors due to low sample size was not pursued further. On the other hand, there were many significant negative correlations between the three measures of cost and post-CS1 interest. Prior research on cost in STEM fields by Perez et al. (2014) found that opportunity cost had a strong positive effect on intention to leave the major. The authors also found that effort cost had a moderate positive effect on intentions to leave. Multiple attempts were made to include effort cost and opportunity cost in the final CS1 model, but path coefficients were small and non-significant in every attempted arrangement. Additionally, the inclusion of these factors decreased the fit of the model to the data in ways that suggested that they should not be a part of the final analysis. During the factor analysis conducted for this study, two of the cost measures loaded with physiological states. Due to the established non-significant relationship of physiological states and self-efficacy in the model the cross-loading provided some additional evidence suggesting that the cost measures would not be significant either. Given that data about costs and physiological states were collected at the midpoint of the semester, the limited findings could be the result of having collected student perspectives on negative motivations for CS before students had experienced the full effect of the CS1 course on these factors. There were significant differences in reported cost by gender when considered outside of the CS1 model in isolation, with women reporting greater costs from participating in CS. These differences could also have played a part in limiting the overall effect of cost in the CS1 model. Further study of costs and stress related to engagement with CS is needed to clarify the role these factors may play in post-CS1 interest and career commitment.

5.1.2 Research Question 1B

For the second part of the first research question the relationship between prior knowledge of computer programming, self-efficacy, and academic achievement was explored in greater detail. The direct relationship between prior knowledge of computer programming and academic achievement was evaluated while simultaneously considering the indirect path from prior knowledge to achievement through self-efficacy. Based on the results of the structural equation modeling, it was determined that there is no indirect relationship from prior knowledge of computer programming through self-efficacy to academic achievement. Only the direct relationship existed from prior knowledge of computer programming to CS1 achievement, and its effects were relatively weak compared to other parts of the model. This does not align with prior work that has suggested a positive relationship between prior knowledge of computer programming and self-efficacy (Byrne & Lyons, 2001; Jones & Burnett, 2008; Margolis & Fisher, 2003; Wiedenbeck et al., 2004). Due to concerns about the adequacy of the CS pre-test for measuring prior knowledge of computer programming, there remain concerns about the strength of this study outcome. It should also be noted that the coefficient estimate for this path was weak and negative, even though it was not significant. The negative relationship would not align with beliefs about student self-efficacy based on work in computer science (Wiedenbeck et al., 2004).

These findings suggest that background in computer programming does not impact a student's belief that they will be successful on computer programming related tasks. One possible explanation for this could be that there is a misalignment between a student's prior programming knowledge and the types of programming tasks in the CS1 course. Roberts et al. (2012) found in a study of Australian CS students that misalignment between expected course content and prior knowledge impacted student interest in CS significantly. Incoming students found the course assessments to be too rigorous and activities did not match their expectations based on what they had been taught in pre-university computer science courses. In cases like this, students may not have developed strong self-efficacy beliefs to be successful in CS1 if their prior engagement with programming was either too facile or involved different kinds of computer science tasks. Another explanation could be a lack of validation of the prior knowledge students have accumulated before taking the CS1 course. As some students learn in informal or independent ways (e.g., online courses, textbooks, etc.), they may not have confidence that they can apply what they have learned in their pre-CS1 experiences. This aligns with beliefs about the development of self-efficacy through mastery experiences (Britner & Pajares, 2006; Usher & Pajares, 2008). This question remains open and needs further work to verify the relationships between prior knowledge of computer programming and self-efficacy.

5.1.3 Research Question 1C

The last part of the first research question was intended to identify ways in which the proposed relationships in the CS1 model varied dependent on gender identification. Prior work exploring self-efficacy for students in STEM courses has suggested differences by gender (Hutchison et al., 2006; Lopez et al., 1997; Zeldin & Pajares). To address these differences, structural regressions were conducted considering gender in a multi-group analysis. This allowed for the exploration of significant differences in path coefficients by gender, and for some conclusions to be drawn regarding the interaction of gender with the latent factors in the CS1 model. As the motivation for this study was to explore areas of inequity in the CS1 experience across gender, this section will mainly address the model for women in the CS1 course.

When fitting the model to the data by gender, the overall fit for women was worse than for the entire population. As will be discussed further in the study limitations, some of the concern over fit can be explained by the low sample size for women. Due to the COVID-19 pandemic, the number of students participating in the study was much lower than would be expected under normal conditions. Despite the sample size concerns, there were some interesting coefficient estimates in the model for women CS1 students that suggest differences in the experience of the CS1 course for women and men. I will discuss these with hopes that future data collection will allow some of these findings to be confirmed.

5.1.3.1 Model Pathways Confirming Prior Research on Women in CS1. When allowing paths to freely estimate between groups, there was one major difference in the path estimates that aligned with prior research. The path from vicarious experiences to self-efficacy had three times as strong an effect for women as for men, and model fit was improved when the path was not constrained to be equal. Women in the CS1 course expressed a greater impact of vicarious learning experiences on their self-efficacy than men did. These experiences might stem from their ability to connect with both the programming successes of their peers and the demonstration of programming problem-solving by teaching assistants and instructors. Unfortunately, the CS1 courses being studied were using a large amount of remote instruction due to the COVID-19 pandemic; due to this, it is unclear precisely why women in the study relied so heavily on vicarious experiences with regards to the development of their self-efficacy. While the vicarious experience items did not specifically mention gender, under normal circumstances I would hypothesize that gender differences in the impact of vicarious experiences on self-efficacy were due to the importance of perceived similarity with peers and instructors demonstrating their programming abilities in the CS1 course. This aligns with prior research that has suggested that similarity with peers positively impacts self-efficacy for women (Beyer, 2014; Rosson et al., 2011; Sax et al., 2018; Ying et al., 2021), and vicarious experiences play a more important role for women than men in STEM courses (Zeldin & Pajares, 2000). Prior research has also shown that women benefit from having women in instructional roles (Cheryan et al., 2011; Cohoon & Aspray, 2006; Rask & Bailey, 2002; Zeldin & Pajares, 2000). Neither course had a woman instructor, but both had several young women serving as teaching assistants who would have led lab instruction and modeled programming tasks. Differences in the model for women in CS1 from this study could be related to these instructional experiences or other attempts to connect women in the CS1 course through group work or study groups.

5.1.3.2 Novel Findings from CS1 Model for Women. Testing the CS1 model with the sample of women participants in the study presented several findings that did not align with prior research. One finding from the model that showed a large effect for women in the course in comparison to men was a strong negative relationship between social persuasions and self-efficacy. While the negative relationship between social persuasions and self-efficacy held for all students, the magnitude of the effect for women was particularly surprising. Prior research on the effects of social supports on student self-belief has suggested that there would be a moderate positive effect and the impact of this effect would be greater for women than for men (Ashcraft et al., 2012; Lin, 2015; Sax et al., 2017; Singh et al., 2007; Zeldin & Pajares, 2000). Most of the indicators for social persuasions revolve around the positive feedback students receive from others, particularly those that hold important and influential roles in the students' lives. It could be argued that the positive encouragement from these individuals manifests as pressure to be successful and remain in the field; I hypothesize that this reduced the self-efficacy of women in the course due to the conflict of this messaging with the feedback from the course itself. While most social support is well-intentioned, and the enrollment of women in CS can be linked to this support, the pressure to persist could have a negative impact on students who do not feel a strong connection to the CS environment once they have begun participating in the CS1 course. Other research has shown some evidence of this effect for undeclared students in first-year CS courses, particularly those students who had success in pre-college STEM courses (Lehman, 2019; Miliszewska et al., 2006). Lehman (2019) found that undeclared students felt great pressure to pursue a STEM major when making college decisions, and pressure to persist once they were in a STEM course of study; however, women reported less pressure to persist than men did. As mentioned in the case for all students, one important source of social support that is important are peers in the CS1 course. Due to the COVID-19 pandemic, students were not interacting with their peers as frequently, and so some of the effects from the measurement in this study are uncertain. Further work should explore how students interpret support messaging particularly in the case when they have committed to pursuing a CS major. The current study did not explore differences in interpretations of support messaging given different levels of prior knowledge, but this is also a factor that could change how students receive information from their social supports.

5.1.4 Research Question 2

The second research question was intended to explore the effects of grade feedback by gender. The processing of feedback is a key part of Bandura's (1997) social cognitive theory. In SCT, information derived from the environment, from interactions with an instructor, or from assessments are used by learners to refine their behaviors. Further, a student's perception of individual progress then sustains their self-efficacy and motivation for tasks in that domain. Prior research has established the role that feedback has on self-efficacy for women in CS1, both in terms of grades feedback (Kapoor & Gardner-McCune, 2018; Kinnunen & Simon, 2011; Lishinski et al., 2016), and feedback derived from the environment of the CS program and CS1 classroom (Beyer, 2014; Biggers et al., 2008; Frieze et al., 2012; Sax et al., 2018). It has been shown that the effects are reciprocal such that negative feedback will lead to further negative outcomes, which then in turn impact self-efficacy and create something akin to a snowball effect (Kinnunen & Simon, 2011; Lishinski et al., 2016).

Results from the path analysis showed gender differences between self-efficacy and interest, and between achievement and interest. In each case, the effects were stronger for women than for men. Specifically, the path coefficient from self-efficacy to interest was three times as strong for women as for men, and the path coefficient from achievement to interest was close to three times as strong for women as for men. These effects are interesting when considered alongside the strong effect of the path from self-efficacy to achievement (which was strong for all students). The data suggests that self-belief in the ability to succeed on computer science tasks has a tremendous impact on whether women in the CS1 course wished to continue studying computer science; however, it is not simply the self-belief that impacts interest. The relationship between CS1 achievement and post-CS1 interest also suggests that the feedback given in the course plays an important role in validating the belief of women in the course that they will continue to be successful, which in turn impacts their post-CS1 interest. Prior research has shown the importance of validation of a student's abilities with regards to their long-term interest and achievement in their studies (Carlone & Johnson, 2007; Gorson & O'Rourke, 2020; Williams & George-Jackson, 2014). It is not surprising that women would look for validation of their abilities in CS as a sign that they should continue in the field, but it is of interest that this same effect does not occur for men in CS1. For men, the relationships between achievement in CS1 and post-CS1 interest were both weaker than for women and statistically non-significant in the context of the overall model. Prior research has found that some of this effect can be explained by a self-critical disposition adopted by more women than men in the CS1 course (Gorson & O'Rourke, 2020). In Gorson and O'Rourke's study, women self-assessed more frequently and were more critical of their performance than men. Other studies have shown that women and men had different perceptions of exam scores, with men being more likely to accept a score of 70% as passing and women seeing higher scores as being unsatisfactory (Zimmerman, 2000). There is further evidence that gender differences in confidence can exist in environments where performance information is not communicated clearly (Schunk & Lilly, 1984). The students in this study were in large CS1 courses and learning in remote settings due to the COVID-19 pandemic. The frequency of feedback and clarity of what feedback means for student ability is likely to have been greatly reduced due to these factors.

While further work needs to be done to understand why there are gender differences with regards to the impact of grade feedback in CS1 on post-CS1 interest, data from this study points to a few potential areas for focus. In the CS1 model, a pre-test of prior knowledge of computer programming was used to measure student background in CS, but additional measures were collected as described in appendices E and F. As expected, there were differences in all the measures of prior experiences with CS in this study except for participation in summer camps and workshops. The pattern of these reported experiences suggests that men in the CS1 course are more likely to have had more CS exposure before electing an undergraduate CS course of study. Additionally, the measure of prior knowledge of computer programming also showed that men were entering the CS1 course with more programming knowledge. If women are currently being encouraged to try CS without as much experience or knowledge of programming, it seems reasonable to suggest that this specific group of low-experience high-support students could be looking for greater validation of their abilities before fully committing to a CS career path. If this is the case, this might explain the greater impacts of negative feedback throughout the course and an important effect from end-of-the-course feedback that could change the long-term interest of students in computer science. In studying the experiences of individual students further, we can identify areas to adapt the CS learning environment with regards to student feedback to mitigate this effect for all low-experience high-support students that may be susceptible to dramatic effects on their self-efficacy.

5.2 Implications

Based on the findings from the study, there are several implications that can be considered for research and practice. In the following sections I will make recommendations for future steps for both CS education researchers and CS educators in the context of the study results. With regards to research, I first focus on the value of the CS1 model and how it aligns within the context of other studies of self-efficacy across domains. After this I discuss the contributions to research on grades feedback in CS1, the role of social supports in the CS1 course, and discuss issues of negative messaging from the CS1 course and CS environment. For practitioners I will briefly discuss steps to address the sources of self-efficacy in the CS1 classroom and the need to re-evaluate the use of feedback.

5.2.1 Implications for Research

This study has highlighted several important factors related to student achievement in the CS1 course and interest for further study of computer science. These outcomes have reinforced previous research on the role of self-efficacy and highlighted the contributions of antecedents to self-efficacy including mastery experiences and vicarious experiences on CS1 outcomes. From these findings, there are several implications for future research related to self-efficacy and its sources for students in the CS1 course, particularly when considering gender as a moderating factor. I will review the contributions made by the study of the CS1 model, and then look at three specific areas of focus for CS education researchers in relation to the outcomes of this study.

The CS1 model proposed in this study was based on prior research focused on the role that self-efficacy and the sources of self-efficacy played for student achievement and interest

(Bandura, 1997; Schunk & Pajares, 2005; Usher & Pajares, 2006; Usher, 2009; Zeldin & Pajares, 2000). In self-efficacy models used to study mathematics, physics, engineering, and other STEM fields, it has been shown that self-efficacy was derived largely from mastery experiences, with vicarious experiences and social persuasions playing a lesser role (Britner & Pajares, 2006; Hutchison et al., 2006; Lopez et al., 1997; Sawtelle, Brewe, Goertzen, & Kramer, 2012; Usher & Pajares, 2006; Zeldin et al., 2008). In studies of gender, self-efficacy, and STEM courses research suggests that self-efficacy for men is more influenced by mastery experiences than the other sources of self-efficacy (Britner & Pajares, 2006; Usher & Pajares, 2006; Usher, 2009), and self-efficacy for women is derived more heavily from vicarious experiences and social persuasions (Lopez et al., 1997; Usher, 2009). For all students, research shows a minimal but not insignificant effect from physiological or emotional states on self-efficacy in STEM courses (Britner & Pajares, 2006; Lopez et al., 1997; Pajares, 1997; Usher & Pajares, 2006; Zeldin et al., 2008). In this study, the CS1 model showed similar results regarding the sources of self-efficacy, with mastery experiences making the most substantial contribution followed by vicarious experiences and social persuasions. The measurement of mastery experiences was potentially conflated with self-efficacy and should be considered with caution. The relationship between social persuasions and self-efficacy was moderate and negative, which did not align with previous models of self-efficacy in STEM courses and serves as a potentially new contribution to the understanding of self-efficacy in computer science. These results were limited by an insufficient sample size that was due in part to the COVID-19 pandemic; all analysis of the results of the structural regressions for the full population and the subset of women students should be evaluated with these limitations in mind. If these results hold for a larger sample of students in CS1, the model makes a significant contribution to the consideration of self-efficacy and its antecedents in the context of computer science.

One of the main outcomes of this study was the important role that grades feedback played for women in CS1 with regards to their future interest in CS courses. The relationship between these two factors was strong and positive and was many times stronger than that for men in the CS1 course, suggesting that grades served as an important source of validation for women in the course with regards to their abilities to succeed in future CS courses. Research on confidence for women in CS1 suggests that many women are susceptible to downturns in self-efficacy from negative feedback early in the course (Beyer, 2014; Frieze & Quesenberry, 2015; Lishinski et al., 2016). Several aspects of the relationship between grades feedback and interest were less clear based on the CS1 model in this study. While prior knowledge of computer programming had a limited impact on CS1 achievement, and essentially no impact on self-efficacy in the course, it is unclear the degree to which grades feedback impacts students based on their prior CS experiences. It should be noted that all results in this study will require further vetting due to sample size limitations. Future research should explore the effects of feedback on student self-belief and interest in CS at multiple time points throughout the course and test the role of prior experiences with CS on those student outcomes. Research suggests that women have less experience with CS before the CS1 course (Bernstein, 1991; Cheryan et al., 2017; Margolis & Fisher, 2003), and so gender-related effects of grade feedback may be entangled with these earlier CS experiences. Beyond disentangling the potential conflation of gender effects and prior experiences with CS, there is also a need for further study of the kinds of feedback that are used in the CS course and how this can be used as a pedagogical tool to reduce fluctuations in student self-efficacy. Efforts in this area should include ways to reduce the instances of attribution of failure on CS assessments to stable factors (e.g., ability) so that students can interpret feedback in a way that positively impacts learning in CS.

An unexpected outcome from the study was the negative relationship between social persuasions and self-efficacy for all students, with this effect being stronger for women. This finding runs contrary to prior research outside of CS, which has shown that students from marginalized groups show greater self-efficacy based on strong positive support from friends, family members, and teachers (Ashcraft et al., 2012; Lehman, 2019; Rosson et al., 2011; Sax et al., 2018). What remains unclear based on this study is how much this effect is tied to any dissonance that students might experience between the positive messages and encouragement that they receive from adult figures in their lives and the feedback from the CS environment that tells them that they are not capable of succeeding in the CS1 course. If these two sources of messages are at odds with one another, it may contribute to internal doubts for students about their abilities in CS. It may be the case that social supports matter most for encouraging students to elect a CS course of study (Beyer et al., 2003; Lin, 2016; Margolis & Fisher, 2003; Petersen et al., 2016; Wang et al., 2015; Zarrett & Malanchuk, 2005), but not for developing self-efficacy in the course itself. In this case, continued support for students provides undue pressure for persistence in CS, as the expectations of teachers and family members for student success do not align with the student's own beliefs about their abilities. Once again, this finding may not hold with an adequate sample for women but given the sample for all students suggests something worth examining more closely. Further work needs to explore the role that social supports might play throughout the student's CS experience and how they might help remove obstacles to persistence. In addition, given the power of mastery experiences on student self-efficacy, there are opportunities to explore whether increased mastery experiences in the CS1 course can help to reduce the negative effect from social persuasions.

A final area for future research stemming from this study involves the negative effects stemming from the CS environment and how these effects manifest for students' interest in CS. In this study, the data suggested strong overlap between physiological and affective states (like stress and anxiety related to the CS1 course) and measures of psychological cost (including opportunity costs and effort costs from CS participation). Due to the lack of clarity between these factors, it was difficult to include either measure in the CS1 model. Taken independently there were signs that the psychological cost measures had a strong negative association with post-CS1 interest. Additionally, while physiological states did not contribute to self-efficacy, there was evidence that they did impact CS1 achievement. Further study of negative components of the CS1 experience is needed, particularly with regards to gender, to identify how these factors impact student interest and whether high self-beliefs are enough to overcome negative affect towards computer science as was suggested by the modeling in this study.

5.2.2 Implications for Practice

For CS educators there are several ways that the outcomes of this study can be applied to improve the environment of CS1 courses and support learning for all students. These include greater focus on developing opportunities for mastery experiences and vicarious experiences and a reevaluation of the use and purpose of grades feedback. These both align with other recommendations from recent work on gender in computer science classrooms. I will review each practice below with connections to this study and to practitioner work.

In this study, mastery experiences and vicarious experiences played important roles in the development of self-efficacy for women in the CS1 course. To leverage the effect of mastery experiences for CS students, I recommend that CS educators increase the number of opportunities for CS1 students to demonstrate mastery and receive feedback from qualified sources particularly during early periods of the CS1 course when the impact on self-efficacy can be most volatile. Several researchers have suggested that successive attainments can help to bolster student self-efficacy and reduce the impacts of occasional task failures (Baker et al., 2007; Usher & Pajares, 2008; Usher, 2009). In designing assessment activities early in the course, research points to the need for authentic mastery experiences over trivial opportunities for success (Pajares, 2006). To provide an example in the context of CS1, I will turn to the task of introducing new programming constructs. In these situations, it is often the case that assessment of program construct functionality is done in the context of a problem-solving activity. The problem-solving component can challenge students that are still developing confidence with their use of the new construct. When students struggle to complete these activities successfully, it presents an opportunity for their self-efficacy for programming to decrease due to poor feedback. Instructors might instead utilize assessments that ask students to read complete code segments utilizing these constructs and report on program output or write small code segments that focus on the programming construct outside of the context of problem solving. The Use-Modify-Create paradigm described above has been recommended by CS education researchers for students with limited programming background as an introductory technique (Lee et al., 2011), and this would help scaffold learning for students in the CS1 course. For students who have not had as much prior experience with computer programming, or who have had poor quality CS education before the CS1 course, this could help to address the impacts that lack of programming knowledge has on student ability to manage the effects of feedback in the CS1 course. As suggested in recent CS education research, women in CS1 report feeling that other students outperform them and can be overwhelmed by the programming tasks they are asked to engage with early in the course (Frieze et al., 2012; Roberts et al., 2012). If students are quick to engage in self-critiques based on implicit and explicit feedback, they may be less inclined to stay in the major (Gorson & O'Rourke, 2020). By introducing more activities that allow students to demonstrate their knowledge and get clear feedback on their opportunities for success, they may be able to use these mastery experiences to buoy their overall self-efficacy for tasks in the CS1 course.

Vicarious experiences play a less impactful role for women in CS1 than mastery experiences but could be used to help enhance self-efficacy during the critical early stage of the CS1 course. Being able to engage in CS tasks with others could help to impact student comfort in the CS environment and help to establish that CS can be less isolating than it may otherwise appear to be (Ying et al., 2021). This could also help to change perceptions of CS as a masculine space with a male-dominant culture (Biggers et al., 2008). To allow students to build closer connections with peers, CS1 courses should be designed to include more activities that encourage social interaction about programming tasks. This aligns with suggested practices for encouraging vicarious learning experiences via modeling which allows for greater social comparison using the model as a diagnostic (Usher, 2009). Social interactions in CS1 could include peer instruction opportunities, paired programming, or opportunities to work in small groups during labs. Research on peer instruction has shown strong benefits for women in CS with regards to engagement and improved social context (Porter & Simon, 2013). Other peer-based social supports like paired programming have been shown to be essential in formulating greater persistence in CS courses (Porter & Simon, 2013; Rosson et al., 2011; Simon & Cutts, 2012). Utilized in tandem with stronger opportunities for mastery experiences, these approaches can help students to build self-efficacy in CS1.

An area that should be reevaluated in the design of CS1 courses is feedback, both in the context of programming tasks, and the use of feedback to communicate progress in the course. This study found evidence that final course grade had a significant impact on the interest of women students in further study of CS. Prior research has also identified the important role of feedback for developing student self-efficacy in CS, with the potential of reciprocal positive or negative effects over time (Kinnunen & Simon, 2011; Lishinski et al., 2016). With regards to programming activities, the feedback of the compiler provides students with an opportunity to evaluate the syntactical correctness of their code. Unfortunately, compiler messages require practice in interpreting before they provide valuable information to the programmer. By increasing the amount of demonstration of program development and refinement, students can start to appreciate the existing forms of feedback in their practice and use this information more effectively (Carver & Risinger, 1987; Kessler & Anderson, 1986). This can also extend to demonstrations of program testing approaches which can help students understand how to use program output to refine their programs and identify logical errors. As shown in this study, grades feedback communicates messages to a student about their ability to be successful in computer science. One recommendation for practice in CS1 courses would be to evaluate the use of grading as feedback for students. For example, instructors and course designers should focus on finding ways to emphasize growth in the course from one activity to the next and provide additional opportunities for low-stakes practice activities that have limited impact on student course grades. For students who use grade feedback as validation of their abilities, particularly in a self-critical way (Gorson & O’Rourke, 2020), this could reduce negative impacts during times of struggle in the CS1 course and reduce the likelihood that negative effects will form a reciprocal effect on student self-efficacy (Lishinski et al., 2016).

5.3 Limitations and Delimitations

The study included several elements related to the data collection and analysis that may limit the conclusions that can be drawn from the results. These include sample size issues, risk of normality violations within the data, and measurement issues surrounding the collection of mastery experience data. In addition, the design of the study involved several choices that were made to reduce the challenge of generalizing the results. Among these design choices were the removal of model factors, the reduction of pre-test questions, and the treatment of ordinal data as continuous data. I will review the study's limitations and delimitations below.

The main limitation of the study was the insufficient sample size for both the main model and the model for women in CS1. As previously highlighted when discussing the design of the study, the goal for data collection was to have a sample sufficient to fully power the structural equation model analysis. Based on Hoelter's critical N the number of CS student participants may have been sufficient; however, other approaches to determining adequate sample size suggest that the main model was underpowered. Specifically, the 182 participants did not meet either the N:q rule of 10 participants per parameter estimated (Jackson, 2003), nor the minimum 450 total participants based on the results of the Monte Carlo simulation. Running a Monte Carlo simulation based on the actual sample size suggests that the model has power of 43.5% which is quite poor. Some of the data collection was impacted by the COVID-19 pandemic, which likely reduced the number of participants in addition to changing the way some of the question items were read by participants (see below). More data was lost due to the inability to collect data during a pre-pandemic semester, which was part of the original data collection plan. Regardless of the reasons for the limited data, the findings must be viewed with some skepticism. Future research will include a larger sample to meet the requirements for adequate power.

In addition, the model using only those students identifying as women was also underpowered. In this case, the model had severe issues as the number of participants (N=70) was less than the number of freely estimated parameters (q=83). Due to the lower percentage participation of women in the CS1 course, the smaller sample was expected. Unfortunately, the data collection was impacted by the approval process for the proposed study, the COVID-19 pandemic, and the enrollment of the CS1 courses which included fewer CS majors than originally anticipated. As mentioned above, the pandemic almost certainly impacted the number of students willing to participate in the study. The estimation of the coefficients in this model has very low confidence and should be considered with extreme scrutiny. Much as in the case of the model for all CS1 participants, efforts will be made to pursue a larger sample to improve the strength of the outcomes.

Another limitation was the potential for normality violations in the data. There were issues detected in the residuals for both dependent variables: CS1 achievement and post-CS1 interest. The data in both cases showed skew and this led to issues in the tests of bivariate normality. To remedy this issue, I used the robust maximum likelihood estimation approach to the structural regressions as described by Yuan et al. (2000). This attempts to account for the violations to normality, but there are no guarantees that it completely removes the effects of the violations. Given the assumption of normality for the use of structural equation modeling, the results should be viewed with some caution.

A final limitation of note was related to the measurement of mastery experiences. The wording of the items for mastery experiences was drawn from a pre-existing instrument, but upon reflection it was not clear that the measures were distinctly different from the ones used to capture self-efficacy. The intention of the mastery experiences items was to ask students about their past accomplishments while the self-efficacy items were designed to capture belief in future performance. I did not draw any special attention to wording in the online survey that participants received that would suggest a difference in tense between the mastery experience items and the self-efficacy items. As the confirmatory factor analysis revealed a high degree of covariance between these factors, it suggests that students may not have differentiated between them. While it was expected that there would be a strong relationship between the two factors, it is unclear based on this data collection whether the strength of this relationship was due to the true relationship or the unclear wording. All results based on the relationship between mastery experiences and self-efficacy should be viewed with some caution.

In addition to the noted limitations, I made several design decisions to delimit the study and narrow the overall scope of the research based on the impact of the inclusion of these elements on the study results. The first decision was to reduce the overall model complexity by removing the task value, cost, and physiological states factors from the model. My initial instinct was to include some of these additional elements to account for motivational factors that both added to and detracted from post-CS1 interest. Based on preliminary analysis of the data which showed either limited effects or confounding results, I then made the decision to omit these factors. The inclusion of additional measures would have decreased the power of an already underpowered model, and some exploratory analysis showed that these factors did not contribute significantly to the model based on anticipated relationships. Further, due to the cross-loading of physiological states and the cost measures it would have been difficult to separate the effects of these factors from one another, complicating their use. While future studies of experiences for women in CS should include these components, it was my opinion that more work was needed to disentangle their effects before being considered for use in this study.

Another design decision was the reduction of the pre-test from the original 26 items down to 10 items. The two reasons for doing so were the potential for test or survey fatigue from study participants, and the reduction of redundant items that were imbalanced in relation to the topics of the course. The original pre-test was lengthy and was to be given over a one-hour period. The reduction in items allowed for the suggested time to be reduced to 30 minutes. In some cases, reducing test items could reduce the internal reliability of the instrument; however, the data collected suggests that this was not necessarily the case for the CS1 students in the sample. The pre-test as it was originally designed also asked multiple questions covering some concepts (e.g., functions and parameter passing), and very few questions covering others (e.g., conditional execution). I elected to remove certain items to create better balance for the reduced set of questions on the pre-test as it was used in this study. The instrument with reduced questions had good reliability but had added complexity in the form of a test-specific programming language which may have impacted measurement of the student prior knowledge of computer programming.

One last decision that was made involved the treatment of the data in the study. Most of the measures were on an ordinal scale using Likert-style items. For the data analysis it was decided that these items should be treated as continuous measures, which is supported by the statistical literature on SEM and measurement (DiStefano, 2002; Kline, 2016; Rhemtulla et al., 2012). Specifically looking at the dependent variables, the choice to use ordinal data would have involved underlying logistic regressions in the structural equation modeling which would have complicated the analysis given the small sample size (Kline, 2016). Many of the available software tools for SEM analysis are unable to produce the output needed to support logistic regressions without dramatic increases to the number of data points used. To improve the clarity of the results and reduce some of the concerns about sample size related to the estimation of path coefficients for non-linear regressions, the dependent variables were treated as if they were continuous measures.

5.4 Conclusions

Understanding the experiences of students in the CS1 course can help to identify ways to increase the participation and retention of women in computer science career pathways. With increased enrollments in university CS programs, and greater numbers of women entering the major, there is the potential to change aspects of the CS culture and environment that negatively impact women and other marginalized students. To make these changes it is important to understand specific ways in which gender impacts the experiences students have in the CS1 course so that motivation for further study remains high and students can engage in their studies to the fullest potential. Research from the last 20 years has focused on identifying some of the factors that impact men and women differently. This study aimed to more closely study these factors in the context of existing learning theory to better understand how they impact the student experience in the CS1 course. Taken from Bandura's (1997) social cognitive theory, the role of self-efficacy and its antecedents provided a strong basis for evaluating the student experience in CS1.

This work contributes to the understanding of experiences in first-year computer science by addressing the role of self-efficacy, the sources of self-efficacy, and prior knowledge of computer programming on CS1 achievement and post-CS1 interest while considering gender as a moderating factor. While initial assumptions regarding the strength of the effect that prior knowledge of computer programming may have on self-efficacy beliefs and overall course achievement were proved to be overstated, the unique experiences of women in the CS1 course were highlighted. Among these are the impact of grades feedback on student interest, the greater role of vicarious experiences and social persuasions on self-efficacy for computer programming for women, and the challenge that social supports may present once students have elected a CS major on their persistence in the face of challenging course material. Despite some noted limitations from the challenge of collecting an adequate sample, and the limiting role of the COVID-19 pandemic, this study has provided a few pathways for further study of the CS1 environment and the experience of women in the course. These include explorations of varying types of feedback on student self-belief throughout the course, the role of social supports beyond the commitment of students to pursuing a CS major, and the psychological costs associated with CS that impact student interest. Overall, the study provides insights into experiences in the CS1 course, areas where the CS1 course may impede student success and interest, and differences in these effects by gender.

APPENDICES

APPENDIX A: TABLES

Table 3.1

Demographic and school information for sample.

Measure	Gender identification				
	Man	Woman	Non-binary	Did not respond	Total
Total	108	70	3	1	182
School					
University A	70	43	3	1	117
University B	38	27	0	0	65
Major					
Computer Science	90	50	2	1	143
Data Science	11	18	1	0	30
CS minor or undeclared	7	2	0	0	9
Race / Ethnicity					
Asian	51	38	0	1	90
Black or African American	0	1	0	0	1
Hispanic or Latinx	9	4	0	0	13
White	35	23	2	1	60
Two or more races	6	4	0	0	10
Did not report	7	0	1	0	9

Table 3.2

Dimensional reduction for sources of self-efficacy, self-efficacy, and cost measures.

Indicator	Component					
	1	2	3	4	5	6
Mastery1	.594					
Mastery2	.604					
Mastery3					-.462	
Mastery4	.667					
Mastery5	.870					
Mastery6	.802					
Vicarious2						.917
Vicarious4						.645
Social1				.689		
Social2				.837		
Social3				.744		
Social4				.822		
Social5				.754		
Social6				.797		
Phys1		.642				
Phys2		.752				
Phys3		.555				
Phys4			.510			
Phys5			.698			
Phys6			.584			
SE1	.741					
SE2	.770					
SE3	.704					
SE4	.877					
SE5	.543					
SE6	.691					

Table 3.2 (cont'd)

Indicator	Component					
	1	2	3	4	5	6
Cost1		.923				
Cost2		.858				
Cost3		.647				
Cost4			.881			
Cost5			.699			
Cost6			.812			
Cost7					.873	
Cost8					.830	
Cost9					.823	

Note. This information was extracted using principal component analysis using promax rotation with Kaiser normalization. The rotation converged in 7 iterations.

Table 3.3

Grade point scale and conversion used for grades data.

Instructor assigned letter grade	GPA
A+, A	4.0
A-	3.7
B+	3.3
B	3.0
B-	2.7
C+	2.3
C	2.0
C-	1.7
D+	1.3
D	1.0
D-	0.7
E	0.0

Table 4.1

Means, standard deviations, and independent samples t statistics for study measures.

Measure	Total		Man		Woman		
	M	SD	M	SD	M	SD	t(176)
PKCP	4.429	2.776	4.713	2.819	3.914	2.685	1.88
ME	3.612	.820	3.796	.688	3.305	.919	4.07**
VE	3.665	.836	3.759	.769	3.536	.910	1.76
SP	3.116	.839	3.228	.802	2.943	.875	2.24*
PS	2.414	.980	2.298	.909	2.633	1.057	-2.26**
SE	4.031	.770	4.188	.667	3.793	.850	3.56**

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy.

*p < .05. **p < .01.

Table 4.2

Correlations between factors in the CS1 model for all students in the sample.

Variable	1	2	3	4	5	6	7	8
1. PKCP	-							
2. ME	.40**	-						
3. VE	.18*	.43**	-					
4. SP	.29**	.59**	.42**	-				
5. PS	-.43**	-.67**	-.39**	-.35**	-			
6. SE	.30**	.81**	.52**	.49**	-.68**	-		
7. Achieve	.33**	.49**	.15	.25**	-.42**	.37**	-	
8. Interest	.26**	.48**	.37**	.31**	-.47**	.47**	.47**	-

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy; Achieve = CS1 achievement; Interest = post-CS1 interest.

*p < .05. **p < .01.

Table 4.3

Correlations with dependent factors in the CS1 model and independent samples Fisher r-to-Z transformations by gender.

Variable	CS1 achievement			Post-CS1 interest		
	Man	Woman	Z	Man	Woman	Z
PKCP	.301	.348	-.34	.187	.308	-.83
ME	.449	.503	-.45	.237	.615	-3.04**
VE	.228	.035	1.26	.236	.445	-1.52
SP	.195	.280	-.58	.134	.426	-2.05*
PS	-.378	-.437	.45	-.342	-.568	1.84
SE	.316	.401	-.62	.191	.646	-3.68**
Achieve	-	-	-	.236	.646	-3.38**

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy; Achieve = CS1 achievement.

*p < .05. **p < .01.

Table 4.4

Correlations for indicators of latent factors in the CS1 model.

Indicator	ME 1	ME 2	ME 3	ME 4	ME 5	ME 6	VE 2	VE 4	SP 1	SP 2	SP 3	SP 4	SP 5	SP 6	SE 1	SE 2	SE 3	SE 4	SE 5	SE 6
ME1	-																			
ME2	.60	-																		
ME3	.58	.42	-																	
ME4	.51	.42	.43	-																
ME5	.51	.50	.48	.51	-															
ME6	.52	.54	.44	.51	.71	-														
VE2	.26	.30	.18	.23	.22	.27	-													
VE4	.32	.27	.30	.41	.32	.34	.51	-												
SP1	.34	.31	.20	.23	.19	.35	.13	.17	-											
SP2	.43	.45	.33	.33	.32	.42	.32	.26	.51	-										
SP3	.46	.41	.30	.41	.38	.41	.35	.33	.37	.67	-									
SP4	.47	.41	.29	.39	.36	.46	.30	.31	.41	.71	.69	-								
SP5	.39	.37	.30	.29	.28	.39	.45	.25	.48	.61	.49	.63	-							
SP6	.40	.42	.25	.39	.30	.45	.30	.31	.50	.58	.54	.58	.68	-						
SE1	.56	.52	.54	.55	.55	.56	.32	.48	.24	.36	.37	.40	.31	.32	-					
SE2	.47	.46	.46	.53	.47	.53	.29	.34	.23	.32	.34	.39	.30	.31	.73	-				
SE3	.50	.49	.45	.52	.51	.47	.34	.46	.20	.35	.39	.35	.21	.24	.64	.62	-			
SE4	.56	.51	.59	.50	.65	.60	.35	.44	.29	.34	.38	.35	.30	.35	.66	.63	.71	-		
SE5	.51	.37	.51	.51	.55	.53	.37	.44	.23	.37	.42	.41	.34	.37	.60	.56	.60	.63	-	
SE6	.66	.51	.63	.51	.56	.48	.35	.43	.24	.38	.39	.33	.35	.35	.69	.57	.62	.73	.67	-

Note. ME = mastery experience; VE = vicarious experience; SP = social persuasion; SE = self-efficacy.

Table 4.5

Tolerance and variance inflation factor statistics for CS1 achievement, post-CS1 interest, and self-efficacy.

Model	Tolerance	Variance inflation factor
CS1 achievement		
Self-efficacy	.911	1.097
Prior knowledge of computer programming	.911	1.097
Post-CS1 interest		
CS1 achievement	.864	1.157
Self-efficacy	.864	1.157
Self-efficacy		
Mastery experiences	.402	2.485
Vicarious experiences	.746	1.341
Social persuasions	.604	1.657
Prior knowledge of computer programming	.781	1.280

Note. Three models were tested (CS1 achievement, Post-CS1 interest, Self-efficacy) with factors that were represented in the CS1 model.

Table 4.6

Skewness and kurtosis values for all indicator items of latent factors in CS1 model.

Indicator	Skewness	Kurtosis
PKCP	.197	-.930
ME1	-.642	-.468
ME2	-.180	-.976
ME3	-.651	-.195
ME4	-.913	.861
ME5	-1.336	1.833
ME6	-.484	-.449
VE2	-.443	-.580
VE4	-.591	.077
SP1	-.159	-.072
SP2	-.033	-.465
SP3	-.354	-.573
SP4	-.077	-.556
SP5	-.022	-.487
SP6	-.021	-.307
SE1	-1.015	.766
SE2	-.513	-.316
SE3	-1.390	1.916
SE4	-.943	.502
SE5	-1.102	1.363
SE6	-.883	.278

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy.

Table 4.7

Partial correlations for latent factors in CS1 model with given conditioning sets.

Independence	Conditioning set	Partial Correlation
ME ⫫ Achieve	SE, VE, SP, PKCP	.292*
VE ⫫ Achieve	SE, ME, SP, PKCP	-.059
SP ⫫ Achieve	SE, ME, VE, PKCP	-.052
PKCP ⫫ Interest	SE, ME, VE, SP, Achieve	.039
ME ⫫ Interest	SE, VE, SP, PKCP, Achieve	.048
VE ⫫ Interest	SE, ME, SP, PKCP, Achieve	.187*
SP ⫫ Interest	SE, ME, VE, PKCP, Achieve	.007

* absolute partial correlation residual greater than 0.1

Table 4.8

Standardized and unstandardized estimates for CS1 model for all students.

Parameter	B	SE	β	Sig. level
Self-efficacy				
ME	.770**	.095	.960**	0
VE	.285**	.099	.262**	.004
SP	-.269*	.136	-.199*	.048
PKCP	-.025	.015	-.098	.081
Achievement				
SE	.484**	.122	.348**	0
PKCP	.078**	.019	.216**	0
Interest				
SE	.355**	.087	.396**	0
Achieve	.194**	.072	.301**	.007

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy.

*$p < .05$. **$p < .01$.

Table 4.9

Direct and indirect effect estimates and 95% confidence intervals for prior knowledge of computer programming on CS1 achievement with self-efficacy as a mediator.

Path	Label	Estimate	95% CI
Indirect effect	(a*b)	-.012	(-.027, 0)
Prior knowledge -> Self-efficacy	(a)	-.025	(-.054, -.001)
Self-efficacy -> Achievement	(b)	.484	(.245, .723)
Prior knowledge -> Achievement	(c’)	.078	(.041, .109)

Table 4.10

Changes in chi-square difference test for model given the constraint of paths to be equal by gender.

Constrained path	$\chi 2$ difference	Sig. Level
A	2.8663	.091
B	3.9432	.047
C	1.2927	.256*
D	.3881	.533*
E	.4247	.515*
F	0	1*
G	3.0404	.081
H	2.6302	.105*

*did not reject the equal-fit hypothesis between constrained and unconstrained models.

Table 4.11

Model fit statistics for CS1 model with increasingly restrictive paths for estimation by gender.

Constrained paths	χ2(df)	CFI	RMSEA	SRMR
Original	640.218 (448)	.907	.069	.064
F, D	640.374 (450)	.908	.069	.064
F, D, E	641.146 (451)	.908	.069	.064
F, D, E, C	642.071 (452)	.908	.069	.064
F, D, E, C, H	645.887 (453)	.907	.069	.068

Table 4.12

Standardized and unstandardized estimates for CS1 model for all women.

Parameter	B	SE	β	Sig. level
Self-efficacy				
ME	.772**	.159	.961**	0
VE	.455*	.231	.381*	.048
SP	-.516	.335	-.315	.123
PKCP	-.037	.021	-.120	.076
Achievement				
SE	.488*	.188	.354*	.010
PKCP	.099**	.037	.235**	.008
Interest				
SE	.466**	.133	.495**	0
Achieve	.296**	.086	.434**	.001

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy.

*p < .05. **p < .01.

Table 4.13

Standardized and unstandardized estimates for CS1 model for all men.

Parameter	B	SE	β	Sig. level
Self-efficacy				
ME	.784**	.171	.953**	0
VE	.150	.112	.173	.181
SP	-.151	.121	-.147	.211
PKCP	-.019	.019	-.098	.295
Achievement				
SE	.508**	.183	.313**	.005
PKCP	.070**	.021	.218**	.001
Interest				
SE	.152	.121	.162	.210
Achieve	.102	.098	.176	.297

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy.

Table F.1

Means, standard deviations, and independent samples t statistics for unused task and cost measures.

Measure	Total		Man		Woman		
	M	SD	M	SD	M	SD	t(176)
UV	4.174	.517	4.169	.532	4.183	.496	-.18
EFF	3.024	1.057	2.911	1.016	3.276	1.060	-2.31*
EMO	2.147	.915	2.096	.859	2.233	1.013	-.97
OPP	2.967	1.135	2.833	1.060	3.210	1.176	-2.22*

Note. UV = utility value; EFF = effort cost; EMO = emotional cost; OPP = opportunity cost.

*p < .05. **p < .01.

Table F.2

Correlations between factors in the CS1 model, utility value, and cost measures for all students in the sample.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. PKCP	-											
2. ME	.40**	-										
3. VE	.18*	.43**	-									
4. SP	.29**	.59**	.42**	-								
5. PS	-.43**	-.67**	-.39**	-.35**	-							
6. SE	.30**	.81**	.52**	.49**	-.68**	-						
7. UV	.14	.19*	.33**	.27**	-.08	.31**	-					
8. EFF	-.34**	-.53**	-.23**	-.26**	.77**	-.46**	.03	-				
9. EMO	-.37**	-.57**	-.40**	-.36**	.76**	-.66*	-.20**	.59**	-			
10. OPP	-.28**	-.61**	-.30**	-.37**	.66**	-.50**	-.01	.61**	.51**	-		
11. Achieve	.33**	.49**	.15	.25**	-.42**	.37**	.13	-.34**	-.34**	-.37**	-	
12. Interest	.26**	.48**	.37**	.31**	-.47**	.47**	.21**	-.32**	-.54**	-.30**	.47**	-

Note. PKCP = prior knowledge of computer programming; ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy; UV = utility value; EFF = effort cost; EMO = emotional cost; OPP = opportunity cost; Achieve = CS1 achievement; Interest = post-CS1 interest.

*p < .05. **p < .01.

Table F.3

Correlations between dependent factors and those measured but not included in the CS1 model with independent samples Fisher r-to-Z transformations by gender.

Variable	CS1 achievement			Post-CS1 interest		
	Man	Woman	Z	Man	Woman	Z
UV	.088	.173	-.55	.177	.272	-.64
EFF	-.343	-.313	-.22	-.245	-.368	-.87
EMO	-.262	-.408	1.05	-.440	-.640	1.83
OPP	-.420	-.307	-.83	-.202	-.335	-.92

Note. UV = utility value; EFF = effort cost; EMO = emotional cost; OPP = opportunity cost.

*p < .05. **p < .01.

Table F.4

Means, standard deviations, and independent samples t statistics for prior experience factors.

Measure	Total		Man		Woman		
	M	SD	M	SD	M	SD	t(176)
AP CS A	.94	1.28	1.06	1.33	.76	1.19	1.58
AP CS P	.46	.97	.46	.98	.43	.93	.23
Non-AP CS	1.45	1.35	1.50	1.34	1.33	1.36	.83
Prog. Lang.	1.78	1.06	1.83	1.03	1.69	1.10	.91
SW/HW	1.14	1.11	1.28	1.10	.91	1.07	2.17*
Club/Group	.83	1.10	.85	1.09	.83	1.13	.14
Online CS	.79	1.05	.85	1.08	.63	.97	1.41
Workshop/Summer	.62	.98	.58	.93	.63	1.04	-.30

Note. AP CS A = took the AP CS A course; AP CS P = took the AP CS Principles course; Non-AP CS = took a non-AP CS course; Prog. Lang. = learned a programming language; SW/HW = completed a software or hardware project; Club/Group = participated in a computing themed club or group; Online CS = took an online CS course or MOOC; Workshop/Summer = participated in a computer science workshop or summer camp.

*p < .05. **p < .01.

Table F.5

Means, standard deviations, and one-way analyses of variance in AP CS A and programming language experiences on CS1 achievement by gender.

Measure	None		Minimal		Moderate		Very High		
	M	SD	M	SD	M	SD	M	SD	
Man									F(3,104)
AP CS A	3.452	.998	4.000	.000	3.264	1.133	3.904	.906	2.467
Prog. Lang.	3.450	.707	3.605	.996	3.470	1.130	3.697	.600	.470
Woman									F(3,66)
AP CS A	3.144	1.276	2.750	1.768	3.633	.498	3.818	.405	1.514
Prog. Lang.	2.731	1.352	2.900	1.543	3.482	.951	3.706	.772	2.894*

Note. AP CS A = took the AP CS A course; Prog. Lang. = learned a programming language.

*$p < .05$.

Table F.6

Correlations for measures of prior knowledge of computer programming and prior experience with computer science.

Prior knowledge measure	Motivational and outcome measures						
	ME	VE	SP	PS	SE	Achieve	Interest
Pre-test	.402**	.176*	.294**	-.429**	.298**	.330**	.261**
AP CS A	.268**	.100	.266**	-.160*	.138	.193**	.144
AP CS P	.035	.030	.040	-.012	.011	.020	.027
Non-AP CS	.192**	.140	.315**	-.124	.154*	.117	.104
Prog. Lang.	.286**	.122	.392**	-.267**	.278**	.192**	.198**
SW/HW	.233**	.095	.339**	-.200**	.161*	.196**	.138
Club/Group	.142	.193**	.221**	-.042	.116	.123	.088
Online CS	.142	.059	.206**	-.143	.144	.113	.055
Workshop/Summer	.023	.029	.083	-.064	.058	.055	-.040

Note. ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; PS = physiological states; SE = self-efficacy; Achieve = CS1 achievement; Interest = post-CS1 interest in CS. AP CS A = took the AP CS A course; AP CS P = took the AP CS Principles course; Non-AP CS = took a non-AP CS course; Prog. Lang. = learned a programming language; SW/HW = completed a software or hardware project; Club/Group = participated in a computing themed club or group; Online CS = took an online CS course or MOOC; Workshop/Summer = participated in a computer science workshop or summer camp.

*$p < .05$. **$p < .01$.

APPENDIX B: FIGURES

Figure 2.1

Model of CS1 achievement and post-CS1 interest.

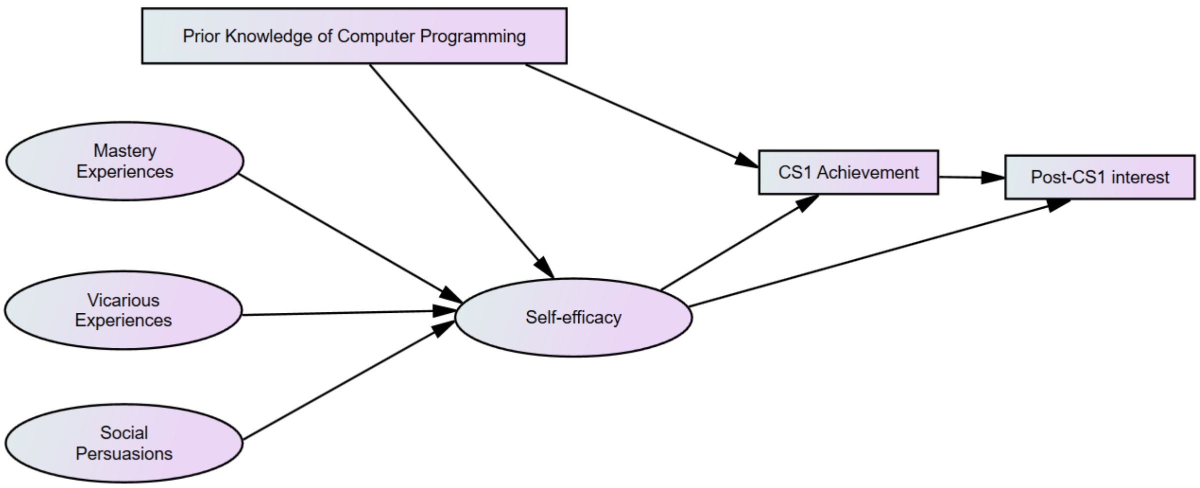
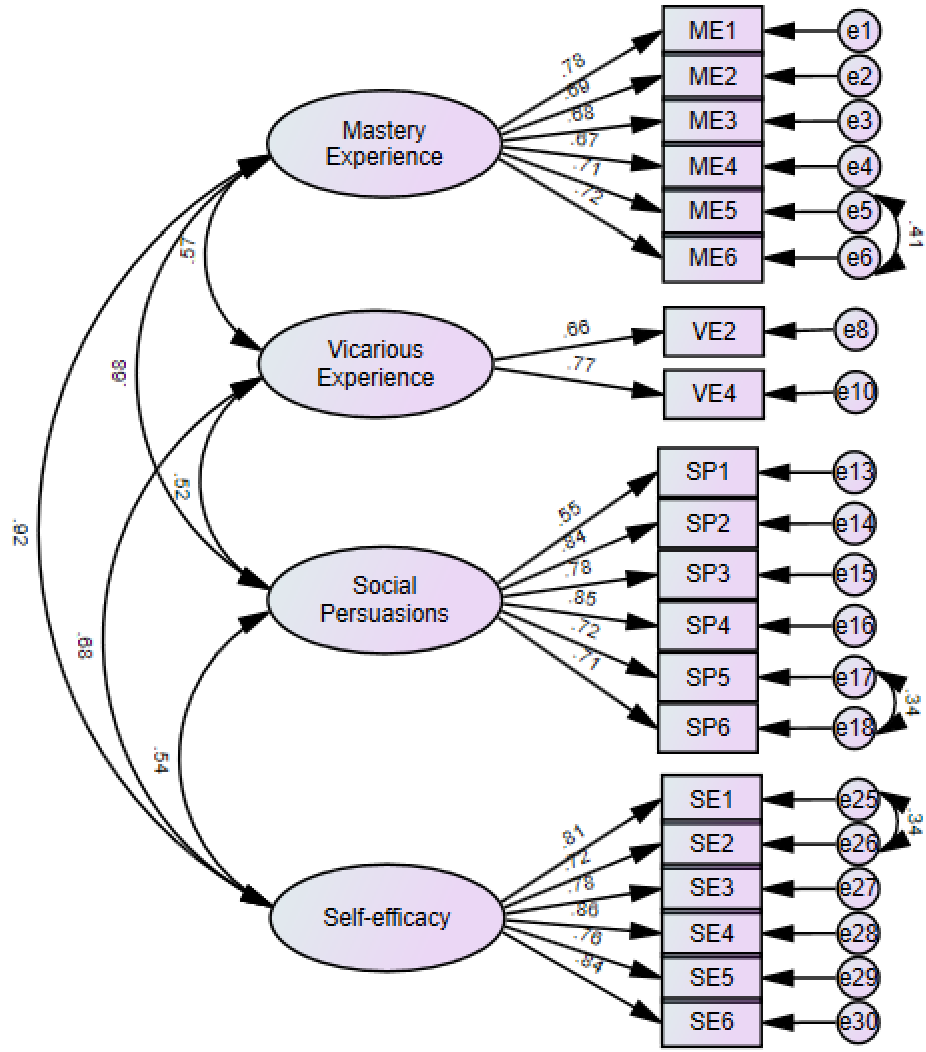


Figure 3.1

Confirmatory factor analysis for motivational factors in CS1 model.



Note. ME = mastery experiences; VE = vicarious experiences; SP = social persuasions; SE = self-efficacy; e = error residual.

Figure 3.2

Data collection timeline for survey instruments and pre-test.

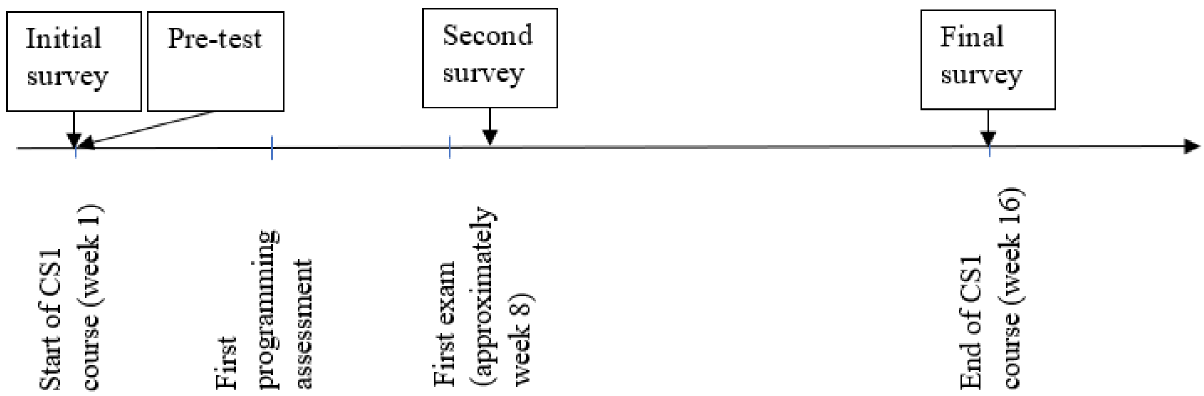
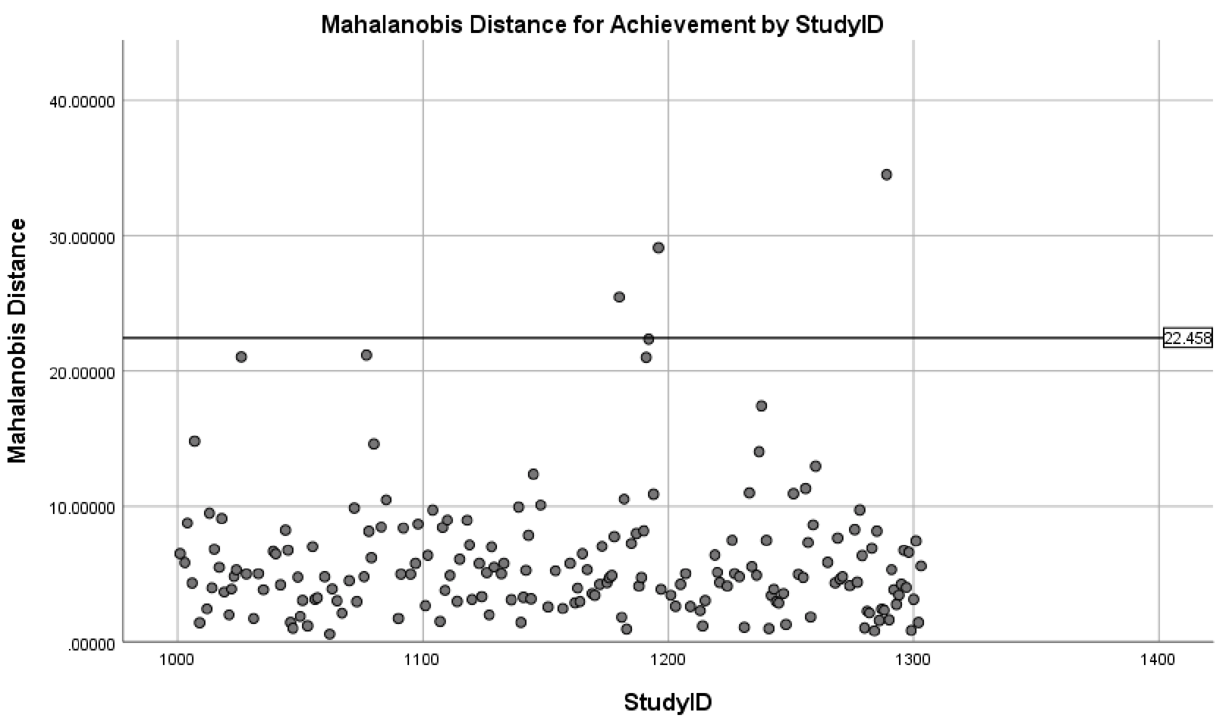


Figure 4.1

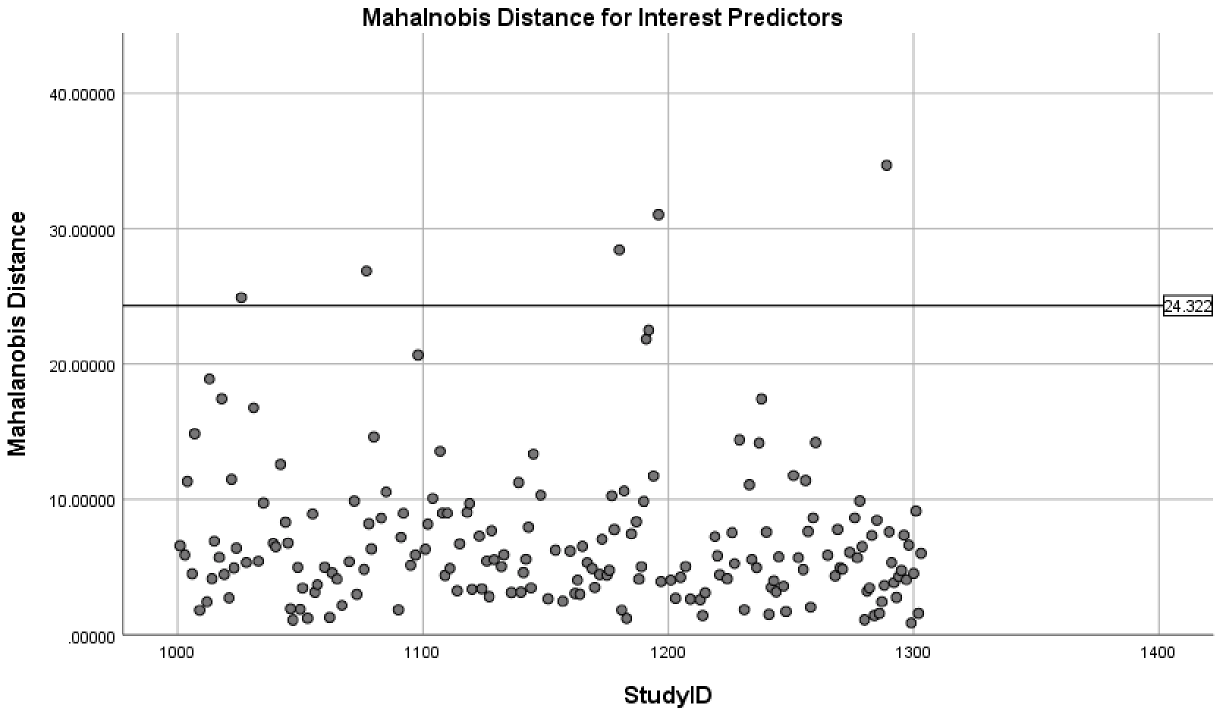
Mahalanobis distances of achievement predictors for all data in sample of CS1 students.



Note. Distances greater than 22.458 are influential points given a chi-square distribution with 6 degrees of freedom at the .1%-level.

Figure 4.2

Mahalanobis distances of interest predictors for all data in sample of CS1 students.



Note. Distances greater than 24.322 are influential points given a chi-square distribution with 7 degrees of freedom at the .1%-level.

Figure 4.3

Residual plot of prior knowledge of computer programming on CS1 achievement.

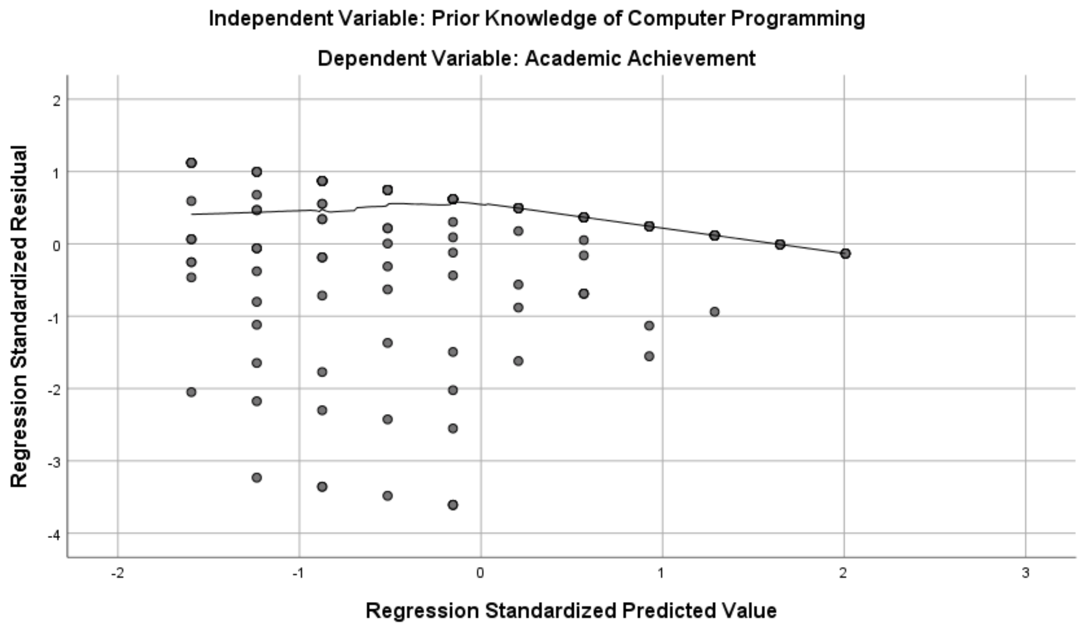


Figure 4.4

Residual plot of self-efficacy on CS1 achievement.

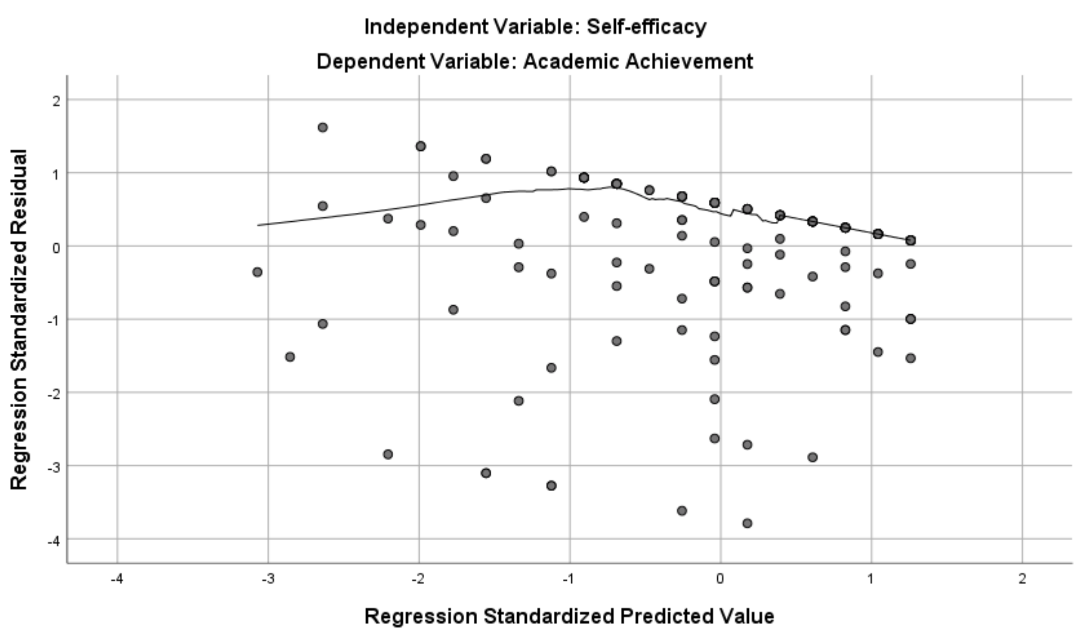


Figure 4.5

Residual plot of CS1 achievement on post-CS1 interest.

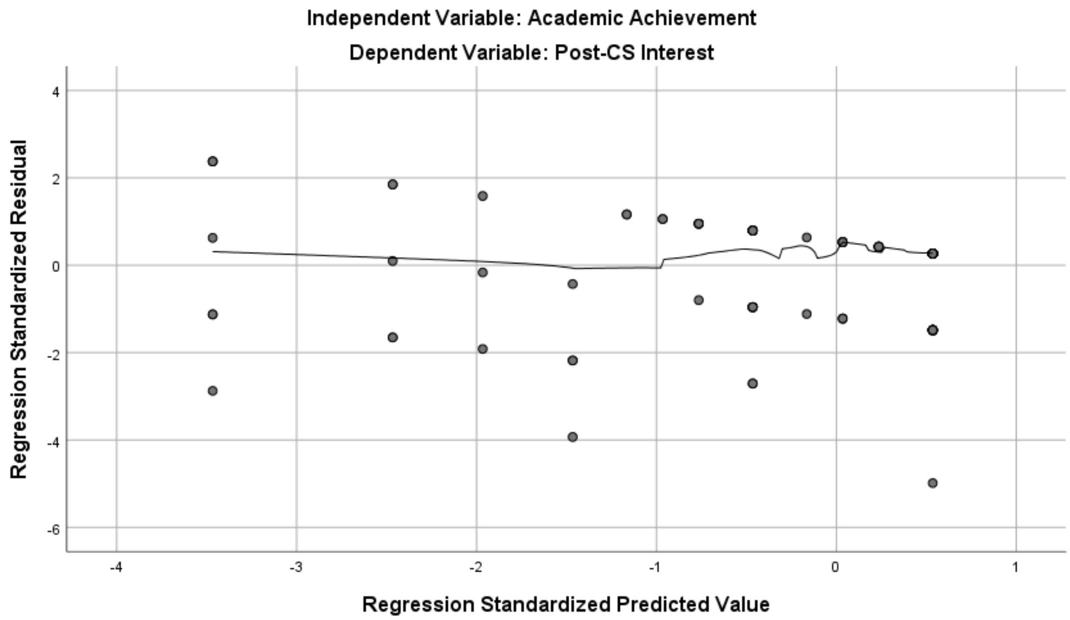


Figure 4.6

Residual plot of self-efficacy on post-CS1 interest.

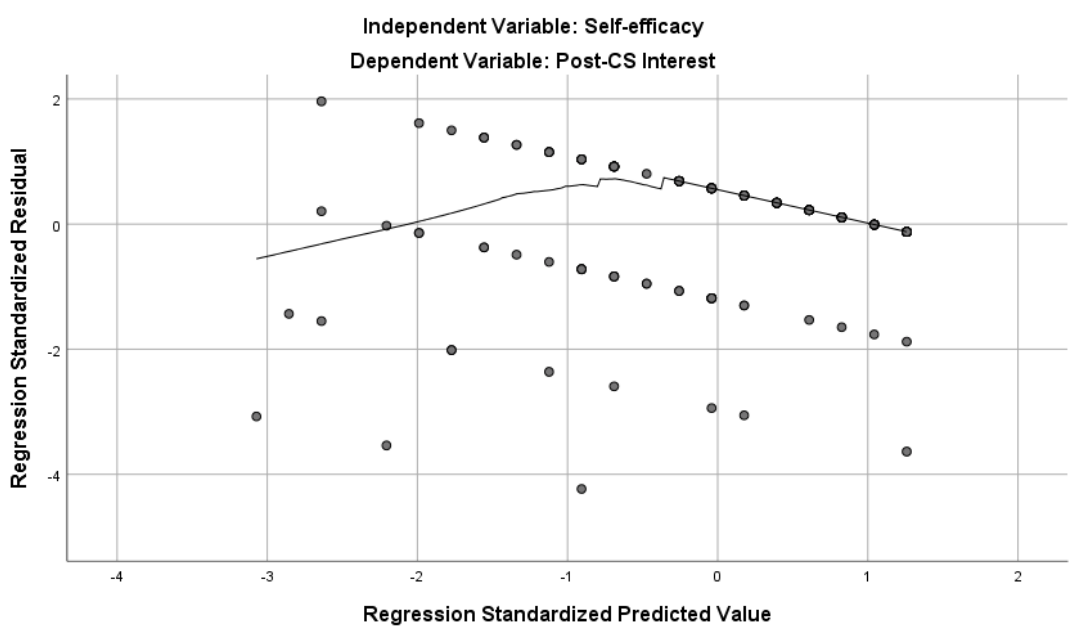


Figure 4.7

Residual plot of prior knowledge of computer programming on self-efficacy.

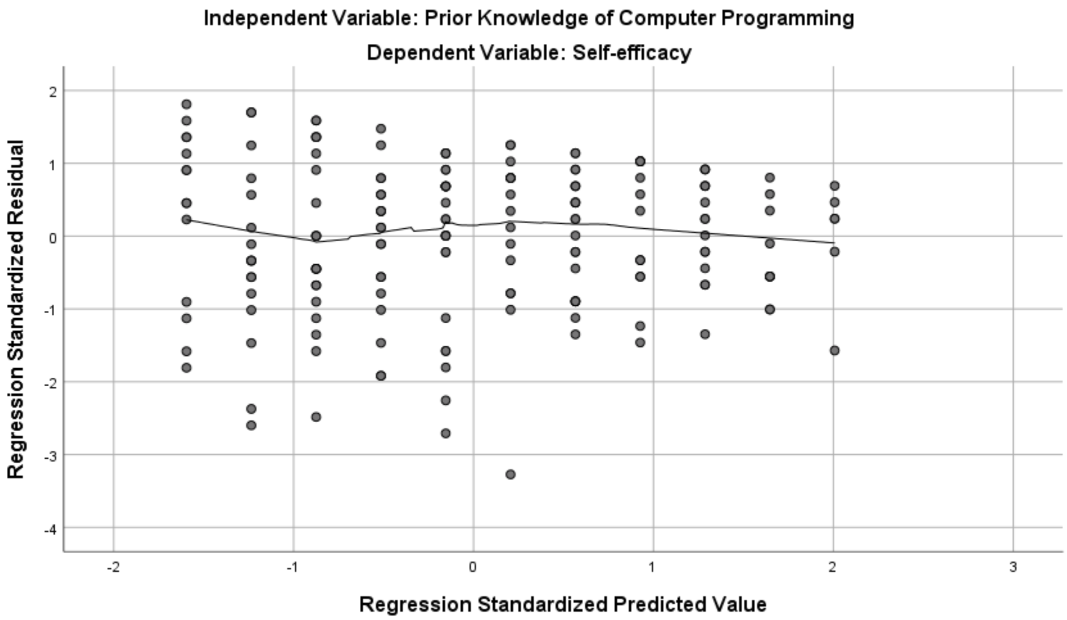


Figure 4.8

Residual plot of mastery experience on self-efficacy.

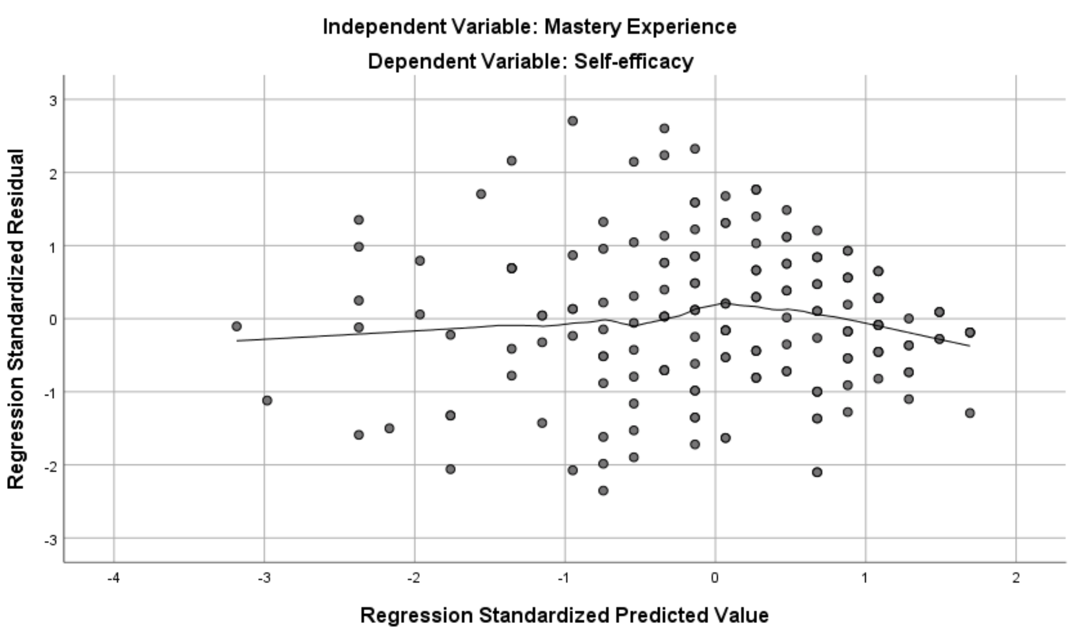


Figure 4.9

Residual plot of vicarious experience on self-efficacy.

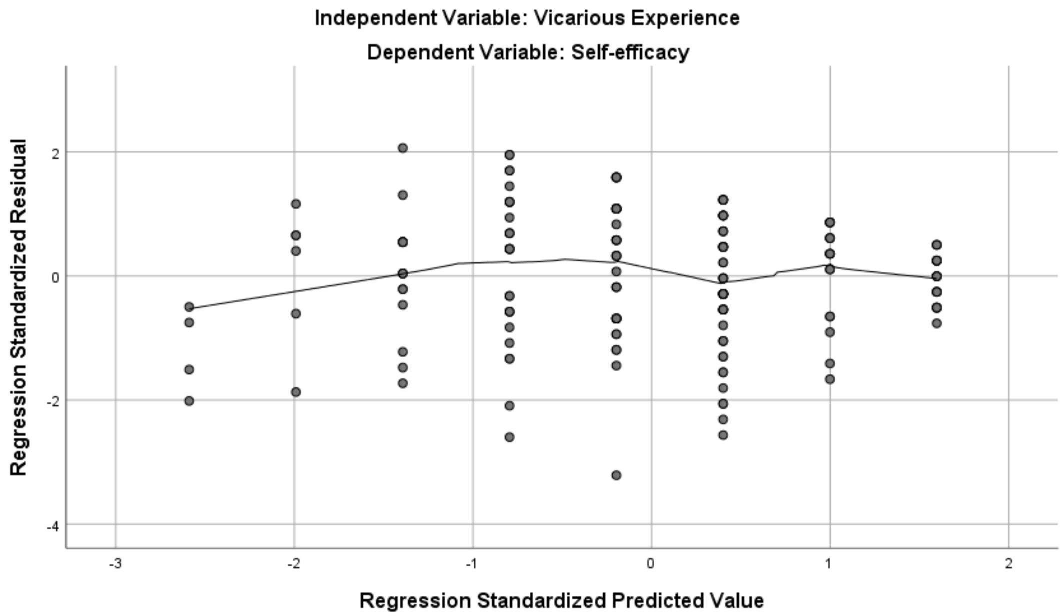


Figure 4.10

Residual plot of social persuasions on self-efficacy.

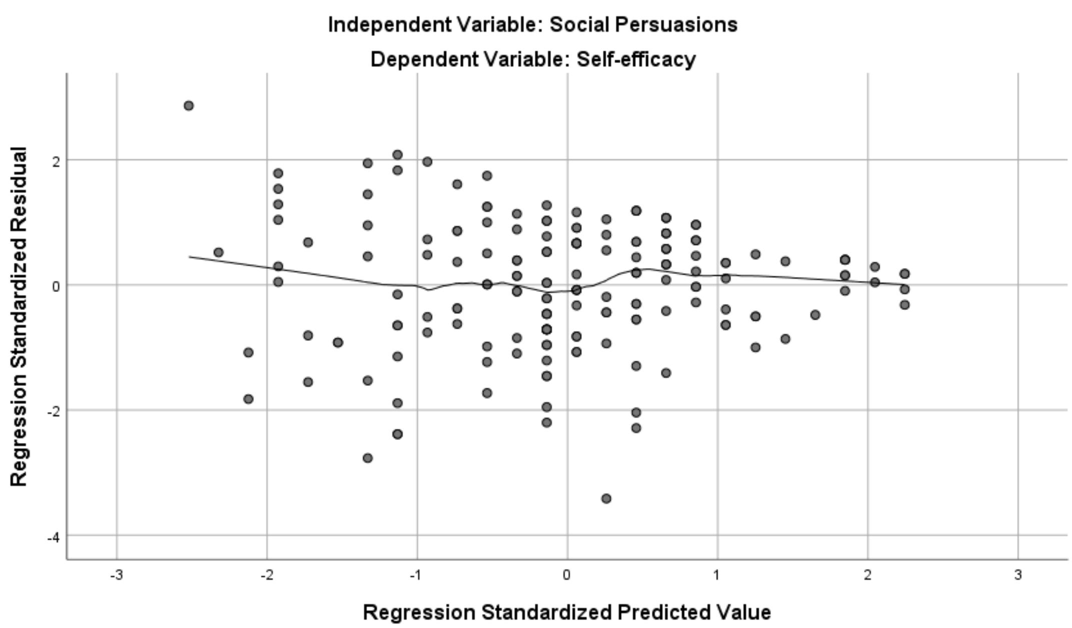


Figure 4.11

Residual plot of physiological states on self-efficacy.

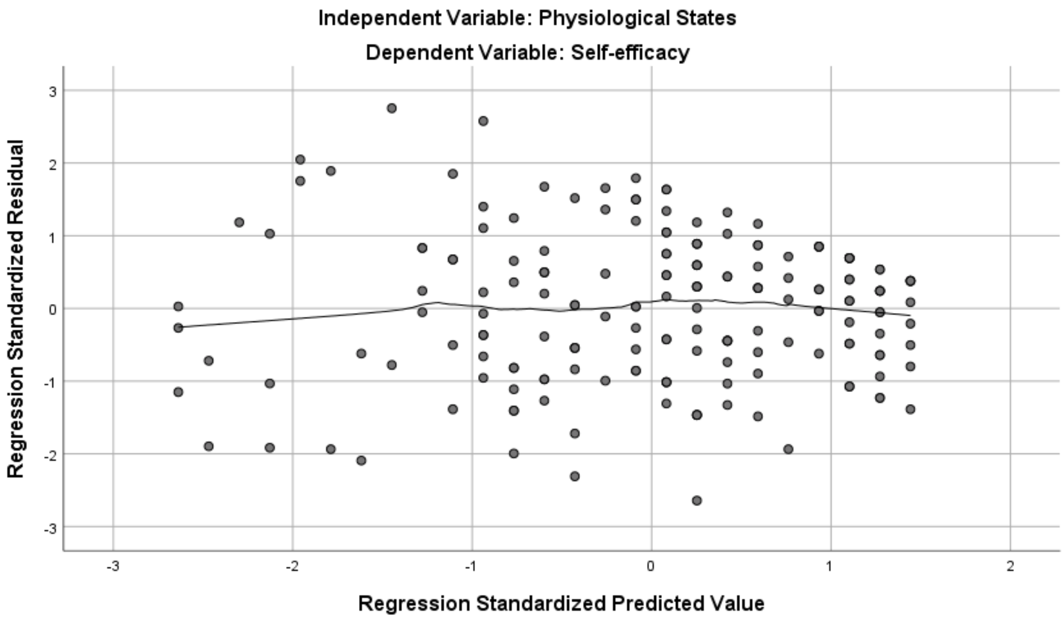


Figure 4.12

Normal Q-Q plot of prior knowledge of computer programming on CS1 achievement.

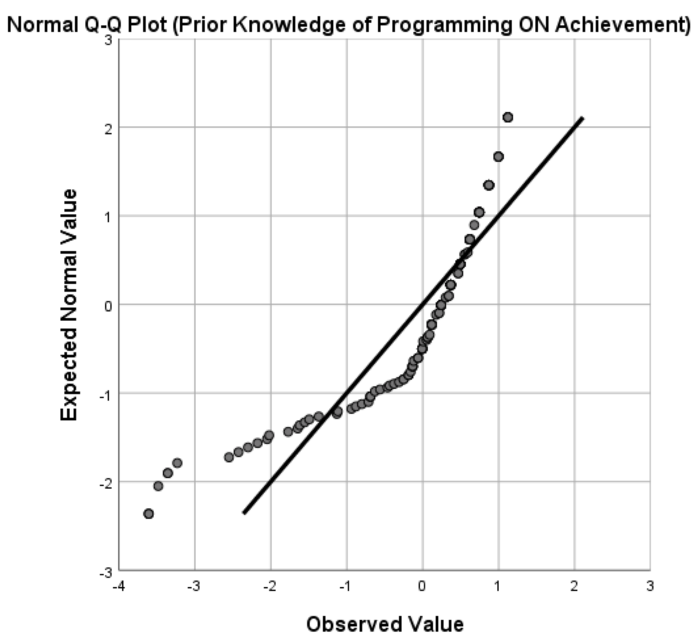


Figure 4.13

Normal Q-Q plot of self-efficacy on CS1 achievement.

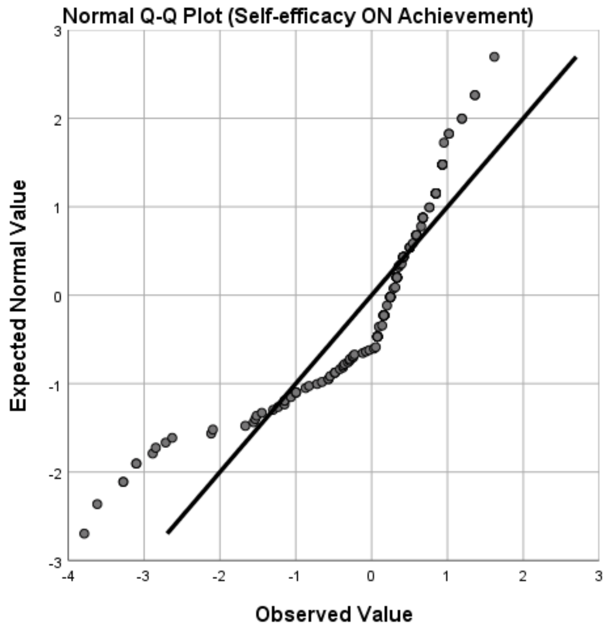


Figure 4.14

Normal Q-Q plot of CS1 achievement on post-CS1 interest.

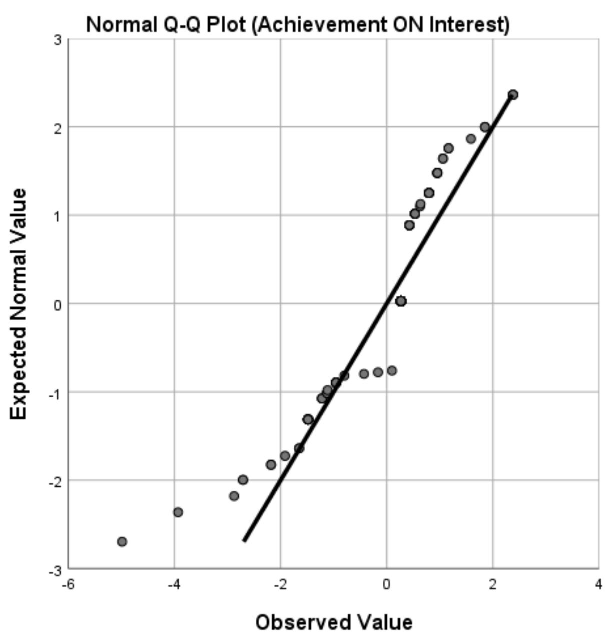


Figure 4.15

Normal Q-Q plot of self-efficacy on post-CS1 interest.

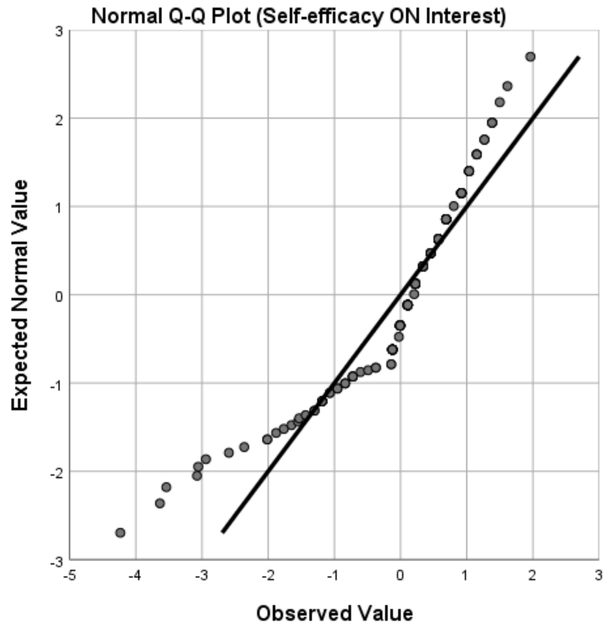


Figure 4.16

Normal Q-Q plot of prior knowledge of computer programming on self-efficacy.

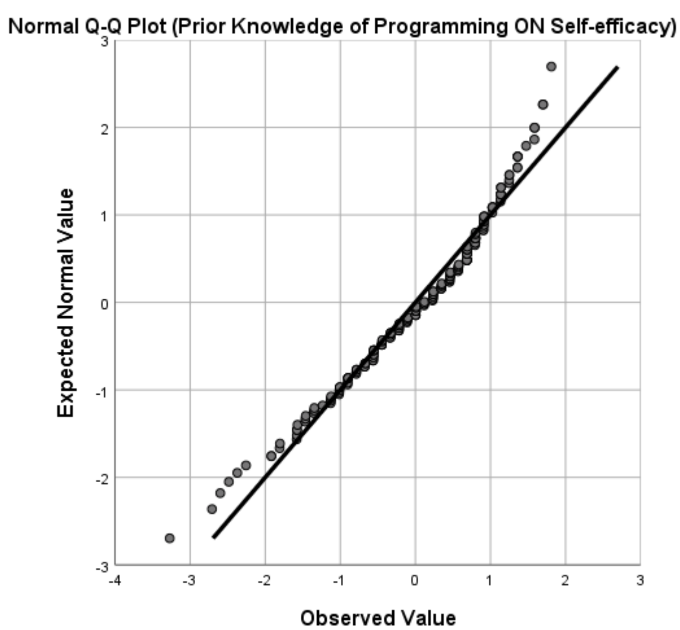


Figure 4.17

Normal Q-Q plot of mastery experience on self-efficacy.

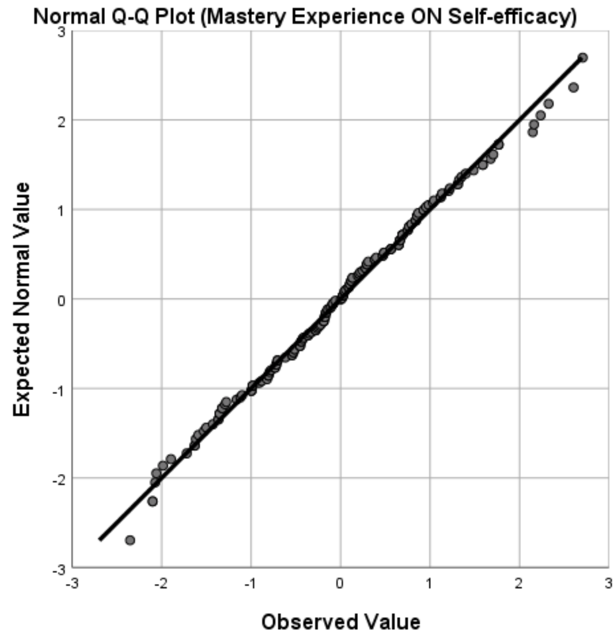


Figure 4.18

Normal Q-Q plot of vicarious experience on self-efficacy.

Figure 4.19

Normal Q-Q plot of social persuasions on self-efficacy.

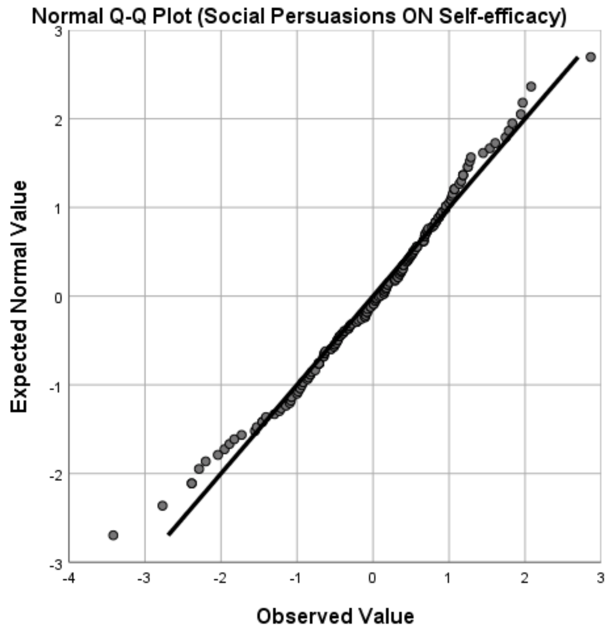


Figure 4.20

Normal Q-Q plot of physiological states on self-efficacy.

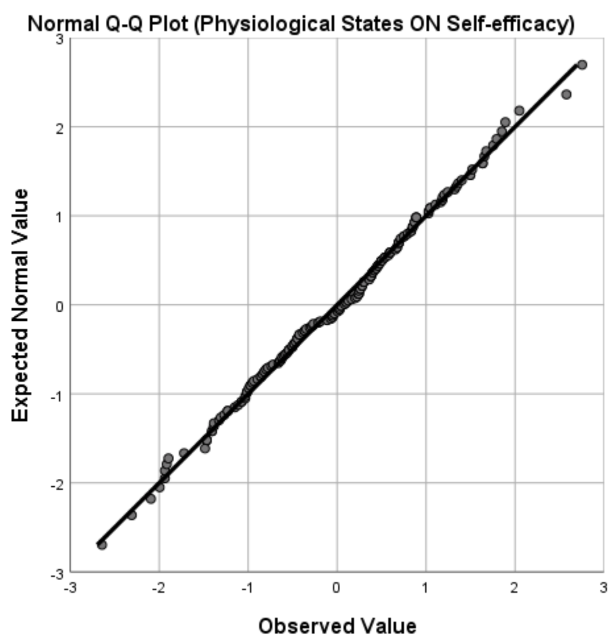


Figure 4.21

Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for all students.

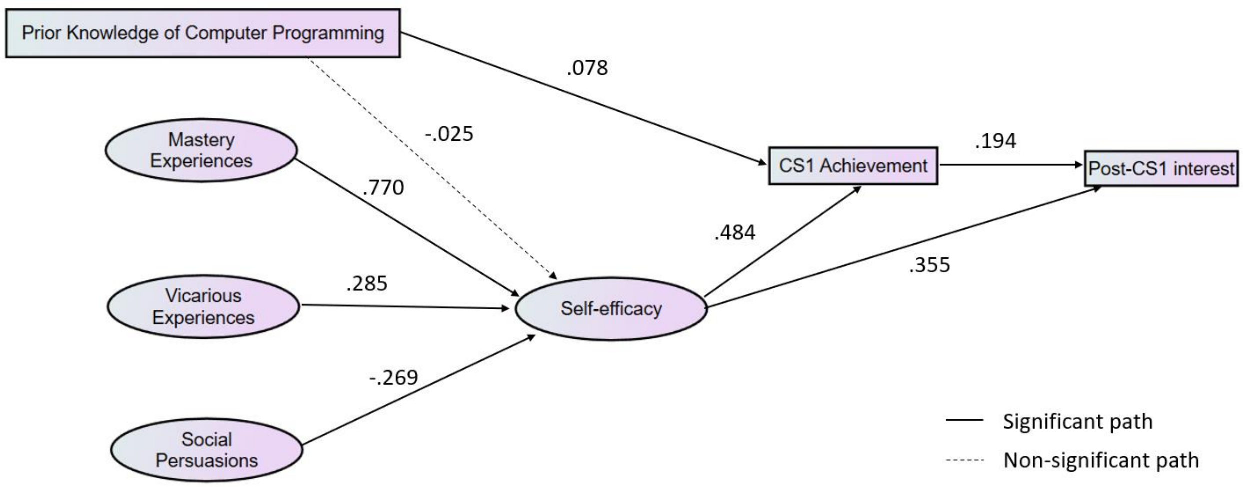


Figure 4.22

Model of CS1 achievement and post-CS1 interest with path labels.

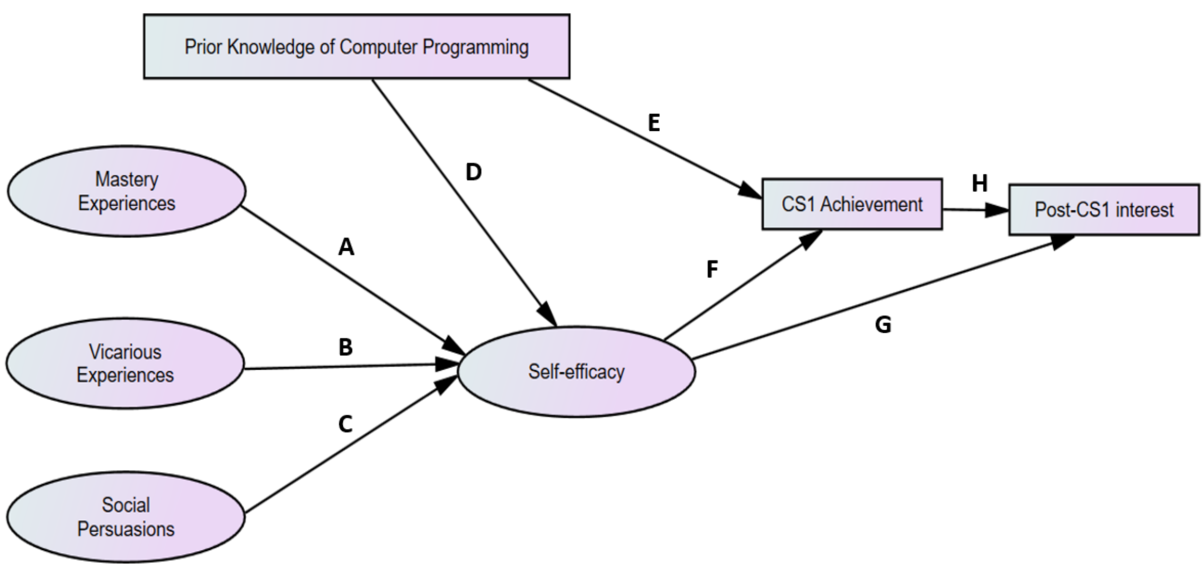


Figure 4.23

Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for women.

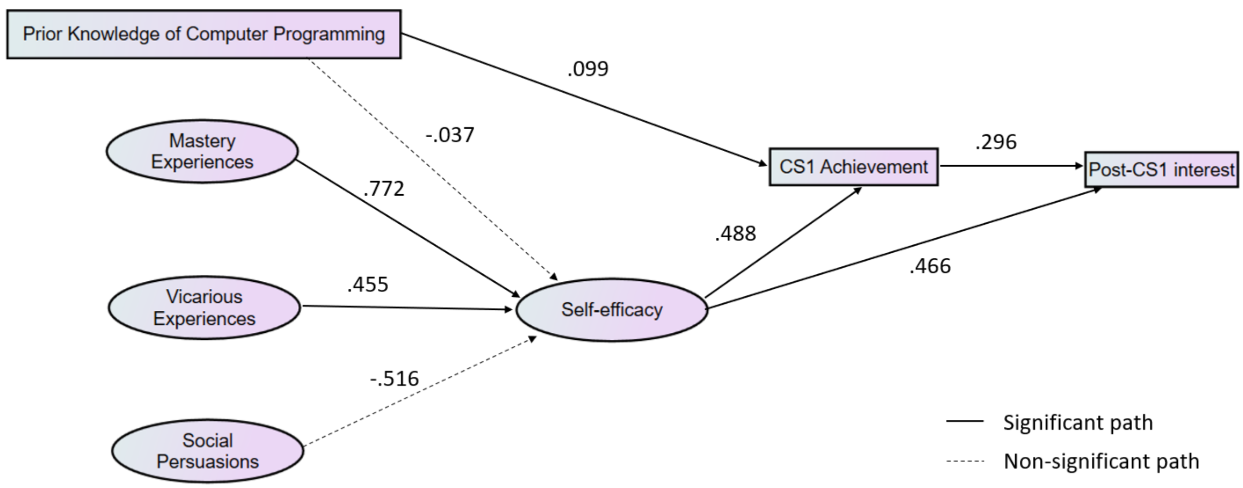
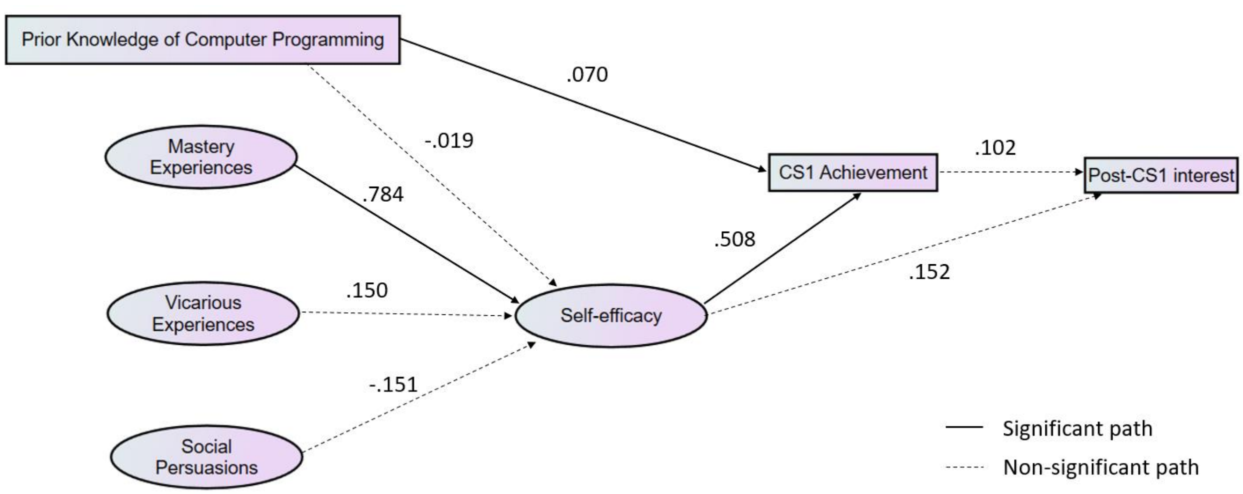


Figure 4.24

Unstandardized parameter estimates for model of CS1 achievement and post-CS1 interest for men.



APPENDIX C: CODEBOOK FOR STUDY MEASURES

Items from the data collection instruments appear below, along with coded response values. Items from the student pre-test do not appear in this document as per the pre-test designers' wishes.

Survey 1 - Demographics, College Major, CS Background

Variable Name	**Item**	**Response Code**
School		University A; University B
Term		FS20; SS21
S1_Major	Declared major / intended area of study - Selected Choice	Computer Science; Data Science; Math, Science, Technology, or Engineering Other
Major_Coded		1 - Computer Science; 2 - Data Science; 3 - Math, Science, Technology or Engineering; 4 - Other
S1_Ethnicity	Are you Hispanic or Latinx?	Yes; No;

Variable Name	Item	Response Code
		Don’t wish to report
S1_Race	Select one or more of the following races: - Selected Choice	American Indian or Alaska Native; Asian; Black or African American; White; Other; Don't wish to report
Race_Ethnicity		0 - Didn’t report; 1 - Hispanic or Latinx; 2 - American Indian or Alaska Native; 3 - Asian; 4 - Black or African American; 5 - White; 6 - Two or more races
S1_Gender_ID	Gender identification: - Selected Choice	Man; Woman; Non-binary; Self-describe; Don't wish to report
Gender_Coded		0 - Man 1 - Woman; 2 - Non-binary;

Variable Name	Item	Response Code
		3 - Self-describe; 4 - Don't wish to report
S1_PE_APCSA	Took the AP Computer Science A course (object-oriented programming)	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high
S1_PE_APCSP	Took the AP Computer Science Principles course (survey of computer science topics)	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high
S1_PE_CS	Took a non-AP course focused on computer programming	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this

Variable Name	Item	Response Code
		experience was very high
S1_PE_PL	Learned a computer programming language	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high
S1_PE_SWHW	Engaged in software or hardware related projects	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high
S1_PE_Informal	Took part in student groups related to computing (e.g., robotics club, competitive programming group, computer club, etc.)	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high

Variable Name	Item	Response Code
S1_PE_Online	Completed an online course related to computing (e.g., MOOC)	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high
S1_PE_Summer	Attended a workshop or summer program focused on computing	0 - No. I did not have this experience.; 1 - Yes. My participation / success in this experience was minimal; 2 - Yes. My participation / success in this experience was moderate; 3 - Yes. My participation / success in this experience was very high

Computer Programming Pre-Test

Variable Name	Item	Response Code
PT_Score	13-item pre-test (not shown at request of pre-test authors); Reduced to 10-items after completion of a CFA	Scores range from 0-10

Survey 2 –Motivational Items

Variable Name	Item	Response Code
S2_SoSE_Mastery1	I make excellent grades on computer science quizzes and exams.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Mastery2	I have always been successful with computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Mastery3_R	Even when I study hard, I	1 - Strongly disagree;

Variable Name	Item	Response Code
	do poorly in computer science.	2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Mastery3	REVERSE CODED	1 - Strongly disagree; 2 - Somewhat agree; 3 - Neither agree nor disagree; 4 - Somewhat disagree; 5 - Strongly disagree
S2_SoSE_Mastery4	I can successfully apply algorithms to novel problems in programming tasks.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Mastery5	I do well on computer science assignments and programming projects.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Mastery6	I do well on even the most difficult computer science assignments	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree;

Variable Name	Item	Response Code
	and programming projects.	4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious1	Seeing adults in my life do well in computer science pushes me to do better.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious2	When I see how my computer science professor or TA solves a problem, I can picture myself solving the problem in the same way.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious3	Seeing students do better than me in computer science pushes me to do better.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious4	When I see how another student solves a problem in computer	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree;

Variable Name	Item	Response Code
	science, I can see myself solving the problem in the same way.	4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious5	I imagine myself working through challenging computer science problems successfully.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Vicarious6	I compete with myself in computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social1	My computer science professor or TA has told me that I am good at learning computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social2	People have told me that I have a talent for computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree;

Variable Name	Item	Response Code
		4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social3	Adults in my family have told me how good I am with computers and computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social4	I have been praised for my ability in computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social5	Other students have told me that I'm good at learning new ideas in computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Social6	My student peers like to work with me in computer science because they think I'm good at it.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree

Variable Name	Item	Response Code
S2_SoSE_Phys1	Just being in computer science class makes me feel stressed and nervous.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Phys2	Doing computer science work takes all of my energy.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Phys3	I start to feel stressed out as soon as I begin my computer science work.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Phys4	My mind goes blank and I am unable to think clearly when doing computer science work.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Phys5	I get depressed when I think about learning	1 - Strongly disagree; 2 - Somewhat disagree;

Variable Name	Item	Response Code
	computer science.	3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SoSE_Phys6	My whole body becomes tense when I have to do computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SE1	I can master the content and skills taught in this computer science class.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SE2	I can master the content in even the most challenging computer science course.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SE3	I can do almost all the work in this computer science class if I don't give up.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree;

Variable Name	Item	Response Code
		5 - Strongly agree
S2_SE4	I can do an excellent job on computer science-related problems and tasks assigned this semester.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SE5	Even if the concepts in this computer science class are hard, I can learn them.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_SE6	I can earn a good grade in my computer science-related courses.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_UV1	Being good at computer science pays off, because it is simply needed for school. UV_School	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree

Variable Name	Item	Response Code
S2_UV2	Computer science is directly applicable in everyday life.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree;
	UV_Daily_Life	5 - Strongly agree
S2_UV3	If I know a lot about computer science, I will leave a good impression on my fellow students.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	UV_Social	
S2_UV4	Good grades in computer science class can be of great value to me later on.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	UV_Job	
S2_UV5	Learning computer science is worthwhile, because it improves my job and career chances.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree;

Variable Name	Item	Response Code
		5 - Strongly agree
	UV_Job	
S2_UV6	Computer science content will help me in my life.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	UV_General	
S2_UV7	I will often need computer science in my life.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	UV_General	
S2_Cost1	I often feel mentally fatigued after doing computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	Effort_cost	
S2_Cost2	Dealing with computer science drains a lot of my energy.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	Effort_cost	

Variable Name	Item	Response Code
S2_Cost3	Learning computer science exhausts me. Effort_cost	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_Cost4	I'd rather not do computer science, because it only worries me. Emotional_cost	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_Cost5	When I deal with computer science, I get annoyed. Emotional_cost	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_Cost6	Computer science is a real burden to me. Emotional_cost	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
S2_Cost7	I have to give up other activities that I like to	1 - Strongly disagree; 2 - Somewhat disagree;

Variable Name	Item	Response Code
	be successful at computer science.	3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	Opportunity_cost	
S2_Cost8	I have to give up a lot to do well at computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	Opportunity_cost	
S2_Cost9	I'd have to sacrifice a lot of free time to be good at computer science.	1 - Strongly disagree; 2 - Somewhat disagree; 3 - Neither agree nor disagree; 4 - Somewhat agree; 5 - Strongly agree
	Opportunity_cost	

Survey 3 - Grade expectation and post-CS1 interest in continued study of CS

Variable Name	Item	Response Code
S3_Grade_Exp	what is your expectation for your final grade in the course?	4 - A; 3 - B; 2 - C; 0 - Below a C

Variable Name	Item	Response Code
S3_Interest	rate your interest in continuing to study computer science.	1 - I will definitely not take more computer science courses; 2 - It is unlikely that I will take more computer science courses; 3 - I am uncertain whether I will take more computer science courses; 4 - It is unlikely that I will take more computer science courses; 5 - I will definitely take more computer science courses
GPA		A+, A 4.0 A- 3.7 B+ 3.3 B 3.0 B- 2.7 C+ 2.3 C 2.0 C- 1.7 D+ 1.3 D 1.0 D- 0.7 E 0.0

APPENDIX D: MPLUS CODE

Code for the Monte Carlo simulation of the CFA used to determine power.

```
TITLE:       CFA MONTE CARLO POWER ANALYSIS CS1 MODEL
MONTECARLO: NAMES ARE ME1-ME6 VE2 VE4 SP1-SP6 SE1-SE6;
            NOBSERVATIONS=182;
            NREPS=2500;
            SEED=3784;
ANALYSIS:   ESTIMATOR = MLR;
MODEL POPULATION:
            MASEXP BY ME1*1 ME2*.928 ME3*.878 ME4*.698 ME5*.748
               ME6*.898;
            VICEXP BY VE2*1 VE4*1.026;
            SOCPER BY SP1*1 SP2*1.664 SP3*1.653 SP4*1.675
               SP5*1.414 SP6*1.325;
            SE BY SE1*1 SE2*1.059 SE3*.894 SE4*1.145 SE5*.869
               SE6*1.097;
            MASEXP@1; VICEXP@1; SOCPER@1; SE@1;
            ME1*.559; ME2*.759; ME3*.588; ME4*.483; ME5*.400;
               ME6*.587;
            VE2*.557; VE4*.342;
            SP1*.647; SP2*.333; SP3*.491; SP4*.319; SP5*.530;
               SP6*.506;
            SE1*.271; SE2*.522; SE3*.264; SE4*.253; SE5*.283;
               SE6*.256;
            MASEXP WITH VICEXP*.333;
            MASEXP WITH SOCPER*.316;
            MASEXP WITH SE*.585;
            VICEXP WITH SOCPER*.185;
            VICEXP WITH SE*.326;
            SOCPER WITH SE*.211;
 MODEL:
            MASEXP BY ME1*1 ME2*.928 ME3*.878 ME4*.698 ME5*.748
               ME6*.898;
            VICEXP BY VE2*1 VE4*1.026;
            SOCPER BY SP1*1 SP2*1.664 SP3*1.653 SP4*1.675
               SP5*1.414 SP6*1.325;
            SE BY SE1*1 SE2*1.059 SE3*.894 SE4*1.145 SE5*.869
               SE6*1.097;
            MASEXP@1; VICEXP@1; SOCPER@1; SE@1;
            ME1*.559; ME2*.759; ME3*.588; ME4*.483; ME5*.400;
               ME6*.587;
            VE2*.557; VE4*.342;
```

```
               SP1*.647; SP2*.333; SP3*.491; SP4*.319; SP5*.530;
                  SP6*.506;
               SE1*.271; SE2*.522; SE3*.264; SE4*.253; SE5*.283;
                  SE6*.256;
               MASEXP WITH VICEXP*.333;
               MASEXP WITH SOCPER*.316;
               MASEXP WITH SE*.585;
               VICEXP WITH SOCPER*.185;
               VICEXP WITH SE*.326;
               SOCPER WITH SE*.211;
OUTPUT:        TECH9;
```

Code for the Monte Carlo simulation of the SR with data generated from the structural regression in the study.

```
TITLE:       SR MONTE CARLO POWER ANALYSIS CS1 MODEL
MONTECARLO:
             NAMES = PKCP ME1-ME6 VE2 VE4 SP1-SP6 SE1-SE6
              INTEREST ACHIEVE;
             NOBSERVATIONS=450;
             NREPS=250;
             SEED=50895;
             POPULATION=REALMC.dat;
             COVERAGE=REALMC.dat;
ANALYSIS:
             TYPE = GENERAL;
             ESTIMATOR = MLR;
MODEL POPULATION:
             MASEXP BY ME1 ME2 ME3 ME4 ME5 ME6;
             VICEXP BY VE2 VE4;
             SOCPER BY SP1 SP2 SP3 SP4 SP5 SP6;
             SE BY SE1 SE2 SE3 SE4 SE5 SE6;
             SE ON MASEXP VICEXP SOCPER;
             ACHIEVE ON PKCP SE;
             INTEREST ON SE ACHIEVE;
             PKCP WITH MASEXP VICEXP SOCPER;
             ME5 WITH ME6;
             SP5 WITH SP6;
             SE1 WITH SE2;
MODEL:
             MASEXP BY ME1 ME2 ME3 ME4 ME5 ME6;
             VICEXP BY VE2 VE4;
             SOCPER BY SP1 SP2 SP3 SP4 SP5 SP6;
             SE BY SE1 SE2 SE3 SE4 SE5 SE6;
             SE ON MASEXP VICEXP SOCPER;
             ACHIEVE ON PKCP SE;
             INTEREST ON SE ACHIEVE;
             PKCP WITH MASEXP VICEXP SOCPER;
             ME5 WITH ME6;
             SP5 WITH SP6;
             SE1 WITH SE2;
OUTPUT:      TECH9;
```

Code for the CFA of motivational factors.

```
TITLE:        CFA MOTIVATIONAL FACTORS CS1 STUDY
DATA:
              file=CS1_FINAL.dat;
VARIABLE:
              NAMES = ID TERM SCHOOL MAJOR RACEETH GENDER APCSA
                APCSP ANYCS PRLANG SWHW INFORM ONLINE SUMMER
                APHIGH PRHIGH PRAVG PKCP ME1-ME6 VE1-VE6 SP1-SP6
                PS1-PS6 SE1-SE6 UV1-UV7 CST1-CST9 GRADEEXP
                ACHIEVE INTEREST;
              USEVAR = ME1-ME6 VE2 VE4 SP1-SP6 SE1-SE6;
              MISSING = all(-999);
ANALYSIS:
              TYPE = general;
              ESTIMATOR = MLR;
MODEL:
              MASTRY BY ME1@1 ME2 ME3 ME4 ME5 ME6;
              VICARS BY VE2@1 VE4;
              SOCPER BY SP1@1 SP2 SP3 SP4 SP5 SP6;
              SE BY SE1@1 SE2 SE3 SE4 SE5 SE6;
              MASTRY WITH VICARS;
              MASTRY WITH SOCPER;
              MASTRY WITH SE;
              VICARS WITH SOCPER;
              VICARS WITH SE;
              SOCPER WITH SE;
              ME5 WITH ME6;
              SP5 WITH SP6;
              SE1 WITH SE2;
OUTPUT:       standardized sampstat modindices;
```

Code for the structural regression of the model of CS1 achievement and post-CS1 interest for all students.

```
TITLE:    STRUCTURAL REGRESSION OF PRIOR CP KNOWLEDGE, SELF-
            EFFICACY, ACHIEVEMENT, AND INTEREST
DATA:
          FILE=CS1_FINAL.dat;
VARIABLE:
          NAMES = ID TERM SCHOOL MAJOR RACEETH GENDER APCSA
            APCSP ANYCS PRLANG SWHW INFORM ONLINE SUMMER APHIGH
            PRHIGH PRAVG PKCP ME1-ME6 VE1-VE6 SP1-SP6 PS1-PS6
            SE1-SE6 UV1-UV7 CST1-CST9 GRADEEXP ACHIEVE INTEREST;
          USEVAR = PKCP ME1-ME6 VE2 VE4 SP1-SP6 SE1-SE6 INTEREST
            ACHIEVE;
          MISSING = all(-999);
ANALYSIS:
          ESTIMATOR = MLR;
MODEL:
          MASTRY BY ME1@1 ME2 ME3 ME4 ME5 ME6;
          VICARS BY VE2@1 VE4;
          SOCPER BY SP1@1 SP2 SP3 SP4 SP5 SP6;
          SE BY SE1@1 SE2 SE3 SE4 SE5 SE6;
          SE ON MASTRY VICARS SOCPER PKCP;
          ACHIEVE ON SE PKCP;
          INTEREST ON SE ACHIEVE;
          PKCP WITH MASTRY VICARS SOCPER;
          ME5 WITH ME6;
          SP5 WITH SP6;
          SE1 WITH SE2;
OUTPUT:   sampstat stdyx tech4 cinterval modindices(5);
SAVEDATA: estimates=REALMC.dat;
```

Code for the structural regression of the model of CS1 achievement and post-CS1 interest for women only.

```
TITLE:    STRUCTURAL REGRESSION OF PRIOR CP KNOWLEDGE, SELF-
            EFFICACY, ACHIEVEMENT, AND INTEREST
DATA:
          FILE=CS1_FINAL.dat;
VARIABLE:
          NAMES = ID TERM SCHOOL MAJOR RACEETH GENDER APCSA
            APCSP ANYCS PRLANG SWHW INFORM ONLINE SUMMER APHIGH
            PRHIGH PRAVG PKCP ME1-ME6 VE1-VE6 SP1-SP6 PS1-PS6
            SE1-SE6 UV1-UV7 CST1-CST9 GRADEEXP ACHIEVE INTEREST;
          USEVAR = PKCP ME1-ME6 VE2 VE4 SP1-SP6 SE1-SE6 INTEREST
            ACHIEVE;
          MISSING = all(-999);
          USEOBS = (GENDER EQ 1);
ANALYSIS:
          TYPE = GENERAL;
          ESTIMATOR = MLR;
MODEL:
          MASTRY BY ME1@1 ME2 ME3 ME4 ME5 ME6;
          VICARS BY VE2@1 VE4;
          SOCPER BY SP1@1 SP2 SP3 SP4 SP5 SP6;
          SE BY SE1@1 SE2 SE3 SE4 SE5 SE6;
          SE ON MASTRY VICARS;
          SE ON SOCPER;
          SE ON PKCP;
          ACHIEVE ON SE;
          ACHIEVE ON PKCP;
          INTEREST ON SE;
          INTEREST ON ACHIEVE;
          PKCP WITH MASTRY VICARS SOCPER;
          ME5 WITH ME6;
          SP5 WITH SP6;
          SE1 WITH SE2;
OUTPUT:   stdyx tech4 cinterval modindices(5);
```

APPENDIX E: LITERATURE SUPPORTING EXPLORATION OF ADDITIONAL CONSTRUCTS

E.1 Other Factors Related to CS1 Achievement and Post-CS1 Interest

The current study of CS1 achievement and post-CS1 interest was focused on self-efficacy, the sources of self-efficacy, and prior knowledge of computer programming. A thorough review of the literature on the CS1 course suggest other factors that could have been included. These include social support for CS (Beyer et al., 2003; Margolis & Fisher, 2003; Petersen et al., 2016; Wang et al., 2015; Zarrett & Malanchuk, 2005), sense of belonging (Cheryan et al., 2017; Lewis et al., 2017; Lewis et al., 2019; Sankar et al., 2015; Tellhed et al., 2017), utility value (Gaspard et al., 2017), and cost (Gaspard et al., 2017; Petersen et al., 2016). Additionally, this study reviewed multiple measures of student background with computing, electing to focus specifically on prior knowledge of computer programming over self-reports of prior experiences with computer science. Lastly, certain cognitive factors have been shown to impact CS1 achievement but do not differ by gender. Due to this, the review of cognitive abilities related to the CS1 course were also excluded from the main text. Further information on the excluded factors can be found in the sections that follow.

E.1.1 Social Supports for Computer Science

Social supports and barriers in computer science can strengthen or weaken student self-efficacy for a given task. The main sources of these supports and barriers are parents, peers, and teachers, and these manifest for students in terms of encouragement or discouragement for participation in CS, modeling of success in CS careers, or promoting stereotypes about the student's abilities with regards to CS. Social supports are important for the development of interest in CS careers (Wang et al., 2015) and persistence in CS programs (Beyer et al., 2003). In STEM subjects, positive family messaging regarding student potential in those subjects had an important impact on women (Astin & Sax, 1996; Moakler Jr & Kim, 2014) and had higher impact when coming from other women within the family (Nauta & Kokaly, 2001). Family support has been suggested as the most important source of social support for women in computer science, impacting both the choice of computer science as a college major (Margolis & Fisher, 2003; Wang et al., 2015), and persistence in the CS major (Beyer et al., 2003). While not as powerful an influence as parental support, there is also evidence that teacher engagement with women in science courses is an important component of encouraging women to choose CS and STEM careers (Google & Gallup, 2016; Leedy et al., 2003).

In this study, social supports were assumed to be captured by the social persuasions factor that was utilized as a part of the sources of self-efficacy. For these reasons, no additional measurement related to other conceptions of social supports were used.

E.1.2 Sense of Belonging and Perceived Similarity in Computer Science

Perceived similarity between peers and sense of belonging are two additional important and related factors that impact students in the CS1 course and broader CS undergraduate environment. Computer science environments are typically described as having masculine and nerdy traits, which negatively impacts interest for women in CS (Cheryan et al., 2009; Cheryan et al., 2017; Inzlicht & Good, 2010; Tellhed et al., 2017). Due to the impact the CS environment has on marginalized students, it is important for these students to see others like them succeeding to encourage further participation. Cheryan and Plaut (2010) found that perceived similarity was a significant mediator in predicting interest in CS for women. It has been proposed that having a peer group in subjects like CS may inoculate students from the effects of negative stereotypes within the field (Dasgupta, 2011). Women in STEM courses are also much more likely to persist in the major given a positive environment that encourages a sense of belonging (Lewis et al., 2017; Petersen et al., 2016; Sax et al., 2018; Tellhed et al., 2017). Prior research has found that a lack of student comfort was an important element in predicting student dropout in CS1 (Kinnunen & Malmi, 2006; Wilson & Shrock, 2001), and that peer groups within CS could help reduce the likelihood of dropout (Petersen et al., 2016). In a study of an all-women's CS classroom, students reported more comfort in collaborating with peers, greater support from peers and instructors, more confidence in acquired CS knowledge, and greater sense of belonging in the classroom (Ying et al., 2021).

Impacting perceived similarity and belonging in CS are perceptions about the nature of computing work and negative stereotypes about CS and CS participants. Several studies have shown that individuals who see CS as being an individual pursuit are less likely to have interest in CS as a career path (Weinberger, 2004; Wilson, 2002). Similarly, students who have a communal goal orientation do not see an alignment of values within the environment of computer science and are less likely to participate (Lewis et al., 2019). For women and other marginalized students, the connection between collaboration and community in the environment of CS has been shown to be particularly important (Sax et al., 2018). The types of careers that women choose are often influenced by gender role stereotypes (Eccles et al., 1990). These stereotypes are often held by parents, which in turn impacts which ideas are communicated to young boys and girls about the roles they will best fit into as students and in the workplace. Based on this Wang et al. (2015) noted that the perception that computer science does not positively impact social causes has a negative effect on women's interest in CS. Other stereotypes about computer science and computer scientists also have a larger impact on women. Beyer et al (2003) found that stereotypes about computer science as a nerdy discipline impacted women more than men. Similarly, Margolis and Fisher (2003) found that the view of computer scientists as “reclusive” and as “hackers” was unappealing to women, which damaged their interest in CS. The impact of negative stereotypes of CS and gender role stereotypes have proven hard to overcome in attracting more women to computing fields.

In this study, the focus was on self-efficacy and the sources of self-efficacy which includes some elements related to belonginess for CS1 students. It was decided that a unique measure of belongingness would not be able to be used as it might overlap with the existing measures in the study.

E.1.3 Task Value for CS and Costs Associated with CS Participation

Wigfield and Eccles (1992) described task value as the interest, importance, or utility of a given task to a specific student. Value for a task consists of subjective beliefs about the reasons for engaging with a given task. If a student values a particular activity highly and can see the relevance of the task in relation to their goals and interests, they are more likely to persist on the task. Value for a task includes four component elements: intrinsic value, utility value, attainment value, and cost (Eccles, 2005, Eccles et al., 1983). Intrinsic value is the enjoyment that is gained from participating in the task. Utility value is the usefulness of the task for helping one achieve their goals. Attainment value is the connection between the task and the individual’s identity. Cost is the potential negative impact from participating in the task. In many fields, task value positively impacts student interest for continued study, but not academic achievement (Bong, 2001; Eccles, 2005; Gaspard et al., 2019; Meece, Eccles, & Wigfield, 1990; Zarrett & Malanchuk, 2005). There is evidence that task value predicts interest in STEM subjects for women (Bong, 2001; Gaspard et al., 2019; Zarrett & Malanchuk, 2005). It has been found that cost impacts college students negatively with regards to achievement and in their interest in pursuing science careers and advanced degrees (Battle & Wigfield, 2003; Perez et al., 2014). Beyer et al. (2004) found that women expressed value for a career in which they could help others, work with people, and have flexible work hours, whereas value in a career for men was expressed as an opportunity to make money. The authors also found that women in their study did not believe that having a family and a career in computer science were compatible, which suggests low value for CS-related tasks. Other studies have found that having a family orientation has a negative relationship to CS (Beyer, 2014; Sax et al., 2017). These differing values impacted the interest of women in taking further computer science courses. Students also cite high effort cost as a chief reason for dropping out during their CS1 course experience (Petersen et al., 2016; Smith et al., 2013).

I chose to measure two specific elements of task value based on work suggesting a relationship between both utility value and cost and the dependent variables in this study. To measure utility value, I considered an initial set of twelve items from an instrument devised by Gaspard et al. (2015). This scale divides the twelve items into sub-groups based on areas of life (e.g., work, school, day-to-day tasks) where students may find the content useful. The authors used a 4-point Likert scale in their work, but I used a 5-point Likert scale (1=Strongly disagree to 5=Strongly agree) to match the other instruments used in this study. A scale reliability score, ρ, was used by Gaspard et al. (2015) in the place of Cronbach's alpha to report internal reliability metrics. These values are as follows: the two items for utility for school ($\rho = .52$), the three items for utility for daily life ($\rho = .83$), the three items for social utility ($\rho = .76$), the two items for utility for job ($\rho = .68$) and the two items for general utility for future life ($\rho = .79$). In this study, I selected seven of the twelve items to capture a single measure of utility value from the sub-scales above. Items were selected to get coverage of all components of utility value while reducing the number of total items that would be used in the survey instrument.

Cost was measured using nine items from the scale developed by Gaspard et al. (2015). Like the measurement of utility value, these items were originally measured using a 4-point scale but will be measured using a 5-point Likert scale (1=Strongly disagree to 5=Strongly agree) in this study. The authors reported reliability for three sub-scales of cost as follows: three items for opportunity cost ($\rho = .83$), four items for effort required ($\rho = .90$), and four items for emotional cost ($\rho = .87$). The items for this measure can be found in appendix C. The three-item effort cost scale had an observed reliability α=.91, the emotional cost scale was α=.85, and the opportunity cost scale was α=.89.

Both task value and cost were considered as part of this study but ultimately excluded due to measurement reasons. Due to the cross loading between three of the physiological states items with effort cost and three of the physiological states items with emotional cost, this provided further evidence to remove both the cost items and physiological states from future consideration in the model. More information on the descriptive and correlational data related to task value and cost in this study can be found in appendix F.

E.1.4 Prior Experiences with Computer Science

Prior experience with computer science is an important factor that impacts student choice and performance in CS. Cheryan et al. (2017) found that a lack of early experience with computing was a predictor of participation in STEM and CS. This effects all students regardless of gender, but the significance of early computing exposure has been shown to be important for girls specifically. Generally, girls have less exposure to computing at an early age (Margolis &

Fisher, 2003). Wang et al (2015) found that having taken a CS course before college had a greater impact on women than on men with regards to pursuit of CS in college. The authors found that women who took the AP CS A exam were 38% more likely to pursue a computing degree in college. In addition, other work has shown that CS1 courses that are misaligned with women's prior experiences with CS was one of the main reasons for their eventual departure from the major (Roberts et al., 2012). The quality of the experiences is also important in determining the degree of interest a student will have with CS. There is evidence that early negative experiences greatly reduce the interest of women in CS, and early positive experiences increase CS interest for women (Bernstein, 1991; Beyer, 2014).

In this study, I collected data about student prior experiences with computer programming and computer science using items based on the CRA Data Buddies survey. Aspects of student experience with computer science that were captured by this self-reported measure were separated into 7 categories. These included experience with the Advanced Placement Computer Science A course, experience with the Advanced Placement Computer Science Principles course, experience with a non-AP computer programming course, having learned a programming language outside of a school setting, having worked on software or hardware projects, having engaged in informal CS experiences such as after-school clubs or robotics teams, and having participated in summer programs focused on CS. For each of these groups, students were asked to rate their experience using a four-point scale (0 - No experience; 1 - Minimal experience / success; 2 - Moderate experience / success; 3 - Very high experience / success).

In this study, data was collected for prior experiences of students with computer science, but this measure was not used in lieu of data collected about prior knowledge of computer programming. As mentioned in the main text, there are measurement concerns regarding the accuracy of self-reports of student experience (Dochy et al., 1999), which are exacerbated in CS due to the tendency for women to devalue their own CS experiences (Ashcraft et al., 2012; Beyer, 2014). Beyond these concerns, the potential for overlapping measurement between prior experience and mastery experience measures (Britner & Pajares, 2006; Schunk & Pajares, 2005) served as a final reason for excluding prior experience measures for computer science in the study. Further information about the prior experiences data can be found in appendix F.

E.1.5 Cognitive Factors in CS1

Cognitive factors are those that describe individual intellectual abilities or an inherent cognitive attribute. Examples of cognitive abilities include general intelligence, quantitative ability, spatial ability, verbal ability, and problem-solving ability. Prior research has found that some of these abilities are related to performance in first-year computer science courses (Bergin & Reilly, 2005; Jones & Burnett, 2008; Lishinski et al., 2016; Parker et al., 2018; Priess & Hyde, 2010; Wilson & Shrock, 2001). Cognitive factors most relevant to CS1 achievement are reviewed along with specific effects due to gender differences.

As computer science derived from mathematics, it is perhaps not surprising that there is a relationship between quantitative ability and performance in computer science. Wilson and Shrock (2001) conducted a study of students in an introductory CS course and found that a self-report of math background had a significant positive relationship to midterm grades in the course. When considering gender, Beyer et al. (2003) found that there were no differences between men and women in computer science courses with regards to quantitative ability, an outcome that has been echoed in other studies (Linn & Hyde, 1989; Priess & Hyde, 2010). Mathematical ability includes several sub-abilities including spatial ability and problem-solving ability (Kruteskii, 1976). While some prior work has shown gender differences in spatial ability (Lawton, 2010), research exploring reasons for these differences attribute them to strategy usage and prior experience with spatial tasks rather than innate ability differences (Glück & Fitting, 2003; Lawton, 2010; Robert & Héroux, 2004). There is some evidence that spatial ability impacts CS performance (Jones & Burnett, 2008), but differences in spatial ability for CS students were not significantly different when considering gender (Parker et al., 2018). Problem-solving ability appears as a significant predictor of CS1 performance, but studies within CS have shown no difference in problem solving ability by gender (Bergin & Reilly, 2005; Lishinski et al., 2016). As there were no significant gender differences found for cognitive abilities in relation to CS1 achievement, these factors were not included for further study.

APPENDIX F: OPERATIONALIZATION AND PRELIMINARY ANALYSIS OF ADDITIONAL CONSTRUCTS

F.1 Additional Data Analysis for Unused Constructs

Some of the constructs went unused in this study in the final CS1 model but were included in survey instruments in the hopes that they could be utilized in some manner. The following section contains the details of the data analysis for these constructs and further details that suggest why they were not included in the main portion of the study.

F.2 Utility Value for CS and Costs Associated with CS Participation

Selected task value items in the data collection for this study were utility value and cost. Cost was further divided into three sub-scales for effort cost, emotional cost, and opportunity cost. The mean utility value response for the total sample was 4.174 (SD=0.517), for men was 4.169 (SD=0.532), and for women was 4.184 (SD=0.496). The difference between women and men was not significant for utility value (t=-0.181, df=176, p=.857) at the 5%-level. Based on the lack of significant differences by gender and the overall high mean value and low variability for the value measure, there were no further attempts to include utility value in the analysis. The mean effort cost response for the total sample was 3.024 (SD=1.057), for men was 2.91 (SD=1.016), and for women was 3.276 (SD=1.06). The difference between women and men was significant for effort cost (t=-2.307, df=176, p=.022) at the 5%-level. The mean emotional cost response for the total sample was 2.147 (SD=0.915), for men was 2.096 (SD=0.859), and for women was 2.233 (SD=1.013). The difference between women and men was not significant for emotional cost (t=-0.973, df=176, p=.332) at the 5%-level. The mean opportunity cost response for the total sample was 2.967 (SD=1.135), for men was 2.833 (SD=1.06), and for women was 3.21 (SD=1.176). The difference between women and men was significant for opportunity cost (t=-2.216, df=176, p=.028) at the 5%-level. This data can be seen in table F.1.

Significant negative relationships existed between CS1 achievement and each of the cost measures: opportunity cost (r=-.370, p<.001), effort cost (r=-.344, p<.001), and emotional cost (r=-.338, p<.001). Utility value had a moderate positive relationship with post-CS1 interest (r=.210, p=.004). Additionally, post-CS1 interest was significantly related to all three cost measures: emotional cost (r=-.544, p<.001), effort cost (r=-.316, p<.001), and opportunity cost (r=-.296, p<.001).

A CFA of the measurement model including the value and cost latent factors had adequate fit with the sample data ($\chi^2(738) = 1129.758$, p<.001; CFI=.922; RMSEA=.054; SRMR=.0578). A combination of these statistics was used to make a final decision regarding the fit of the measurement model. This model used N=182 data points and was estimated using maximum likelihood with IBM Amos software. While the fit was reasonable, there were some concerns regarding convergent validity of the measurement model. Specifically, there were several strong correlations represented between latent factors: physiological states and effort cost (r=.87); social persuasions and emotional cost (r=.86); social persuasions and effort cost (0.83). To reduce the concern about some of these elements, the task value measures were not included in the model.

In addition to concerns over measurement, the cost items were also excluded from the CS1 model and subsequent structural regressions. Based on the evidence that there were strong correlations between the dependent variables and the cost measures (see tables F.2 and F.3), it was expected that the inclusion of the cost measures would enhance the CS1 model.

Unfortunately, attempts to include relationships between emotional cost and post-CS1 interest, and all three cost metrics and achievement greatly reduced the fit of the model to the data and produced non-significant path coefficient estimates. For these reasons, the cost measures were not included in the final analysis of the CS1 model.

F.3 Prior Experience with Computer Programming

There is no single way to collect information on student background with computer science. Student presence in K-12 computer science courses does not guarantee that the student was successful in those courses. Students may not have understood all course concepts and the quality of the course may not meet high standards. This could lead to future misconceptions about major concepts for students studying computer science. Tests of computer science knowledge may not capture the ability of students to succeed on computer science or computer programming tasks for a variety of reasons which include but are not limited to poor content coverage, misleading question stems, or test anxiety. As it was important in this study to capture a quality measure of student background with computer science and computer programming, multiple measurement tools were considered. These were a measure of prior knowledge of computer programming, and measures of prior experiences with computer science. The measures of prior experiences with computer science were not selected for use in the study. Details about their measure and further analysis supporting their exclusion follows.

F.3.1 Descriptive, Correlational, and Inferential Statistics for Prior Experiences of CS

Student prior experience with computer science was measured by eight items reflecting the level of participation of the students in formal and informal learning experiences focused on CS. For men in the sample, the mean self-report value for having taken the AP CS A course was 1.06 (SD=1.327), having taken the AP CS Principles course was 0.46 (SD=0.980), having taken a non-AP course focused on computer programming was 1.5 (SD=1.343), having learned a computer programming language was 1.83 (SD=1.028), engaged in software of hardware projects was 1.28 (SD=1.101), took part in computing-focused student groups was 0.85 (SD=1.109), completed an online CS course as 0.85 (SD=1.075), and attended a workshop or summer program focused on CS was 0.58 (SD=0.929). For women in the sample, the mean value for AP CS A was 0.76 (SD=1.185), for AP CS Principles was 0.43 (SD=0.926), non-AP CS course was 1.33 (SD=1.359), programming language was 1.69 (SD=1.097), software / hardware projects was 0.91 (SD=1.073), computing student groups was 0.83 (SD=1.129), online CS course was 0.63 (SD=0.966), workshop or summer CS program was 0.63 (SD=1.038). The only significant difference by gender at the 5%-level among these eight experience measures was for participation in software or hardware related projects. Men in the study were significantly more likely to engage in these activities.

It was anticipated that students who had learned how to program in any context before having taken the CS1 course would see the benefit of those prior experiences in their course performance. To test this basic assumption, I conducted simple inferential statistics for these variables for the entire sample of students as well as by gender group. For all students, I separated them into two groups by previous experience learning a programming language. Students who reported a "moderate" or "very high" level of experience were in one group, while those reporting "minimal" or "none" were in the second group. The mean GPA score for students with more exposure to a programming language was 3.574 (SD=0.9) while those students with less previous programming experience had a mean GPA of 3.243 (SD=1.15). The difference in GPA scores between these groups was significant (t=2.117; p=0.036) at the 5%-level. Descriptive statistics for these groups and inferential statistics can be seen in table F.4.

To consider potential gender differences in experiences with programming languages and the relationship to GPA, I conducted one-way ANOVA for men and women in the CS1 course. Student GPA was computed for each reported level of programming experience, and then statistics were computed both between and within groups. For women in the course, mean GPA increased for each higher level of programming experience. These differences across reported levels of experience were significant (F=2.894; p=0.042) at the 5%-level for women. For men in the CS1 course, the mean GPA scores were almost identical for the "moderate" and "none" groups and were just slightly lower than the "very high" and "minimal" groups. Differences across reported levels of experience were not significant (F=.470; p=.704) at any level for men in the study. ANOVA statistics for each group can be found in table F.5.

F.3.2 Comparing Measures of Background with Computing

Using the data collected in the study, I evaluated a subset of the measures for student prior knowledge of computer programming with the goal of selecting one measure that would best capture the knowledge and experiences that students brought to their first computer science course at the university level. The student pre-test of programming knowledge was one attempt to collect this information, and the self-reported measures of prior experiences with various computing-related courses and activities were another. Of the collection of self-reported prior experience measures, I chose to focus on the items that addressed programming specifically. This included experience with the AP CS A course which used the Java programming language and focused on a similar curriculum to the CS1 course, experience in a non-AP computer programming course, and experience learning a programming language in any setting. To evaluate these measures, I ran descriptive and inferential statistics, as well as correlational statistics and simple regressions using student achievement as the dependent variable. All the data used in evaluating these measures can be seen in table F.6.

Looking at data from the entire sample, the pre-test scores were significantly correlated with mastery experiences ($r=.402$, $p<.001$), vicarious experiences ($r=.176$, $p=.017$), social persuasions ($r=.294$, $p<.001$), physiological states ($r=-.429$, $p<.001$), self-efficacy ($r=.298$, $p<.001$), academic achievement ($r=.33$, $p<.001$), and post-CS1 interest ($r=.261$, $p<.001$). This represents all the variables selected for use in the model. For the self-report of AP CS A course participation there were significant correlations with mastery experiences ($r=.268$, $p<.001$), social persuasions ($r=.266$, $p<.001$), physiological states ($r=-.016$, $p=.031$), and academic achievement ($r=.193$, $p=.01$). For experience learning a programming language, there were significant correlations with mastery experiences ($r=.286$, $p<.001$), social persuasions ($r=.392$, $p<.001$), physiological states ($r=-.267$, $p<.001$), self-efficacy ($r=.278$, $p<.001$), academic achievement ($r=.192$, $p=.01$), and post-CS1 interest ($r=.198$, $p=.01$). For the measure of experience in a non-AP computer science course there were no significant relationships with the dependent variables and so I did not pursue this measure further.

There were no significant differences when considering gender for pre-test score, AP CS A participation, or having learned a programming language; however, women scored lower or participated at a lower rate than men for each of these measures. For the AP CS A and programming language measures I conducted one-way ANOVA with CS1 achievement as the main effect. There was no significant effect on academic achievement for either women ($F=1.514$, $p=.219$) or men ($F=2.467$, $p=.066$) based on participation in the AP CS A course. For women, having learned a programming language impacted achievement in the course ($F=2.894$, $p=.042$) but this did not hold for men ($F=.470$, $p=.704$). Full information for the analysis of variance can be seen in table F.5. Based on these results, I only considered learning of a programming language from the self-report statistics and not AP CS A or other non-programming experiences.

Despite concerns about the pre-test instrument and its use in capturing student knowledge of computer programming ability, I elected to use this measure in this study. A post-hoc analysis appears to justify my selection. The correlational and inferential statistics pointed to the pre-test of computer programming knowledge having stronger relationships with the dependent variables. Additionally, while the measure may have flaws, it avoids reliance on individual students to assess their engagement with specific programming-related activities. I referred to the programming language metric in one instance where additional information was needed to support an argument for differences in experience by gender.

F.4 Additional Findings

Most of the findings in this study were driven by the exploration of the model of CS1 achievement and post-CS1 interest and potential moderation by gender. In the preparation for the study, data to support several other latent factors was collected. This data had some interesting preliminary outcomes which may influence future work on gender inequities in the CS1 course. I review these findings below.

For all students in the CS1 course there were significant relationships between the three measures of cost and both academic achievement and post-CS1 interest. Effort cost and opportunity cost showed a strong negative correlation with interest, and each had significantly stronger relationship for women than for men. Prior work has suggested that psychological cost has a negative effect on interest in STEM subjects (Ball et al., 2019; Huang et al., 2018; Perez et al., 2014; Perez et al., 2019). There was no work looking specifically at these effects in the context of computer science. More work should explore how cost is weighed in early CS1 course experiences as it might provide evidence of first-year CS courses as an early departure point. Specifically, future work should examine how students' prior programming experience contributes to their performance on programming assessments and contributes to their perceptions of the costs associated with study of CS.

In addition to cost differences, there were several other areas where the data from this study showed that women and men differed in the CS1 course. It appears across multiple measures of prior CS experiences that men are entering CS programs with more CS experiences and expressed greater quality in those experiences than women are. This aligns with research on gender differences in CS, with men having more access to CS and entering college with more confidence in their computing abilities (Clegg & Trayhurt, 2000; Denner et al., 2014; Frieze & Quesenberry, 2015; Margolis & Fisher, 2003). In this study, men reported more mastery experiences and were supported more in their pursuits. Self-efficacy also showed significant gender differences with men reporting higher overall self-efficacy. While there were non-significant differences for many of the prior experience measures, all but one of those measures featured a higher self-report value for men than for women. Taken together, the gender differences in the number of and quality of CS experiences prior to undergraduate studies speaks to lingering issues with support for girls in computing career paths. There has been a major push to engage girls with computing content in K-12 classrooms (College Board, 2020; Code.org et al., 2020; Google & Gallup, 2016), yet the effects of this push from schools, universities, and national organizations does not appear to have changed the culture significantly. Future work in pre-university computer science should explore the impact of formative experiences on girls' interest in computing and identify avenues for greater validation of the girls' abilities in computer science.

Lastly, it should be noted that there were differences in both achievement and interest for women and men in CS1. The achievement difference in terms of grade point average was not significant, but men in this sample averaged higher grades than women did. Post-CS1 interest was high for both groups, but men showed significantly greater interest in continuing in CS. One measure of prior experience collected in this study was the effort to learn at least one programming language before college. Across gender a moderate or very high experience with learning a programming language led to significant differences in achievement in the CS1 course. For women, the average CS1 course performance rose by several tenths of a GPA point with each higher level of experience with computer programming. The model at the center of the study showed some evidence that prior knowledge of computer programming has a higher influence on course performance for women than men, but further work should explore self-assessments of computer programming knowledge to better understand the differences between the quality of CS experiences for men and women before entering the CS1 course.

WORKS CITED

WORKS CITED

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. https://doi.org/10.1037/0033-2909.103.3.411

Ashcraft, C., McLain, B., & Eger, E. (2016). *Women in tech: The facts*. National Center for Women & Information Technology. https://wpassets.ncwit.org/wp-content/uploads/2021/05/13193304/ncwit_women-in-it_2016-full-report_final-web06012016.pdf

Association of Computing Machinery. (2013). *Computer science curricula 2013*. https://www.acm.org/binaries/content/assets/education/cs2013_web_final.pdf

Astin, H. S., & Sax, L. J. (1996). Undergraduate women in science: Personal and environmental influences on the development of scientific talent. In C. Davis, A.B. Ginorio, C.S. Hollenshead, B.B. Lazarus, P.M. Rayman (Eds.), *The equity equation: Fostering the advancement of women in the sciences, mathematics, and engineering*, (pp. 96-121). Wiley.

Ball, C., Huang, K. T., Rikard, R. V., & Cotten, S. R. (2019). The emotional costs of computers: An expectancy-value theory analysis of predominantly low-socioeconomic status minority students' STEM attitudes. *Information, Communication & Society, 22*(1), 105-128. https://doi.org/10.1080/1369118X.2017.1355403

Bandura, A. (1986). *Social foundations of thought and action: Asocial cognitive theory.* Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W.H. Freeman & Company.

Baker, D., Krause, S., Yaşar, Ş., Roberts, C., & Robinson-Kurpius, S. (2007). An intervention to address gender issues in a course on design, engineering, and technology for science educators. *Journal of Engineering Education*, *96*(3), 213-226. https://doi.org/10.1002/j.2168-9830.2007.tb00931.x

Barker, L. J., McDowell, C., & Kalahar, K. (2009). Exploring factors that influence computer science introductory course students to persist in the major. *ACM SIGCSE Bulletin,41*(1), 153-157. https://doi.org/10.1145/1539024.1508923

Battle, A., & Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *Journal of Vocational Behavior*, *62*(1), 56–75. https://doi.org/10.1016/S0001-8791(02)00037-4

Bejar, D. (2004). *Your Blues.* Talitres Records.

Bennedsen, J., & Caspersen, M.E. (2007). Failure rates in introductory programming. *ACM SIGCSE Bulletin, 39*(2), 32-36. https://doi.org/10.1145/1272848.1272879

Bergin, S., & Reilly, R. (2005). Programming: Factors that influence success. *ACM SIGCSE Bulletin, 37*(1), 411-415. https://doi.org/10.1145/1047124.1047480

Bernstein, D. (1991). Comfort and experience with computing: Are they the same for women and men? *ACM SIGCSE Bulletin, 23*(3), 57-60. https://doi.org/10.1145/126459.126472

Beyer, S. (2014). Why are women underrepresented in computer science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education, 24*(2-3), 153-192. https://doi.org/10.1080/08993408.2014.963363

Beyer, S., Rynes, K., & Haller, S. (2004). Deterrents to women taking computer science courses. *IEEE Technology and Society Magazine, 23*(1), 21-28. https://doi.org/10.1109/MTAS.2004.1273468

Beyer, S., Rynes, K., Perrault, J., Hay, K., & Haller, S. (2003). Gender differences in computer science students. *ACM SIGCSE Bulletin, 35*(1), 49-53. https://doi.org/10.1145/792548.611930

Biggers, M., Brauer, A., & Yilmaz, T. (2008). Student perceptions of computer science: A retention study comparing graduating seniors with CS leavers. *ACM SIGCSE Bulletin, 40*(1), 402-406. https://doi.org/10.1145/1352322.1352274

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology, 26*(4), 553-570. https://doi.org/10.1006/ceps.2000.1048

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1–40. https://doi.org/10.1023/A:1021302408382

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching, 43*(5), 485-499. https://doi.org/10.1002/tea.20131

Byrne, P., & Lyons, G. (2001). The effect of student attributes on success in programming. *ACM SIGCSE Bulletin, 33*(3), 49-52. https://doi.org/10.1145/507758.377467

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching, 44*(8), 1187-1218. https://doi.org/10.1002/tea.20237

Carver, S., & Risinger, S. (1987). Improving children's debugging skills. In G. Olson, S. Sheppard & E. Soloway (Eds.), *Empirical Studies of Programmers: Second Workshop* (pp. 147–171). Ablex.

Cheryan, S., & Plaut, V. C. (2010). Explaining underrepresentation: A theory of precluded interest. *Sex Roles*, *63*, 475-488. https://doi.org/10.1007/s11199-010-9835-x

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science, 2*(6), 656-664. https://doi.org/10.1177/1948550611405218

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*(1), 1-35. https://doi.org/10.1037/bul0000052

Clegg, S., & Trayhurt, D. (2000). Gender and computing: Not the same old problem. *British Educational Research Journal, 26*(1), 75–89. https://doi.org/10.1080/014119200109525

Code.org, CSTA, & ECEP Alliance. (2020). *2020 State of computer science education: Illuminating disparities.* https://advocacy.code.org/2020_state_of_cs.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Erlbaum.

Cohoon, J. M. (2001). Toward improving female retention in the computer science major. *Communications of the ACM*, *44*(5), 108-114. https://doi.org/10.1145/374308.374367

Cohoon, J. M., & Aspray, W. (2006). *Women and information technology: Research on underrepresentation*. The MIT Press.

College Board (2020). *AP Program Participation and Performance Data 2020* https://research.collegeboard.org/programs/ap/data/participation/ap-2020

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry*, *22*(4), 231-246. https://doi.org/10.1080/1047840X.2011.607313

Denner, J., Werner, L., O'Connor, L., & Glassman, J. (2014). Community college men and women: A test of three widely held beliefs about who pursues computer science. *Community College Review, 42*(4), 342–362. https://doi.org/10.1177/0091552114535624

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(3), 327-346. https://doi.org/10.1207/S15328007SEM0903_2

Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. Elliott & C. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 105-121). Guilford Publications.

Eccles J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and Achievement Motivation* (pp. 75–146). W. H. Freeman.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46*(2), 183-201. https://doi.org/10.1111/j.1540-4560.1990.tb01929.x

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*(4), 519-533. https://doi.org/10.1037/0012-1649.38.4.519

Frieze, C., & Quesenberry, J. (2015). *Kicking butt in computer science: Women in computing at Carnegie Mellon University.* Dog Ear Publishing.

Frieze, C., Quesenberry, J. L., Kemp, E., & Velázquez, A. (2012). Diversity or difference? New research supports the case for a cultural perspective on women in computing. *Journal of Science Education and Technology, 21*, 423–439. https://doi.org/10.1007/s10956-011-9335-y

Gaspard, H., Dicke, A. L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663-677. https://doi.org/https://doi.org/10.1037/edu0000003

Gaspard, H., Wille, E., Wormington, S. V., & Hulleman, C. S. (2019). How are upper secondary school students' expectancy-value profiles associated with achievement and university STEM major? A cross-domain comparison. *Contemporary Educational Psychology, 58*, 149-162. https://doi.org/10.1016/j.cedpsych.2019.02.005

Glück, J., & Fitting, S. (2003). Spatial strategy selection: Interesting incremental information. *International Journal of Testing, 3*(3), 293–308. https://doi.org/10.1207/S15327574IJT0303_7

Google Inc. & Gallup Inc. (2016). Diversity gaps in computer science: Exploring the underrepresentation of girls, Blacks and Hispanics. http://goo.gl/PG34aH

Goold, A., & Rimmer, R. (2000). Factors affecting performance in first-year computing. *ACM SIGCSE Bulletin, 32*(2), 39-43. https://doi.org/10.1145/355354.355369

Gorson, J., & O'Rourke, E. (2020). Why do CS1 students think they're bad at programming? Investigating self-efficacy and self-assessments at three universities. *ICER '20:*

Proceedings of the 2020 ACM Conference on International Computing Education Research (pp. 170-181). https://doi.org/10.1145/3372782.3406273

Hagan, D., & Markham, S. (2000). Does it help to have some programming experience before beginning a computing degree program? *ACM SIGCSE Bulletin*, *32*(3), 25-28. https://doi.org/10.1145/353519.343063

Hoelter, J. W. (1983). The effects of role evaluation and commitment on identity salience. *Social Psychology Quarterly, 46*(2), 140-147. https://doi.org/10.2307/3033850

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huang X., Hong W., Bernacki M. (2018) The psychological cost of college math: Digital learning behaviors, outcomes, and genders differences. In C. Stephanidis (Ed.) Communications in Computer and Information Science: Vol 852 (pp. 43-50). Springer. https://doi.org/10.1007/978-3-319-92285-0_7

Hutchison, M. A., Follman, D. K., Sumpter, M., & Bodner, G. M. (2006). Factors influencing the self-efficacy beliefs of first-year engineering students. *Journal of Engineering Education, 95*(1), 39-47. https://doi.org/10.1002/j.2168-9830.2006.tb00876.x

Inzlicht, M., & Good, C. (2006). How environments can threaten academic performance, self-knowledge, and sense of belonging. In S. Levin & C. Van Laar (Eds.), *Stigma and Group Inequality* (pp. 143-164). Psychology Press.

Jones, S., & Burnett, G. (2008). Spatial ability and learning to program. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, 4*(1), 47-61. http://urn.fi/URN:NBNfi:jyu-200804151352

Kapoor, A., & Gardner-McCune, C. (2018). Considerations for switching: Exploring factors behind CS students' desire to leave a CS major. *ITiCSE 2018: Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, 290–295. https://doi.org/10.1145/3197091.3197113

Keller, J. (2007). Stereotype threat in classroom settings: The interaction effect of domain identification, task difficulty and stereotype threat on female students' math performance. *British Journal of Educational Psychology*, *77*(2), 323–338. https://doi.org/10.1348/000709906X113662

Kessler, C., & Anderson, J. (1986). A model of novice debugging in LISP. In E. Soloway & S. Iyengar (Eds.), *Empirical Studies of Programmers* (pp. 198–212). Ablex.

Kinnunen, P., & Malmi, L. (2006, September). Why students drop out CS1 course? *ICER '06: Proceedings of the Second International Workshop on Computing Education Research,* 97-108. https://doi.org/10.1145/1151588.1151604

Kinnunen, P., & Simon, B. (2011). CS majors' self-efficacy perceptions in CS1: Results in light of social cognitive theory. *ICER '11: Proceedings of the Seventh International Workshop on Computing Education Research*, 19-26. https://doi.org/10.1145/2016911.2016917

Kline, R.B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.

Kruteskii, V. A. (1976). *The psychology of mathematics ability in school children.* The University of Chicago Press.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*, 151-160. https://doi.org/10.1007/BF02288391

Lawton, C. (2010). Gender, spatial ability, and wayfinding. In J.C. Chrisler & D.R. McCreary (Eds.), *Handbook of Gender Research in Psychology* (pp. 317-342). Springer.

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., & Werner, L. (2011). Computational thinking for youth in practice. *ACM Inroads*, *2*(1), 32-37. https://doi.org/10.1145/1929887.1929902

Leedy, M. G., LaLonde, D., & Runk, K. (2003). Gender equity in mathematics: Beliefs of students, parents, and teachers. *School Science and Mathematics*, *103*(6), 285-292. https://doi.org/10.1111/j.1949-8594.2003.tb18151.x

Lehman, K. J. (2019). An untapped recruitment pool: Undecided students in CS1 courses. *2019 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT),* 1-8. https://doi.org/10.1109/RESPECT46404.2019.8985882

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1), 79-122. https://doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. In D. Brown (Ed.), *Career Choice and Development* (pp. 255-311). Wiley.

Lewis, C., Bruno, P., Raygoza, J., & Wang, J. (2019). Alignment of goals and perceptions of computing predicts students' sense of belonging in computing. *ICER '19: Proceedings of the 2019 ACM Conference on International Computing Education Research,*11-19. https://doi.org/10.1145/3291279.3339426

Lewis, K. L., Stout, J. G., Finkelstein, N. D., Pollock, S. J., Miyake, A., Cohen, G. L., & Ito, T. A. (2017). Fitting in to move forward: Belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychology of Women Quarterly, 41*(4), 420-436. https://doi.org/10.1177/0361684317720186

Lin, G. Y. (2015). Self-efficacy beliefs and their sources in undergraduate computing disciplines: An examination of gender and persistence. *Journal of Educational Computing Research, 53*(4), 540–561. https://doi.org/10.1177/0735633115608440

Linn, M. C., & Hyde, J. S. (1989). Gender, mathematics, and science. *Educational Researcher, 18*(8), 17-27. https://doi.org/10.3102/0013189X018008017

Lishinski, A., Yadav, A., Good, J., & Enbody, R. (2016). Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. *ICER '16: Proceedings of the 2016 ACM Conference on International Computing Education Research,* 211-220. https://doi.org/10.1145/2960310.2960329

Lopez, F. G., Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Role of social–cognitive expectations in high school students' mathematics-related interest and performance. *Journal of Counseling Psychology, 44*(1), 44–52. https://doi.org/10.1037/0022-0167.44.1.44

Margolis, J., & Fisher, A. (2003). *Unlocking the clubhouse: Women in computing*. MIT press.

Marra, R. M., Rodgers, K. A., Shen, D., & Bogue, B. (2009). Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy. *Journal of Engineering Education, 98*(1), 27-38. https://doi.org/10.1002/j.2168-9830.2009.tb01003.x

Master, A., Cheryan, S., & Meltzoff, A. N. (2016). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology, 108*(3), 424-437. https://doi.org/10.1037/edu0000061

McCauley, R., Fitzgerald, S., Lewandowski, G., Murphy, L., Simon, B., Thomas, L., & Zander, C. (2008). Debugging: a review of the literature from an educational perspective. *Computer Science Education, 18*(2), 67-92. https://doi.org/10.1080/08993400802114581

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*(1), 60-70. https://doi.org/10.1037/0022-0663.82.1.60

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. University of Michigan.

Miliszewska, I., Barker, G., Henderson, F., & Sztendur, E. (2006). The issue of gender equity in computer science – what students say. *Journal of Information Technology Education: Research, 5*(1), 107-120. https://www.learntechlib.org/p/111535/

Moakler Jr, M. W., & Kim, M. M. (2014). College major choice in STEM: Revisiting confidence and demographic factors. *The Career Development Quarterly, 62*(2), 128-142. https://doi.org/10.1002/j.2161-0045.2014.00075.x

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(4), 599-620. https://doi-org.proxy1.cl.msu.edu/10.1207/S15328007SEM0904_8

National Center for Women and Information Technology (2019). *By the Numbers*. https://www.ncwit.org/resources/numbers

National Center for Science and Engineering Statistics. (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019.* National Science Foundation. https://www.nsf.gov/statistics/wmpd/

Nauta, M. M., & Kokaly, M. L. (2001). Assessing role model influences on students' academic and vocational decisions. *Journal of Career Assessment, 9*(1), 81-99. https://doi.org/10.1177/106907270100900106

Order, N. & Division, J. (2011). *Total: From Joy Division to New Order.* Rhino Records.

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr, & P. R. Pintrich (Eds.), *Advances in Motivation and Achievement* (pp. 1–49). JAI Press.

Pajares, F. (2006). Self-efficacy during childhood and adolescence. In F. Pajares & T.C. Urdan (Eds.), *Self-efficacy Beliefs of Adolescents* (pp. 339-367). Information Age Publishing.

Palmer, D. H. (2006). Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education, 36*, 337-353. https://doi.org/10.1007/s11165-005-9007-0

Pantic, K., & Clarke-Midura, J. (2019). Factors that influence retention of women in the computer science major: A systematic literature review. *Journal of Women and Minorities in Science and Engineering, 25*(2), 119-145. https://doi.org/10.1615/JWomenMinorScienEng.2019024384

Pappas, I. O., Giannakos, M. N., & Jaccheri, L. (2016). Investigating factors influencing students' intention to dropout computer science studies. *ITiCSE '16: Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education,* 198-203. https://doi.org/10.1145/2899415.2899455

Parker, M. C., Guzdial, M., & Engleman, S. (2016). Replication, validation, and use of a language independent CS1 knowledge assessment. *ICER '16: Proceedings of the 2016 ACM Conference on International Computing Education Research,* 93-101. https://doi.org/10.1145/2960310.2960316

Parker, M. C., Solomon, A., Pritchett, B., Illingworth, D. A., Marguilieux, L. E., & Guzdial, M. (2018). Socioeconomic status and computer science achievement: Spatial ability as a mediating variable in a novel model of understanding. *ICER '18: Proceedings of the 2018 ACM Conference on International Computing Education Research*, 97-105. https://doi.org/10.1145/3230977.3230987

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology, 106*(1), 315–329. https://doi.org/10.1037/a0034027

Perez, T., Wormington, S. V., Barger, M. M., Schwartz-Bloom, R. D., Lee, Y., & Linnenbrink-Garcia, L. (2019). Science expectancy, value, and cost profiles and their proximal and distal relations to undergraduate science, technology, engineering, and math persistence. *Science Education, 103*(2), 264–286. https://doi.org/10.1002/sce.21490

Petersen, A., Craig, M., Campbell, J., & Tafliovich, A. (2016). Revisiting why students drop CS1. *ICER '16: Proceedings of the 16th International Conference on Computing Education Research,* 71-80. https://doi.org/10.1145/2999541.2999552

Porter, L., & Simon, B. (2013). Retaining nearly one-third more majors with a trio of instructional best practices in CS1. *SIGCSE '13: Proceeding of the 44th ACM Technical Symposium on Computer Science Education,* 165-170. https://doi.org/10.1145/2445196.2445248

Priess, H.A., & Hyde, J.S. (2010). Gender and academic abilities and preferences. In J.C. Chrisler & D.R. McCreary (Eds.), *Handbook of Gender Research in Psychology* (pp. 297-316). Springer.

Pryor, J.H., Hurtado, S., DeAngelo, L. Palucki Blake, L. & Tran, S. (2009). *The American freshman: National norms fall 2009.* Higher Education Research Institute, UCLA.

Quille, K., & Bergin, S. (2019). CS1: How will they do? How can we help? A decade of research and practice. *Computer Science Education, 29*(2-3), 254-282. https://doi.org/10.1080/08993408.2019.1612679

Ramalingam, V., LaBelle, D., & Wiedenbeck, S. (2004). Self-efficacy and mental models in learning to program. *ITiCSE '04: Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, 171-175. https://doi.org/https://doi.org/10.1145/1007996.1008042

Rask, K. N., & Bailey, E. M. (2002). Are faculty role models? Evidence from major choice in an undergraduate institution. *The Journal of Economic Education, 33*(2), 99-124. https://doi.org/10.1080/00220480209596461

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. https://doi.org/10.1037/a0029315

Robert, M., & Héroux, G. (2004). Visuo-spatial play experience: forerunner of visuo-spatial achievement in preadolescent and adolescent boys and girls? *Infant and Child Development, 13*(1), 49-78. https://doi.org/10.1002/icd.336

Roberts, M. R. H., McGill, T. J., & Hyland, P. N. (2012). Attrition from Australian ICT degrees - Why women leave. ACE '12: Proceedings of the 14th Australasian Computing Education Conference, 15–24. https://ro.uow.edu.au/infopapers/1870/

Rosson, M. B., Carroll, J. M., & Sinha, H. (2011). Orientation of undergraduates toward careers in the computer and information sciences: Gender, self-efficacy and social support. ACM Transactions on Computing Education, 11(3), 1–23. https://doi.org/10.1145/2037276.2037278

Rubio, C., Osca, A., Recio, P., Urien, B., & Peiró, J. M. (2015). Work-family conflict, self-efficacy, and emotional exhaustion: A test of longitudinal effects. *Revista de Psicología del Trabajo y de las Organizaciones, 31*(3), 147-154. https://doi.org/10.1016/j.rpto.2015.06.004

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education, 96*(3), 411-427. https://doi.org/10.1002/sce.21007

Sands, P., & Capobianco, B. (2020). *The impact of mentoring on computing identity for women in computer science.* Unpublished manuscript.

Sankar, P., Gilmartin, J., & Sobel, M. (2015). An examination of belongingness and confidence among female computer science students. *ACM SIGCAS Computers and Society,45*(2), 7-10. https://doi.org/10.1145/2809957.2809960

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50,* 83-90. https://doi.org/10.1007/BF02294150

Sawtelle, V., Brewe, E., Goertzen, R. M., & Kramer, L. H. (2012). Identifying events that impact self-efficacy in physics learning. *Physics Education Research, 8*(2), 1-18. https://doi.org/10.1103/PhysRevSTPER.8.020111

Sawtelle, V., Brewe, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching, 49*(9), 1096-1121. https://doi.org/10.1002/tea.21050

Sax, L. J., Blaney, J. M., Lehman, K. J., Rodriguez, S. L., George, K. L., & Zavala, C. (2018). Sense of belonging in computing: The role of introductory courses for women and underrepresented minority students. *Social Sciences, 7*(8), 122-145. https://doi.org/10.3390/socsci7080122

Sax, L. J., Lehman, K. J., Jacobs, J. A., Kanny, M. A., Lim, G., Monje-Paulson, L., & Zimmerman, H. B. (2017). Anatomy of an enduring gender gap: The evolution of women's participation in computer science. *The Journal of Higher Education*, *88*(2), 258–293. https://doi.org/10.1080/00221546.2016.1257306

Schunk, D. H., & Lilly, M. W. (1984). Sex differences in self-efficacy and attributions: Influence of performance feedback. *The Journal of Early Adolescence, 4*(3), 203-213. https://doi.org/10.1177/0272431684043004

Schunk, D.H., Meece, J.L., Pintrich, P.R. (2014). *Motivation in education: theory, research, and applications.* Pearson.

Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. In A. Elliot & C.S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 85-104). Guilford Press.

Sheard, J., Carbone, A., Markham, S., Hurst, A. J., Casey, D., & Avram, C. (2008). Performance and progression of first year ICT students. *ACE '08: Proceedings of the tenth conference on Australasian computing education, 78*, 119-127. https://doi.org/10.5555/1379249.1379261

Simon, B., & Cutts, Q. (2012). Peer instruction: A teaching method to foster deep understanding. *Communications of the ACM*, *55*(2), 27-29. https://doi.org/10.1145/2076450.2076459

Singh, K., Allen, K. R., Scheckler, R., & Darlington, L. (2007). Women in computer-related majors: A critical synthesis of research and theory from 1994 to 2005. *Review of Educational Research, 77*(4), 500–533. https://doi.org/10.3102/0034654307309919

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Personality and Social Psychology Bulletin, 39*(2), 131-143. https://doi.org/10.1177/0146167212468332

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. https://doi.org/10.1006/jesp.1998.1373

Stolzenberg, E. B., Eagan, M. K., Aragon, M. C., Cesar-Davis, N. M., Jacobo, S., Couch, V., & Rios-Aguilar, C. (2019). *The American freshman: National norms fall 2017*. Higher Education Research Institute, UCLA.

Tafliovich, A., Campbell, J., & Petersen, A. (2013). A student perspective on prior experience in CS1. *SIGCSE '13: Proceedings of the 44th ACM Technical Symposium on Computer Science Education*, 239-244. https://doi.org/10.1145/2445196.2445270

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED majors. *Sex Roles*, *77*, 86-96. https://doi.org/10.1007/s11199-016-0694-y

Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education, 18*(2), 345-368. https://www.learntechlib.org/primary/p/32311/

Usher, E. L. (2009). Sources of middle school students' self-efficacy in mathematics: A qualitative investigation. *American Educational Research Journal, 46*(1), 275-314. https://doi.org/10.3102/0002831208324517

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751-796. https://doi.org/10.3102/0034654308321456

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology, 34*(1), 89-101. https://doi.org/10.1016/j.cedpsych.2008.09.002

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111-133. https://doi.org/10.1207/s15326985ep3902_3

Vile, K. (2013). *Wakin on a pretty daze.* Matador Records.

Wang, J., Hong, H., Ravitz, J., & Ivory, M. (2015). Gender differences in factors influencing pursuit of computer science and related fields. *ITiCSE '15: Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, 117-122. https://doi.org/10.1145/2729094.2742611

Watson, C., Li., F.W. (2014). Failure rates in introductory programming revisited. *ITiCSE'14: Proceedings of the 2014 ACM Conference on Innovation and Technology in Computer Science Education*, 39-44. https://doi.org/10.1145/2591708.2591749

Watson, C., Li, F. W., & Godwin, J. L. (2014). No tests required: comparing traditional and dynamic predictors of programming success. *SIGCSE '14: Proceedings of the 45th ACM*

Technical Symposium on Computer Science Education, 469-474. https://doi.org/10.1145/2538862.2538930

Weinberger, C. J. (2004). Just ask! Why surveyed women did not pursue IT courses or careers. *IEEE Technology and Society Magazine, 23*(2), 28-35. https://doi.org/10.1109/MTAS.2004.1304399

Weisgram, E. S., & Bigler, R. S. (2006). The role of attitudes and intervention in high school girls' interest in computer science. *Journal of Women and Minorities in Science and Engineering, 12*(4), 325-336. https://doi.org/10.1615/JWomenMinorScienEng.v12.i4.40

Wiedenbeck, S., LaBelle, D., & Kain, V. N. R. (2004). Factors affecting course outcomes in introductory programming. In E. Dunican, & T.R.G. Green (Eds.), *Proceedings of the 16th annual Psychology of Programming Interest Group* (pp. 97-110). PPIG. https://www.ppig.org/papers/2004-ppig-16th-wiedenbeck/

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review,12*(3), 265-310. https://doi.org/10.1016/0273-2297(92)90011-P

Wilcox, C., & Lionelle, A. (2018). Quantifying the benefits of prior programming experience in an introductory computer science course. *SIGCSE '18: Proceedings of the 49th ACM Technical Symposium on Computer Science Education,* 80-85. https://doi.org/10.1145/3159450.3159480

Williams, M. M., & George-Jackson, C. (2014). Using and doing science: Gender, self-efficacy, and science identity of undergraduate students in STEM. *Journal of Women and Minorities in Science and Engineering, 20*(2), 99-126. https://doi.org/10.1615/JWomenMinorScienEng.2014004477

Wilson, B. C. (2002). A study of factors promoting success in computer science including gender differences. *Computer Science Education, 12*(1-2), 141-164. https://doi.org/10.1076/csed.12.1.141.8211

Wilson, B. C., & Shrock, S. (2001). Contributing to success in an introductory computer science course: A study of twelve factors. *ACM SIGCSE Bulletin,33*(1), 184-188. https://doi.org/10.1145/366413.364581

Xie, B., Davidson, M. J., Li, M., & Ko, A. J. (2019). An item response theory evaluation of a language-independent CS1 knowledge assessment. *SIGCSE '19: Proceedings of the 50th ACM Technical Symposium on Computer Science Education,* 699-705. https://doi.org/10.1145/3287324.3287370

Ying, K. M., Rodríguez, F. J., Dibble, A. L., Martin, A. C., Boyer, K. E., Thomas, S. V., & Gilbert, J. E. (2021). Confidence, connection, and comfort: Reports from an all-women's

CS1 class. *SIGCSE '21: Proceedings of the 52nd ACM Technical Symposium on Computer Science Education,* 699-705. https://doi.org/10.1145/3408877.3432548

Yuan, K. H., Bentler, P. M., & Zhang, W. (2005). The effect of skewness and kurtosis on mean and covariance structure analysis: The univariate case and its multivariate implication. *Sociological Methods & Research, 34*(2), 240-258. https://doi.org/10.1177/0049124105280200

Yuan, K. H., Chan, W., & Bentler, P. M. (2000). Robust transformation with applications to structural equation modelling. *British Journal of Mathematical and Statistical Psychology, 53*(1), 31-50. https://doi.org/10.1348/000711000159169

Zarrett, N. R., & Malanchuk, O. (2005). Who's computing? Gender and race differences in young adults' decisions to pursue an information technology career. *New Directions for Child and Adolescent Development, 2005*(110), 65-84. https://doi.org/10.1002/cd.150

Zeldin, A. L., Britner, S. L., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 45*(9), 1036-1058. https://doi.org/10.1002/tea.20195

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37*(1), 215–246. https://doi.org/10.3102/00028312037001215

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82-91. https://doi.org/10.1006/ceps.1999.1016

Zingaro, D. (2015). Examining interest and grades in Computer Science 1: A study of pedagogy and achievement goals. *ACM Transactions on Computing Education, 15*(3), 1-18. https://doi.org/10.1145/2802752