Title

Establishing an Empirical Foundation for a Theory of Student Learning and Success in CS1

Author

Salguero, Adrian

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Establishing an Empirical Foundation for a Theory of Student Learning and Success in CS1

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Computer Science

by

Adrian Salguero

Committee in charge:

Professor William G. Griswold, Co-Chair
Professor Leo Porter, Co-Chair
Professor Gail Heyman
Professor Beth Simon

2024

The Dissertation of Adrian Salguero is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

I want to thank my mother and father, Amelia and Pascual, for their endless love and support for me throughout my life. I can never repay all the sacrifices you've made for me and your strong work ethic inspires me everyday.

I want to thank my brothers Salvador and Eric, my sister Marlene, my sister-in-law Julia, and my nephew and niece, Sebastian and Juliana. Your unwavering support for me throughout my life have given me the drive to push ever higher. Whenever I felt down or unsure I could always count on you all to be there for me.

Finally, I want to thank the amazing friends that have supported me through my journey at UCSD. Naming you all would extend this section far too long. Through you all I learned and explored more about who I was as a person. Whether it was to go watch a baseball game, grab food together, or simply lounge around you provided me with laughs, love, and wonderful memories that I will forever cherish.

I would not have made it so far without the support and love of you all. Whether through blood or not, you are family and I thank you all for being a part of my life and this journey.

Love,

Adrian

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I want to thank my advisors, Leo Porter and Bill Griswold for the opportunity to work with them across many years and projects and for nurturing my love and passion for computer science education research. I also want to thank Beth Simon, Christine Alvarado, Gerald Soosai Raj, Sophia Krause-Levy, and Sander Valster. I have learned so much from you all and thank you for being a part of my graduate school journey.

I would also like to thank John Graham Lewis, Eileen Kodama, Elaine Blomeyer, and Nancy Cox-Konopelski for mentoring and guiding me through high school and undergrad. I would not be where I am today without you motivation and mentorship.

Most importantly, I want to thank my parents, friends, and the rest of my family who have have always supported me.

Chapter 1, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold,

Christine Alvarado, and Leo Porter. "Understanding Sources of Student Struggle in Early Computer Science Courses", 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. "Applying CS0/CS1 Student Success Factors and Outcomes to Biggs' 3P Educational Model". The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. "A Longitudinal Evaluation of a Best Practices CS1", 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. "Understanding Sources of Student Struggle in Early Computer Science Courses", 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. "Applying CS0/CS1 Student Success Factors and Outcomes to Biggs' 3P Educational Model". The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in the Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. "Applying CS0/CS1 Student Success Factors and Outcomes to Biggs' 3P Educational Model". The dissertation author is the primary investigator and author of this paper.

Chapter 5, in part, is a reprint of the material as it appears in the Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. "Understanding Sources of Student Struggle in Early Computer Science Courses". The dissertation author is the primary investigator and author of this paper.

Chapter 6, in part, is a reprint of the material as it appears in the Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”. The dissertation author is the primary investigator and author of this paper.

Chapter 7, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Chapter 8, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

VITA

2018	Bachelor of Science in Computer Science, University of California Santa Cruz
2021	Master of Science in Computer Science, University of California San Diego
2024	Doctor of Philosohpy in Computer Science, University of California San Diego

PUBLICATIONS

Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter, "Applying CS0/CS1 Student Success Factors and Outcomes to Biggs' 3P Educational Model", *In Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024)*, 1168-1174, 2024.

Adrian Salguero, "Establishing and Empirical Foundation for a Theory of Student Learning and Success in CS1", *In Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 2 (ICER 2023)*, 52-54, 2023.

Ismael Villegas Molina, Adrian Salguero, Shera Zhong, and Adalbert Gerald Soosai Raj, "The Effects of Spanish-English Bilingual Instruction in a CS0 Course for High School Students", *In Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V.1 (ITiCSE 2023)*, 75-81, 2023.

Sophia Krause-Levy, Adrian Salguero, Rachel S. Lim, Hayden McTavish, Jelena Trajkovic, Leo Porter, and William G. Griswold, "Instructor Perspectives on Prerequisite Courses in Computing" *In Proceedings of the 54th ACM Technical Symposium on Computer Science Education (SIGCSE 2023)*, 277–283, 2023.

Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter, "Understanding Sources of Student Struggle in Early Computer Science Courses", *In Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021)*, 319-333, 2021.

Sander Valstar, Sophia Krause-Levy, Adrian Salguero, Leo Porter, and William G. Griswold, "Proficiency in Basic Data Structures among Various Subpopulations of Students at Different Stages in a CS Program", *In Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2021)*, 429–435, 2021.

McKenna Lewis, Zhanchong Deng, Sophia Krause-Levy, Adrian Salguero, William G. Griswold, Leo Porter, and Christine Alvarado, "Exploring Student Experiences in Early Computing Courses during Emergency Remote Teaching" *In Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2021)*, 88–94, 2021.

Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter, “A Longitudinal Evaluation of a Best Practices CS1” *In Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER 2020)*, 182-193, 2020.

ABSTRACT OF THE DISSERTATION

Establishing an Empirical Foundation for a Theory of Student Learning and Success in CS1

by

Adrian Salguero

Doctor of Philosophy in Computer Science

University of California San Diego, 2024

Professor William G. Griswold, Co-Chair
Professor Leo Porter, Co-Chair

Introductory computer programming (i.e. CS1) is the entry point into the computer science major at higher education institutions worldwide. It introduces foundational concepts to students that are then built upon in future courses. Computer science as a whole has struggled to attract and retain students in the major, particularly women and students from minority backgrounds. Researchers have explored reasons as to why students are struggling to complete their CS1 course and remain in the major. Current research has uncovered a wide range of different factors which have been studied mostly in isolation from each other, thus not being able to provide a full view of the CS1 student experience. This dissertation presents an (1)

exploration of the current state of CS1 research and its organization into an existing theoretical framework, (2) a **holistic** view of student experiences in early computing courses, and (3) the benefits and role of **pedagogical** best practices in supporting student learning in CS1.

Regarding **exploration**: we surveyed the current state of CS1 literature to gain an understanding of what factors related to student CS1 success have been studied and mapped the findings to an existing student learning framework: Biggs' 3P model. We find that the factors explored included areas such as internal student factors, instructional practices, the student's learning process, and many more.

Regarding **holistic**: we collected student experience surveys from students of the first four courses students at UCSD take during their computer science program. We found lower performing students report higher levels of stress across multiple areas compared to their higher performing peers.

Regarding **pedagogy**: we analyzed 12 years of longitudinal student data to understand how the implementation of a trio of best practices (peer instruction, pair programming, and media computation) in introductory programming courses at UCSD demonstrated improvement to student performance and retention within the program. Regression models and significance testing provide evidence that the presence of these best practices were significant in these positive outcomes. Despite these improvements, certain groups such as women and minorities were still failing and being retained at lower rates in the major than their peers.

Chapter 1

Introduction

1.1 Motivation

Historically, computer science has a history of high failure rates, low retention rates of students interested in computing [13, 128], and students not learning what instructors expect [81]. As such, researchers have sought to understand why some students struggle in computing courses by examining a myriad of factors that might be connected to student success [11, 83, 157, 161]. In the United States, the context of our work, there also exist substantial divides by gender and race among those who major in computer science [38] and as well as those who have access to computing education before college [86, 87]. In addition, Black, LatinX, and Native American and women students who pursue computing in university are retained at lower rates than White and Asian men [172].

The rise of the dot-coms at the end of the 1990s correlated with a surge in interest in computer science and related fields, with the number of bachelor's degrees in computing peaking in the United States in 2003 [99]. Following that peak, the number of students interested in computer science suffered a precipitous decline. To respond to that decline, computer science programs began to focus on how best to attract and retain students into the computer science discipline. Within this context, our institution began a concerted effort to overhaul its CS1 course for students without prior experience in order to reduce failure rates, improve retention of students in computing majors, and to attract new students to the major [112].

According to the United States Bureau of Labor Statistics, the projected growth for computing-related job occupations are expected to grow at a faster average rate between 2022 and 2032, with about 377,500 openings projected each year [105]. Furthermore, the median average wage for these types of jobs are higher than the median annual wage for all occupations [105]. The combination of high expected job growth and financial opportunity has made computer science a highly sought after degree. In 2021-2022 academic year, the National Center for Education Statistics reported 108,503 bachelors degrees in computer science were awarded, an 8 percent increase from the prior year and the most ever awarded in a year [35]. However, women only made up about 23 percent of the awarded majors and BLNPI students made up about 21 percent of the awarded majors [35, 36].

For many students, introductory computer science is the course that introduces them to the world of computing. For many others, it is notoriously the course that deters them from considering computer science as an academic or professional pursuit. Researchers have examined a myriad of factors that can influence student success in computer science and retention in high institution computer science major programs. Prior studies in computer science education have investigated the impact of a variety of factors on students' outcomes. Some of these factors include prior experience [161], sense of belonging [109, 102, 131, 71, 155, 89, 90, 94], interest in the material [22], self efficacy [79], study skills [74], and alignment between student goals and the perceived goals of the field [71].

CS1 is important as it teaches the fundamentals of programming and computer science that will be built upon in future courses, thus making it so popular among researchers. In 2010, Robins defined a hypothesis as to why students struggle in their CS1 courses called, "Learning Edge Momentum (LEM)" [122]. This hypothesis claims that students who understand a concept will be more likely to understand later concepts that build upon the earlier concept. Over time, this learning will compound and students will begin to understand material dependent on how well they understand foundational knowledge. If there exists a misunderstanding or the student simply does not grasp the foundational concept, it is more likely they will struggle with future concepts. This hypothesis highlights the importance of students obtaining a strong CS1 foundational knowledge which are impacted by the factors mentioned above. Researchers and instructors must understand how these factors are impacting student learning to ensure strong early foundational knowledge in a students' computational education. It has therefore led researchers to begin uncovering which factors are tied to student success and struggle in these courses.

1.1.1 Understanding CS1 Success and Struggle

A popular area of CS1 research has focused on understanding what characteristics of learners, educators, and environments are associated with student success in these courses. Studies have discovered a variety of mutable and immutable characteristics and experiences that correlate with or predict student CS1 success. For example, one of the most consistent predictive factors of CS1 success is prior programming experience; many studies conclude that students entering with prior programming or computing experience tend to outperform their peers who enter with little to no experience [160, 49, 158, 161, 18]. Additional factors of success that have been identified through research span across cognitive and spatial skills [27, 78, 4, 16, 24, 56, 84, 88, 34, 107], behavioral [74, 42, 96], and mindset [97, 135, 119, 120, 65].

Within these broad areas, research has explored the effect of specific factors. For example, certain cognitive abilities, such as mental models and performance on specific cognitive tasks (e.g. spatial reasoning), have been associated with success in CS1 [16, 56, 84, 24, 34, 107, 88]. Other studies have shown how a student's sense of belonging, self-efficacy, and confidence is associated with their performance and enjoyment of CS1 [59, 77, 2, 62, 95]. Additional research has focused on how pedagogy can improve student learning in CS1 [128, 139, 111, 2]. Overall, success characteristics in CS1 cover a broad range of topics, and it is difficult to determine how these characteristics interact and whether some are more significant than others. For example, while sense of belonging and self-efficacy are both related to success, some research suggests that sense of belonging is an outcome of self-efficacy rather than a causal predictor of success [79, 77].

A nuanced understanding of how these factors interact allows researchers and educators to focus on factors that affect success (i.e., self-efficacy).

There have been some studies focusing on the interactions between different factors and how they influence CS1 student outcomes [126, 78, 141, 157]. A study conducted by Lishinski et al. looked at differences in student programming performance based on their gender, motivation, goals, and levels of self-efficacy [79]. They discovered goal orientation and metacognitive exercises impact student self-efficacy, which predicts students' CS1 outcomes. Another study by Rountree et al. used decision tree classifiers to predict success or failure in a CS1 course, concluding that groupings of factors such as age, prior background, and desired grades were predictive of success and failure in the course [126]. Early work conducted by Wilson and Shrock studied twelve factors that contributed towards success in a CS1 course [161]. Their findings indicated that factors such as comfort level, math background, and prior computing experience were positively associated with performance while attributing success and failure to luck was negatively associated with performance.

Many studies have also looked at the experiences of certain demographic groups in CS1. Research has identified that women and members of certain demographic groups, such as Black and LatinX students, significantly struggle to succeed and be retained in CS programs [128, 21, 57]. This work has led researchers to explore what reasons are causing CS attrition amongst these students. It has pointed to factors such as low self-efficacy, lack of prior programming experience, and low interest in the field as reasons for these negative outcomes, which echoes a recent conclusion that students who struggle typically struggle with multiple risk factors [123, 90, 33, 129]. The range of factors students appear to be impacted by has led us to search for a framework to begin organizing these research results.

1.2 Role of Theory in Computing Education

Despite the vast amount of potentially interacting and compounding factors impacting CS1 students, no theoretical framework has been developed within the computing education community to fully explain their experiences. Theory has and continues to play an important role in computing education research. Learning theories and frameworks, such as Bandura's theory of self-efficacy [6] and Dweck's mindset theory [14], are used across many studies.

The role of theory in this field is important as it provides guidance as to how students are learning concepts and what qualities are tied to their learning. However there are cases where theories developed in other fields and adopted in computing studies do not translate perfectly. For example, studies focusing on achievement goal theory demonstrated how only mastery goals were tied to success in CS1 and performance goals were not [171].

Of these factors, some have consistent and replicated outcomes, such as prior experience being positively correlated with higher performance or women generally entering CS0/CS1 with less prior experience than men [49, 160, 158, 18]. However, other factors are understudied or produce inconsistent results, limiting our knowledge of how to best support students. To our knowledge, there is no taxonomy that attempts to survey and organize the extensive amount of CS1 literature in a way that organizes factors into categories, highlights what has been heavily studied, explores the relationships between factors, and identifies what gaps exist. Working towards some framework will develop a strong picture of what is impacting student learning and ways instructors can better support their students.

1.3 Thesis Statement

[Thesis Statement] Prior work in CS1 research has primarily looked at one or two factors related to student outcomes. An **exploration** of the introductory computing research literature through the lens of an educational model designed to frame student learning, demonstrated how a multitude of factors within this space are contributing to student learning and success in CS1 courses. Two key areas identified in this literature review included instructional and student related factors. We hypothesize the student experience in CS1 courses are impacted by a compounded effect of both student experiences and the pedagogical techniques employed in these courses. We further hypothesize there is no single reason why students are struggling within their CS1 courses, and the extent and impacts of these struggles are greater effecting students from underrepresented groups in computing (women, Black, LatinX, Native American, Pacific Islander (BLNPI), and transfer students). We believe a **holistic** approach towards surveying students will uncover various sources of struggle that are impacting students' ability to prosper in their early computing courses. Furthermore, we hypothesize that best practice **pedagogy** can positively impact student learning, but is not sufficient to help all students.

These hypotheses are addressed by the following research questions studied in this dissertation:

- **Exploration:** What factors related to CS1 student success have been explored in the literature?
- **Holistic:** What factors do students in early computer science courses, including CS1, report through course experience surveys as being sources of struggle and barriers to their success and experience in their courses?
- **Pedagogy:** How does implementing multiple best practice pedagogical approaches in a CS1 course improve outcomes?

The following chapters are organized in the order above: exploration, holistic, and pedagogy. In regards to exploration, we discovered that the CS1 research literature is vast but the results can be organized into an existing educational framework, Biggs' 3P model. We find the majority of studies focus on one or two factors and these factors focus primarily on student and instructor factors and their direct correlation to some CS1 outcome such as pass rates or retention. In terms of holistic, students reported a wide range of struggles impacting their ability to perform in their early computing courses, including CS1. Lower performing students reported higher stress across more areas than their high performing peers. In terms of pedagogy, we found promising results of implementing a trio of best practices in a CS1 course. However, despite overall improvements in course pass rates and major retention, women and BLNPI students were still underperforming compared to their peers. In the discussion and conclusion we will revisit these research questions and summarize the key findings presented in this dissertation.

1.4 Acknowledgements

Chapter 1, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Chapter 2

Background and Related Work

This chapter discusses work focusing on improving student performance in CS1, literature reviews related to CS1, factors impacting student performance, pedagogical interventions, and a description of the Biggs' 3P educational model.

2.1 Pedagogical Best Practices

The manner in which material is delivered is clearly important for student learning. For much of history, the traditional lecture-style approach of an instructor speaking to students in front of a classroom has been the primary method of instruction across universities. However, work has demonstrated that this approach may not be the most efficient in helping students learn and researchers have sought to develop new more effective ways of teaching material.

Pedagogy development and evaluation is a large part of computer science education research. Different styles of instruction have been developed and refined over years. For the work covered within this dissertation, we focus on three well-studied pedagocial practices: peer instruction, media computation, and pair programming. In this section we describe prior work related to the relevant best practices and the theoretical underpinning for these practices.

Prior work has focused on ways we could improve student performance and retention in CS1 courses. The majority have focused on some form of pedagogical intervention, usually around some form of active or team-based learning. In 2018, Latulipe et. al. looked at the longitudinal effects of implementing a flipped classroom approach in CS1 [67]. They studied student performance and retention metrics during and after CS1 and found CS1 performance to be higher for students in the flipped classroom version. They also found women and underrepresented minority students who took the flipped classroom version had a higher one year retention in the major. Another study in 2013 by Porter and Simon demonstrated the impact multiple best practice pedagogies can have in CS1 performance [111]. Using a trio of best practice pedagogies (peer instruction, media computation, and pair programming), helped improve pass rates in CS1 and retained nearly one third more majors in the program.

2.1.1 Peer Instruction

Peer Instruction [25] is an interactive pedagogical approach where students prepare for lecture by reading relevant material, then attend lecture prepared to discuss with peers and the instructor [170]. In class, the instructor poses several multiple choice questions; for each, students individually think about the question, discuss with peers for several minutes, and vote on the correct answer. Peer instruction has been implemented in early mathematics and physics courses, demonstrating increased student mastery of the material [26].

Peer Instruction has been studied extensively in computing, finding that Peer Instruction: is valued by students in lower- and upper-division courses at both large research-focused universities and small liberal arts colleges [70, 114, 117]; shifts students' classroom engagement from passive to interactive [137]; results in in-class learning, both from peers [113] and from the instructor [168]; reduces failure rates [28, 115]; results in improved final exam scores [138, 167]; and provides data useful to researchers for identifying key concepts and struggling students [73, 76, 75, 116].

Recently, Porter and Simon described the factors that led to Peer Instruction being adopted by portions of the computing education community [112]. However, some barriers continue to exist for some faculty wishing to adopt Peer Instruction in their courses [17]. For more detailed summaries of Peer Instruction, please see Simon et al. [140] and Porter and Simon [112].

Peer Instruction is strongly grounded in constructivist learning theory, specifically socioconstructive learning [156]. Peer Instruction replaces lecture from "sage on the stage" with instructor as "guide on the side." Through carefully crafted multiple-choice questions (targeting students' zone of proximal development), Peer Instruction provides students with a scaffolded opportunity to challenge themselves with new concepts and address common misconceptions. The peer discussion phase allows them to develop their own understanding through discussion and explanations to each other. Under the "Interactive, Constructive, Active, and Passive" Framework proposed by Chi and Wylie [23], they connect student behaviors during active learning to cognitive engagement. A study in CS found that students in traditional lectures report lower levels of engagement than those in a Peer Instruction class [137]. While we know of no research that explicitly measures the impact of Peer Instruction on specific social-psychological measures such as growth mindset and social belonging, there seem clear connections. The vote-discuss-vote process of learning in class demonstrates that we can all grow our understanding [14]. Involving students in discussion, becoming acquainted with several other students in the class, and observing other students' thinking could impact students' sense of self-belonging [162]. Peer Instruction also offers a form of cognitive apprenticeship in learning to analyze programs. By bringing analysis and discussion of code, which is often hidden behind the focus on code writing in introductory programming classes, Peer Instruction supports the "enculturation of students into authentic practices through activity and social interaction" [19].

2.1.2 Media Computation

Media Computation is a contextualized computing curriculum developed to teach programming to a broader range of students than solely those focused on becoming computing professionals [44, 48, 121]. The curriculum teaches standard introductory programming concepts in the context of manipulating images and sounds [44, 47].

Adoption of media computation has been connected to reduced failure rates [45, 110] as well as improved retention of students in computing [149, 142, 110]. A goal of the media computation curriculum was to improve outcomes for women in computing [45] and researchers found that women find the course more motivating than prior offerings [37], appreciate the media context [121], and may appreciate the opportunities for creative expression [5]. Passing rates were found to be balanced between men and women in the media computation version of the course [45].

Media Computation as a context for introducing computer programming is also grounded in the theory of situated cognition [68]. It recognizes that (as of the early 2000's) computers had, in general society, moved from being considered a computational tool to being a generally used communication tool. With Media Computation, students experience programming as "situated in activity bound to social, cultural and physical contexts" [43]. Additionally, as described in Forte and Guzdial [37], media computation-based assignments allow for much more creativity than traditional computational-focused assignments. These may improve learning as "the constructionist approach to learning asserts that people learn particularly well when they are engaged in constructing a public artifact that is personally meaningful" [37]. For extended summaries of the work on media computation, please see Guzdial [45, 46].

2.1.3 Pair Programming

Pair Programming is a cooperative learning approach that involves students working closely together while programming by alternating roles as "driver" (the person controlling the keyboard and mouse) and "navigator" (the person providing guidance and suggestions) [159]. Pair programming has been extensively studied in computing with three meta reviews appearing between 2011 and 2019 [51, 130, 153]. These meta-reviews find that pair programming is broadly associated with improved outcomes in computer science, including increased enjoyment and satisfaction [51] as well as improved grades on programming assignments, overall grades, and pass rates [153]. Pair programming has also been shown to improve retention of women in computing [91].

Not all pairings of students are equally effective as students with similar programming skills appear to be more successful [51]. Moreover, a recent qualitative study explored the components of inequitable pairings [72]. A longstanding concern with pair programming is that some students may not contribute enough to the pair to learn effectively [130, 166], however findings remain mixed.

Pair programming, and possible benefits from its use, can be explored through the lens of several theories of learning also seen in the two previous best practices. Pair programming is not just a constructivist learning approach but specifically a socioconstructivist one [156]. Pair programming, if explained to students as an industry-practice, can also contribute to students' sense of social belonging, although it is not really a form of legitimate peripheral participation as their community of practice only contains other novices [68]. Please see Simon et al. [140] for a detailed summary of Pair Programming.

2.1.4 Combining Best Practices

One prior study has examined the combined effects of Peer Instruction, Media Computation, and Pair Programming [110]. In their study, the authors report that the offering of the course reduced failure rates and improved retention of majors 1-year after taking the course. Our present study examines a larger set of students from across a longer time period and does so with enough time passed to evaluate long-term outcomes for students including retention in the major through graduation, grades in following courses for the major, and time-to-degree.

2.2 Other Factors Impacting Student Performance

A variety of factors associated with student success have been studied by the community. Factors associated with student success include prior experience [161], sense of belonging [109, 102, 131, 71, 155, 89, 90, 94], interest in course content [22], achievement goals [171], self efficacy [79], study skills [74], and alignment between student goals and the perceived goals of the field [71]. Which students succeed may be impacted by the structure of CS courses or curricula, as the way courses are taught [115, 112, 128], the learning environment [8, 7, 39, 40, 132], and the competitive nature of the program [102] (among others). In this section we review the theory and prior work related to the factors most closely related to those examined in our current work.

2.3 Social-Emotional Factors

Situated learning theory asserts that there exists an important relationship between the act of learning and the social context in which the learning occurs [69]. Legitimate Peripheral Participation says that learning occurs within a community and that "...mastery of knowledge and skill requires newcomers to move toward full participation in the sociocultural practices of a community" [69]. From this lens, if a student feels they do not belong in that community, it may be difficult for them to fully participate in the community, potentially limiting their ability to make connections with their peers and potentially impacting their performance in the course.

Studies in computing have shown how sense of belonging has been associated with various important student factors such as performance, retention, motivation, and persistence in computing, with race and gender being significant predictors of feelings of belonging [155, 41, 90, 131, 71, 63]. A study focusing on why students leave CS1 shows that sense of belonging from lack of social groups can impact the student experience [109]. Understanding how student belonging can be an influence might help in retaining more students. A recent study focused on how providing students networking, outreach, and mentoring opportunities positively influenced their sense of belonging in computing [94]. Apart from hurting a student's academic success, this lack of belonging has also been tied to feelings of depression [50].

Other studies have focused on self-efficacy in computer science and its association with anxiety, success, and interest in the field [32, 93, 161]. The concept of self-efficacy makes the claim that individuals with high self-efficacy are likely to face difficult challenges while those with low self-efficacy will try to avoid them [6]. Studies have shown that women in computing struggle with self-efficacy, which can impact their performance [12, 15].

Prior studies that have observed student outside obligations, such as work and family, have focused on how work-life balance is associated to well-being, anxiety and depression. Although we know these issues can have an impact on student performance, to our knowledge, no prior study has directly associated these outside obligations to student exam performance.

2.4 Student Goals or Behaviors Associated with Success

Goal congruity theory argues that student personal goals should be aligned with opportunities to achieve those goals in a field in order for students to enter and persist in the field. When these goals are not aligned, students are more likely to exit the field [29, 30]. In a recent study, Lewis et.al., demonstrated how high communal goals led to a lower sense of belonging among computing students [71]. Since communal goals have been shown to be heavily endorsed among women and BLN+ students, two groups that computing education struggles to retain [143, 151], a lack of alignment between goals and perceived opportunities in CS (and STEM) may be a source for the lack of retention of women and BLN+ students.

Bandura's Social Cognitive Theory asserts that learning occurs in a social context within a dynamic interaction between people, the environment, and behavior [66]. Focused on the behaviors of students, a recent study observing behaviors of high and low performers in a CS1 course [74] identified different help-seeking behaviors among low and high performing students in terms of who they approached (friends, instructional staff) and the type of help they sought. The availability of help for students is a potential structural (environmental) barrier included in this study.

Achievement goals are students' goals for success in a particular setting. The goals that students' approach (mastery versus performance) have been shown to impact student success in computing courses [169, 171]. Although this area of goal types are relatively new in its application of CS, studies found that students who are focused more on mastering the material (mastery goals) are more likely to perform better in their CS courses and express interest in pursuing CS as a field than students who want to outperform their peers (performance goals).

2.4.1 Challenges due to Students' Personal Lives

A few studies have examined the role of students' personal lives on their success, but this work has generally not been specific to CS. A study of Scottish further education students found that full-time students in the study worked in a job between 16-20 hours per week and had different strategies for coping with the demands of job, school, family, and friends [80]. In a study of Midwestern college students, Sprung and Rogers found that students with a better work-life balance had lower levels of stress and depression [145]. Stratton et al. examined how different obligations impact post-secondary student enrollment and drop-out rates [147] and found that outside employment, marriage, work-study aid, and grants were all connected to the likelihood of staying in the program.

In computer science, previous studies have not examined the link between outside obligations and success explicitly, though a few studies have examined the effect of multiple different barriers to student retention and motivation. One study examined how multiple factors, such as personal values, teaching quality, satisfaction with their learning, and student motivations impacted retention in CS [106]. Another study in CS used a four-factor model to describe student motivations found that students' enjoyment of completing academic tasks and the utility of studying CS were most important to motivation, whereas factors such as the influence of friends and family, were less important [127].

2.5 Relevant Literature Reviews in CS1

With such an extensive amount of literature within CS1 research, many literature and meta-reviews have emerged to organize and record the state of research in all these subareas. The Cambridge Handbook on Computing Education Research is an example of a resource the community has produced to summarize years of work [112]. The book summarizes CSEd history, methods, theories, related fields, and reviews of popular areas of study to date. Some topics covered in the book include computational thinking, systemic issues in CSEd (such as plagiarism and diversity), foundational statistical and qualitative methods for research, and pedagogical approaches to teaching CS. Literature reviews have focused on condensing years of introductory computer science research[10, 82, 108, 118, 92].

Despite such a vast amount of literature, critiques have emerged about how theory is used in CSEd research. A study by Malmi et al. that looked at multiple publications from CSEd venues over the past several years found that nearly half of the papers did not build off any prior theoretical framework[85]. Other studies have highlighted the importance of theoretical frameworks in research. Niss discusses the importance of theory in math education and highlights important benefits theory provides to a research field[103]. Overall, they discuss how theoretical foundations provide some consistency and guidance in explaining behaviors and phenomenon in a field, along with providing guidance and insight into interpreting results. This low use of theory in CSEd has created a concern that this field is building itself on a dispersed theoretical foundation as opposed to a concentrated stable one[85]. In order to address this issue, our work explores how the current state of literature about CS1 success and struggle can be mapped to a well-known educational experience model, Biggs' 3P model.

2.6 Biggs' 3P Model

Biggs' 3P (Presage-Process-Product) model is an educational model that structures the student learning experience as relationships between potentially interacting factors across three phases: presage, process, and product.

The presage phase refers to factors related to students prior to engaging in learning. This phase includes both student-related factors and instructor/teaching-related factors. Some examples of student-centered presage factors include prior experience, mindset, demographics, and their preferred learning approaches. Because Biggs' model focuses primarily on the student in the learning process phase, non-student factors that set up a learning experience from the instructor's perspective are also considered presage. Some examples include teaching philosophy, departmental policies and norms, and the types of assignments and tasks the instructor designs.

The process phase refers to factors related to the actual act of student learning. Biggs defines these factors in terms of a student's learning approach, which they classify as being either a deep or surface learning approach. When a student is presented with a task (assignment, reading, etc.), they decide how to go about completing the task. Based on the technique or process chosen, it likely can be classified as either a deep or surface learning approach. A deep approach refers to actively seeking to understand the material or subject while surface learning is more passive, in which a student focuses on simply reproducing the material without necessarily fully understanding it. Prior work in CSEd has shown the benefits of a deep learning approach over a surface learning approach when measuring programming performance [55, 165, 134]. In short, the process phase describes the techniques (active or surface) that a student decides to use when engaged with a task.

The product phase is the final phase of Biggs' 3P model and is defined as the resulting experiences and outcomes a student achieves after the learning experience. These outcomes can represent different aspects of student success and experiences, such as performance (e.g., grade), affect (e.g., emotional responses), and preparedness (e.g., skills).

Prior CS1 studies have focused on many different types of factors and their interactions in relation to student outcomes. The Biggs' 3P model can capture this dynamic interaction and interrelationship that appears to exist during a student's CS1 experience. The model defines learning as an ever-evolving process where a student approaches a learning situation with some mindset based on existing qualities, experiences something within a learning context, and emerges with some change in qualities or skills that can influence the next learning experience. We believe that mapping current CS1 research to this framework can provide important insights into the status of this research area.

2.7 Acknowledgements

Chapter 2, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. "A Longitudinal Evaluation of a Best Practices CS1", 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. "Understanding Sources of Student Struggle in Early Computer Science Courses", 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. "Applying CS0/CS1 Student Success Factors and Outcomes to Biggs' 3P Educational Model". The dissertation author was the primary investigator and author of this paper.

Chapter 3

Study Context

This chapter describes the context of the data sets used in this dissertation. Chapter 4 is a literature review on CS1 literature so the context of the papers will be discussed in that chapter. However, chapters 5 and 6 are quantitative studies conducted at the University of California San Diego. Context for these chapters will be discussed here.

3.1 Study Context for Chapters 5 and 6

3.1.1 University Context and Demographics

These studies were carried out at the University of California San Diego, which is a large US research university (R1). The university runs on a quarter system, meaning courses are meet for 10 weeks of instruction followed by a week of final exams.

3.1.2 Course Descriptions

The courses in these studies include CS1-NPE (No Prior Experience), CS1-PE (Prior Experience), CS2, and Computer Organization (CompOrg). University course catalog descriptions are listed below [104].

1. **Introduction to Programming and Computational Problem-Solving I (CS1/CS1-NPE).** This is an introductory course designed for students who do not have prior programming experience. The current course description on the UCSD catalog is: “Introductory course

for students interested in computer science and programming. Basics of programming including variables, conditionals, loops, functions/methods. Structured data storage such as arrays/lists and dictionaries, including data mutation. Hands-on experience with designing, writing, hand-tracing, compiling or interpreting, executing, testing, and debugging programs. Students solve relevant computational problems using a high-level programming language.""

2. **Introduction to Programming and Computational Problem-Solving: Accelerated Pace (CS1-PE).** This is the accelerated version of the introductory course and assumes students have some prior experience with programming. The current course description on the UCSD catalog is: "Accelerated introductory programming including an object-oriented approach. Covers basic programming topics from CSE 8A including variables, conditionals, loops, functions/methods, structured data storage, and mutation. Also covers topics from CSE 8B including the Java programming language, class design, interfaces, basic class hierarchies, recursion, event-based programming, and file I/O. Basics of command-line navigation for file management and running programs.""

3. **Basic Data Structures and Object-Oriented Design (CS2).** The current course description on the UCSD catalog is: "Use and implementation of basic data structures including linked lists, stacks, and queues. Use of advanced structures such as binary trees and hash tables. Object-oriented design including interfaces, polymorphism, encapsulation, abstract data types, pre-/post-conditions. Recursion. Uses Java and Java Collections.

4. **Computer Organization and Systems Programming (CompOrg).** The current course description on the UCSD catalog is: "Introduction to organization of modern digital computers-understanding the various components of a computer and their interrelationships. Study of a specific architecture/machine with emphasis on systems programming in C and Assembly languages in a UNIX environment.

3.1.3 Student Population Characteristics

The courses in these studies are all required to be taken by students in order to earn a computer science degree, whether that be a major or minor. Students from computing adjacent majors, such as data science and math-CS, are also required to take these courses. Students who do not plan on majoring in any computing related major can also take CS1 to satisfy a general education requirement needed to graduate.

3.2 Acknowledgements

Chapter 3, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Chapter 4

[Exploration] Exploring the Realm of CS0/CS1 Computer Science Education Research

4.1 Introduction

For many students, introductory computer science (CS0/CS1) is the course that introduces them to the world of computing. For many others, it is notoriously the course that deters them from considering computer science as an academic or professional pursuit. In an attempt to understand and address student experience and performance, researchers have sought to study the impacts of learner characteristics, interventions, and pedagogies on student learning and experiences in CS0/CS1. The factors themselves can range across many categories such as sense of belonging[62], self-efficacy[102, 59], instructional practices[98, 136, 128], peer learning[115, 139], cognitive ability[4, 78], and achievement goals[171, 169]. Of these factors, some have consistent and replicated outcomes, such as prior experience being positively correlated with higher performance or women generally entering CS0/CS1 with less prior experience than men[49, 160, 158, 18]. However, other factors are understudied or produce inconsistent results, limiting our knowledge of how to best support students. To our knowledge, there is no taxonomy that attempts to survey and organize the extensive amount of CS0/CS1 literature in a way that organizes factors into categories, highlights what has been heavily studied, explores the relationships between factors, and identifies what gaps exist.

For this study, we used Biggs' 3P model because of its dynamic qualities and relationships between three phases: presage, process, and product. At its core, this model describes the learning process as an interaction of many factors from different sources of origin and quality. We use this model as a framework to organize the current CS0/CS1 literature by mapping papers in our review to the different components of the 3P model in order to highlight what aspects of the student experience have been studied. Our results of this study are as follows:

1. The research landscape does not sufficiently take a student's learning process and engagement into account. While outcomes are heavily studied, the actual learning process a student has with the material is seldom evaluated.

2. Many studies do not focus on understanding how student factors are evolving or how those factors are specifically impacting how students engage with materials or how that engagement evolves and changes.

4.2 Study Design

4.2.1 Research Questions

Our study addresses the following through the lens of the 3P model:

- How do CS1 research papers map onto the Biggs' 3P model? Which areas have high concentrations of papers? Which areas have low concentration of papers?

4.2.2 Paper Selection

The goal of this project is to explore a sample of the CS0/CS1 literature and map papers to the 3P model. To do this, we searched for relevant literature related to success or struggles in CS0/CS1. Key terms were selected based on desired student outcomes (retention/performance), student characteristics (self-efficacy and BLNPI/URM), and pedagogy. Pedagogy was selected because the research team sees this as a research area that actively aims to improve student

Table 4.1. Set of key terms that were used to scrape research papers online. These values were for X in the query, "Factors influencing X in Y".

Interest	Key Terms (X)
Retention	retention, attainment, success, failure, drop outs
Self-Efficacy	self-efficacy, sense of belonging, motivation, mindset
Performance	course performance, exam performance, assignment performance, programming performance
BLNPI (URM)	minority student, women, underrepresented student
Pedagogy	pedagogy, peer instruction, media computation, pair programming, team based learning, best practices, live coding, active learning, flipped classroom, POGIL

outcomes. Synonyms and phrases to represent these key terms were then brainstormed amongst the research team.

We created a Python script to query Google Scholar using the following search phrase, "Factors influencing [X] in [Y]". We decided to use Google Scholar as it acts as an aggregator of papers across databases. Although we acknowledge that not all papers will be selected, our intention is to get a representative sample, not all existing papers. The possible values of X are key terms related to popular areas or outcomes from CSEd literature. The possible values of Y are context terms that describe introductory CS which include *introductory computer science course*, *CS1*, *CS0*, and *introductory programming*. Refer to Table 4.1 to see the key terms for X that were searched. An example of a query built from the terms is, “Factors influencing *drop outs* in *CS1*", where “drop outs" is the X value in the interest of retention, and “CS1" is the Y value.

After running the script, we collected 2090 papers in our initial dataset. The research team then read the titles and abstracts of the papers and removed any that had no clear connection to CS0/CS1 or were duplicates. This reduced the number of papers to 947.

A key aspect of this work is that we wanted to examine traditional university/post-secondary CS0/CS1 contexts. CS0/CS1 has gained popularity and has expanded into the K-12 and online space throughout the years. To remain within scope, there were certain criteria used for inclusion of papers.

Academically Published

The paper must have been published in an academic conference or journal as an article. We omitted graduate thesis, dissertation, working group, poster presentation, abstract-only, and doctoral symposium papers. We believe work found in these articles will be available in other research articles by the same author. Textbooks were also not considered as we believe the information found in them would be based on relevant literature that our search would include.

Post-Secondary CS0/CS1

Because we want this work to focus primarily on traditional post-secondary introductory CS offerings, we omitted any research paper that reported on other types of offerings. Papers on offerings of CS0/CS1 in the K-12 or MOOC space were removed from our analysis. We believe the settings of these offerings create new challenges that are unique to that setting and not typical in traditional offerings. Additionally, we only included papers that included samples from CS0/CS1 exclusively. If other courses were studied in the same paper, we only included the paper in our analysis if the results for CS0/CS1 participants were independent of the other course samples.

Experience Reports

We did not include experience reports if they did not provide sufficient statistical analysis or did not provide sufficient exploratory/explanatory power in their results sections. This criterion also excludes studies that report on a learning tool's or intervention's development and mentions future plans for implementation.

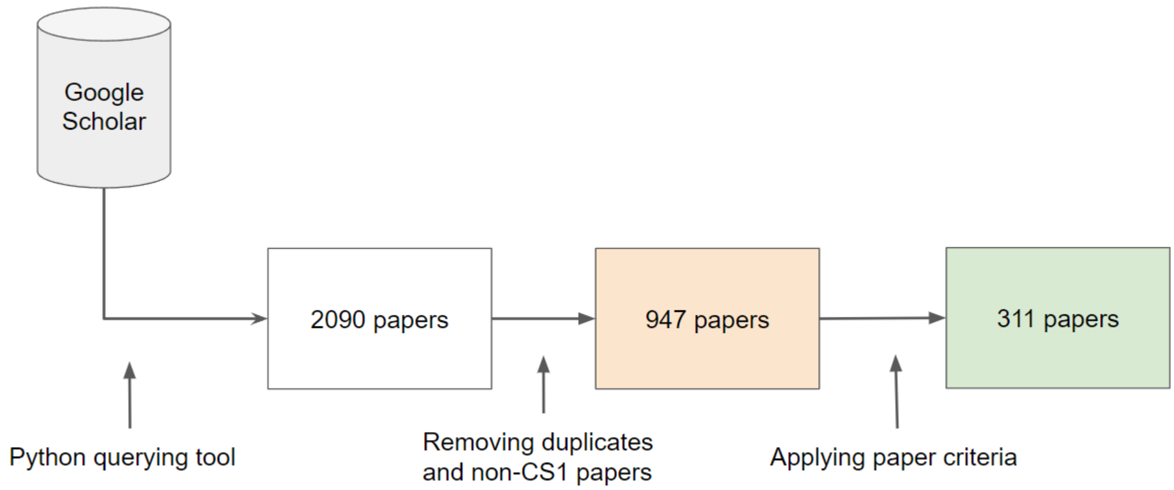


Figure 4.1. Number of papers after each phase of our sampling procedure.

Qualitative, Quantitative, and Mixed Methods

We included papers that used qualitative, quantitative, or mixed-method approaches. However, we required that qualitative studies must include more analysis aside from simply stating feedback from participants. Some level of qualitative analysis, such as coding responses or phenomenography must be used to be included. Quantitative studies must provide some form of inferential statistical or empirical analysis to be included, such as significance testing, regression coefficients, or effect sizes.

Once these criteria were applied, our final sample included 311 papers. Figure 4.1 illustrates how many papers we had at each phase of our literature search. The full list of papers can be found at https://github.com/adsalgue/SIGCSE2024_3PModelPaper.git.

4.2.3 Analysis Process

The following sections describe the process of analyzing the 311 papers in our final dataset. We describe the content we gathered from each paper and how papers were grouped and analyzed.

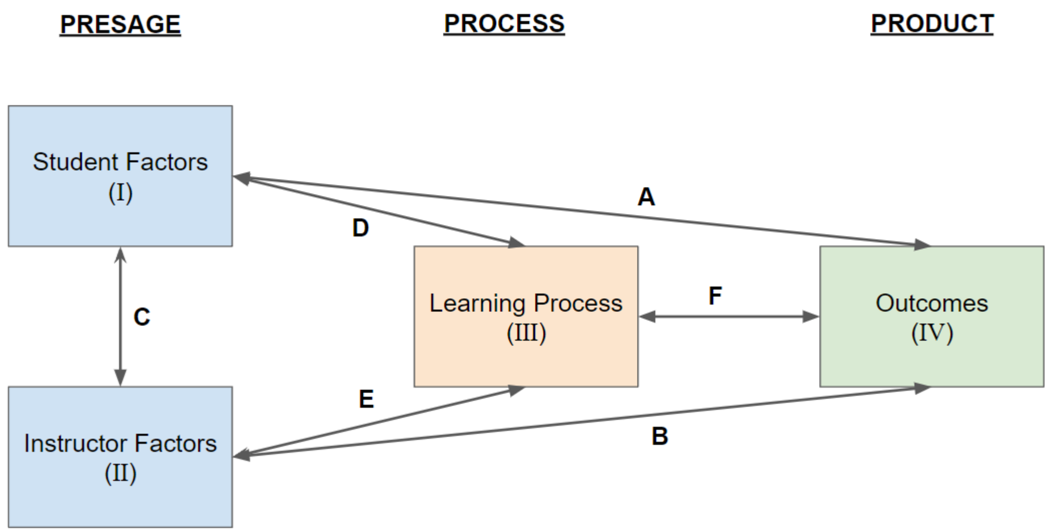


Figure 4.2. Labeled Biggs' 3P model. I-IV are the nodes and aspects that make up each phase. A-F denote the relationships between these groupings. Since Biggs' model is dynamic, each of these factors across the phases can potentially interact and influence each other.

Data Gathered Per Paper

For each paper we recorded the relationships, outcomes, and conclusions made by the paper. Relationships between factors and outcomes of the paper were the most important pieces of information we recorded and are at the core of our results. Although many papers report on many relationships, we only considered statistically significant outcomes in quantitative studies. In qualitative studies, due to the high likelihood of there being no statistical testing, we decided to include the relationships reported while marking results being qualitative in nature. Once again, qualitative studies were required to have some level of qualitative analysis, such as coding responses. Other information such as sample size, sample characteristics, data collection methods, and data analysis methods were recorded but not part of our mapping to the 3P model.

4.2.4 Analyzing Papers

Mapping to Biggs' 3P Model

For our mapping process, we read and recorded information from 311 research papers. Of those studies, 297 reported significant findings. We only consider papers with significant findings in our 3P model mapping. For example, if a study looked at gender and pedagogy in relation to student CS0/CS1 performance, but only gender was statistically significant, we only include gender in our mapping and not pedagogy. After recording all the necessary information from our papers, we mapped the significant outcomes and their respective factors onto the 3P model as seen in Figure 4.2. We will use graph notation of nodes and edges to discuss our results. For the mapping we define the nodes (I-IV) as the following:

- **I - Student Factors**: These are factors related to the qualities students possess across a variety of different areas such as technical skills, psychological, cognitive, and self-perceived.
- **II - Instructor Factors**: These factors are related to the conditions of the course that instructors employ and have some level of control over such as assignments, pedagogy, learning environment, and evaluation metrics for student performance.
- **III - Learning Process**: These factors relate to things involved in the learning of materials during the course. These include factors related to how students engage with the learning process and material such as learning approaches and various in-class behaviors such as help seeking.
- **IV - Outcomes**: These are the different types of outcomes students can achieve after a learning experience such as grades or emotional responses.

The edges (A-F) are meant to represent associations or relationships between the types of factors above. These edges do not indicate that the relationship is causal, rather that simply one exists. These edges will highlight the majority of outcomes reported in the papers of our literature review.

Factors were mapped onto a phase and outcomes onto one of the arrows in the 3P model. For example, if a paper focused on measuring the relationship between sense of belonging prior to the class and their final exam results, this would map to the Presage and Product phases along with the arrow connecting the two. Once we recorded all the necessary information, we used negotiated agreement in the mapping process. The first author made an initial pass through each of the study outcomes and mapped each factor to one of the nodes seen in Biggs' 3P model (i.e. Student Factors, Instructor Factors, Learning Process, and Outcomes). After the first author mapped all the factors, the second author reviewed those mappings. At this time, authors one and two discussed the mappings of each paper's outcomes based on the node definitions seen above. Initial agreement on factor mapping is labeled as a raw agreement. When author two disagreed with the mapping done by author one, a discussion was held to reach a consensus on which mapping would be correct. This involved discussing what nodes best define the factors in the paper's outcomes. Once again, heavy use was made of the node definitions. When a consensus was reached between authors one and two, this was labeled as a negotiated agreement. In the cases where no consensus was reached despite much discussion, these were labeled as disagreements.

The same process was conducted for mapping each paper's outcomes to the edges of the 3P model. After completing the negotiated agreement for the nodes, the first author created mappings of edges for each paper on their own. The second author was then brought in to discuss edge mappings. Table 4.2 and 4.3 lists the results of our negotiated agreement process of mapping papers to the nodes and edges of the 3P model respectively.

Table 4.2. Results of negotiated agreement process for 3P nodes.

	Count	Proportion
Raw Agreement	228	77%
Negotiated Agreement	64	22%
Disagreement	4	1%

Table 4.3. Results of negotiated agreement process for 3P edges.

	Count	Proportion
Raw Agreement	234	79%
Negotiated Agreement	64	21%
Disagreement	0	0%

Table 4.4. Number of papers mapped to each node in Biggs' 3P model.

Node	Count	Proportion
Student Factors (I)	171	58%
Instructor Factors (II)	161	54%
Learning Process (III)	103	35%
Outcomes (IV)	268	90%

Table 4.5. Number of papers mapped to each edge in Biggs' 3P model.

Edge	Count	Proportion
Student Factors <-> Outcomes	146	49%
Instructor Factors <-> Outcomes	132	44%
Student Factors <-> Instructor Factors	3	1%
Student Factors <-> Learning Process	29	10%
Instructor Factors <-> Learning Process	32	11%
Learning Process <-> Outcomes	56	19%

4.3 Results

4.3.1 Node Analysis

Table 4.4 shows where the counts of papers that had significant findings fall in each of the nodes of the 3P model. We see that over half the papers reported conclusions involving Student Factors, Instructor Factors, and Outcomes. Papers that fell into student factors focused on areas such as gender, race/ethnicity, sense of belonging, prior experience, attitudes, and cognitive abilities. It is surprising that just over half of our papers had significant findings on Student Factors. Considering how CS0/CS1 research is focused heavily on the student experience, we would have expected more papers to fall under this category. Papers that fell into the Instructor Factors category were primarily focused on student responses to different pedagogy, curriculum, interventions, or types of assignments. Papers that fell into Outcomes focused on some sort of assessment (exam, programming assignments, test score, etc.) or on a student's change in ability or attitude. It is understandable that a large proportion of papers map onto Outcomes as a shared goal of CS0/CS1 research is to examine how students perform at the end on some sort of performance or qualitative metric. Finally, less than half of the papers had findings related to the student learning process in CS0/CS1. This is an important finding because it shows not many papers explored how a student reacts and learns during a course, only the final overall outcome. This is an important aspect of the student experience that future studies should focus on, primarily which factors impact students in their learning process and how certain factors cause them to directly engage (or disengage) with CS0/CS1 material.

4.3.2 Edge Analysis

Table 4.5 shows the count and proportion of edges that were mapped using significant relationships and outcomes reported in our sample of papers. Note that each paper often mapped more to one edge, which is why the proportions do not sum to 100%. Our mapping of the edges to Biggs' 3P model indicates large gaps in the types of significant relationships and outcomes being reported on in the CS0/CS1 literature.

The majority of papers mapped to the edges connected Student Factors to Outcomes and Instructor Factors to Outcomes. These connections are understandable as many papers in CS0/CS1 focus on how student factors (sense of belonging, CS interest, race/ethnicity, and prior programming experience) relate to their performance or other metrics in a CS0/CS1 course. Similarly, with correlations between Instructor Factors and Outcomes, pedagogical interventions and improvements are a major part of CS0/CS1 research, which would fall under the instructor factor node. The research primarily focuses on how pedagogy can impact student outcomes such as pass rates, retention, and interest. Therefore, a high proportion of papers exploring such relationships was expected.

However, when we begin exploring the edges related to the student learning process node (edges D-F), we see a large drop in the number of mapped papers. This indicates that CS0/CS1 research is not focusing on identifying significant correlations between student learning processes and other factors. Many papers are missing observations and data about how specifically certain factors are influencing student engagement and how they choose to approach material. Is a certain pedagogy or student personal factor causing them to engage more with the material and build interest, in turn allowing them to employ deeper learning techniques? Although assumptions can be made, additional observations and data gathering/analysis will be needed. This is crucial information that requires more research as it is important to understand how factors impact the learning process during the course and if they are truly responsible for the observed outcomes. An updated Biggs' 3P model showing our mapping results can be found in Figure 4.3.

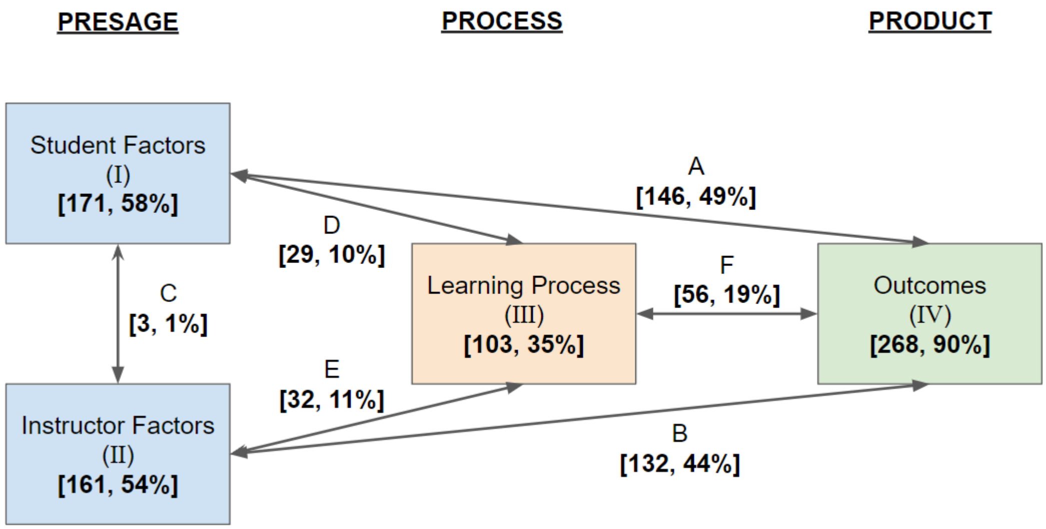


Figure 4.3. Updated Biggs' 3P model with our results from our node and edge analysis in the format [n, %] where n is the raw number of papers mapped and % is the percentage in relation to the entire sample.

4.4 Discussion

Our results highlight areas that have been studied in CS0/CS1 research and opportunities to further our knowledge in future studies.

4.4.1 Presage and Outcomes

Our findings indicate that many studies focus on understanding the relationship between a student's presage factors and outcomes. Researchers appear to use these factors as a prediction tool to see how students will perform on performance metrics such as exams or grades. Although this information is valuable, we believe it will be invaluable to the community to understand the interaction among these presage factors and how those directly impact the student learning process. Although papers do not outright ignore the learning process of the student, data gathering and analysis within this phase is small compared to the Presage and Product phases.

4.4.2 Importance of Learning Process

The processes students employ when learning is a key aspect of the student experience. Their level of engagement along with the techniques they employ merits closer investigation as it can provide a clearer picture of their experience in a CS0/CS1 course. Although certain presage factors appear to be correlated with student success, the research community needs to explore how these factors influence student decisions during the course rather than only identifying correlations to outcomes. For example, one study can focus on understanding how prior knowledge impacts coping mechanisms during the course and how those mechanisms are evolving during specific milestones in the course.

4.4.3 Commonly Appearing Factors

During this project, the research team recognized many common factors with similar outcomes. We noticed heavy evidence showing factors such as prior experience and active learning to be associated with positive outcomes. Meanwhile, factors such as low sense of belonging, low interest, and lack of prior experience were associated with negative outcomes. Given the abundance of studies on these topics, our findings suggest further work on these topics should examine the relationship between these factors and the student learning process to better refine our understanding of the role these factors play.

4.4.4 Call for Theory

The application of this work to Biggs' 3P model aims to highlight the pressing need for a CS-specific theory to explain the CS0/CS1 student experience. Currently, there are some popular theories being used in the field such as Bandura's theory of self-efficacy, achievement goal theory, and constructivism, to name a few. This overarching analysis shows that CS0/CS1 research literature is fractured into small pockets of interrelated studies. We encourage future work to focus on consolidating this current body of research into a common theoretical framework to support a broader understanding of how various factors relate to student learning and subsequent outcomes.

4.5 Conclusion

In this study, we mapped 297 out of our 311 original CS1 papers onto Biggs' 3P educational model that describes student learning. The majority of papers focused on instructor and student factors and their direct relationship with outcomes and did not include the actual process of a student engaging and learning the material. We found that a lower concentration of papers focused on the actual learning processes and techniques a student employs in learning CS during a course.

Overall, our study indicated several areas for future work in this field. We have seen in prior work these factors can have interacting effects on students so now it is the time to begin formulating and organizing all this research into cohesive frameworks that consider the interactions among different types of factors and how they affect learning throughout the learning process. In the future this could perhaps lead to the development of a CS0/CS1-specific theory to understand the impact factors and their interactions have on student learning in these courses.

4.6 Acknowledgements

This work was supported in part by NSFaward #2121592, as well as a UCSD Sloan Scholar Fellowship, a UCSD STARS Fellowship, and a Gates Millennium Scholarship. Chapter 4, in part, is a reprint of the material as it appears in the Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1

Student Success Factors and Outcomes to Biggs' 3P Educational Model". The dissertation author is the primary investigator and author of this paper.

Chapter 5

[Holistic] Student Experiences in Early Computer Science Courses

Computing education is well known to suffer from poor retention of students interested in computing [13, 128] and students not learning what instructors expect [81]. As such, researchers have sought to understand why some students struggle in computing courses by examining a myriad of factors that might be connected to student success [11, 83, 157, 161]. In the United States (the context for our work), there also exists substantial divides by gender and race among those who major in computer science [38] and in those who have access to computing education before college [86, 87]. In addition, Black, LatinX, and Native American and women students who pursue computing in university are retained at lower rates than White and Asian men [172].

Given these challenges, prior studies in computer science education have investigated the impact of a variety of factors on students' outcomes. Some of these factors include prior experience [161], sense of belonging [109, 102, 131, 71, 155, 89, 90, 94], interest in the material [22], self efficacy [79], study skills [74], and alignment between student goals and the perceived goals of the field [71].

These previous studies typically study a single factor in isolation, making it difficult to gain a holistic view of multiple factors that might interfere with a student's learning. Some students might be affected by one factor and other students a different one. Yet other students may be affected by more than one factor at the same time, perhaps compounding their negative effects.

Single-factor studies cannot see this larger picture. Similarly, many studies have focused on a single course [109, 74, 125, 171] (typically CS1), perhaps masking challenges that evolve over time, or are unique to a specific topic or timing in the curriculum. Finally, few of these studies have examined the role of students' personal lives (e.g., outside work commitments, family obligations) on their success; existing work in this area is not specific to CS [80, 145]. These limitations might partially explain why some interventions directed at mitigating a particular factor (e.g., meta-cognitive skills or growth mindset) have not had greater impact [146, 20, 60].

The present study investigates how multiple potential sources of student struggle relate to outcomes across multiple computer science courses at a research-intensive university in the United States. Specifically, we investigated how non-academic barriers (e.g., illness, family care) might interact with social emotional and social cognitive barriers (e.g., lack of belonging, lack of peer networks) or structural barriers (e.g., lack of peer support, inability to get help). Drawing on previous studies and our own experience as instructors, we developed a broad survey instrument that was compact enough to be administered regularly with high participation rates. Instructors administered it multiple times throughout the Fall 2019 term in four programming-intensive courses ranging from CS1 to Computer Organization. We then employed Exploratory Factor Analysis to cluster questions into cohesive categories of factors. Four factors emerged: personal obligations, lack of sense of belonging, in-class confusion, and lack of confidence. We analyzed how these factors related to student performance on the final exam in each course. Finally, we examined differences by gender, race/ethnicity, and matriculation status,[1] again relating reported factors to performance on the final exam. Our results can be summarized as follows:

- When students struggle, they often struggle on multiple fronts. Over 70% of students in the lowest quartile of final exam performance report high levels of stress due to at least one of the four identified factors compared with less than half (30%) in the highest quartile. Over 50% of students in the lowest quartile report high levels of stress for two or more

[1] In the United States, 2-year community colleges provide high school graduates with an opportunity to prepare themselves for acceptance to a 4-year university, for the latter half of their studies, at a lower net cost.

factors. By comparison, half as many students in the next higher quartile report struggling with multiple factors.

- For a given level of performance, women students; Black, LatinX, Native American, and Pacific Islander (BLN+) students; and transfer students report slightly more factors overall interfering with their performance, especially in the lowest quartile of final exam performance. However, which factor interferes the most depends on the demographic.

- Survey response rates for students in the lowest quartile of performance were significantly lower than for higher performing students. This may be a signal of struggle, rather than having given up, because these students still persisted to take the final exam. Because their voices weren't heard, our other results regarding students in the lowest quartile of performance are likely conservative. We believe this result is not unique to our work and exposes an under-acknowledged threat to this type of research (participation rates may vary within subgroups).

- Students in the CS1 course for students with no prior experience reported considerably more struggle than students in the CS1 course for students with prior experience, across all levels of exam performance. This result suggests that computing students with no prior experience need additional support.

5.1 Study Design

5.1.1 Research Questions

As motivated in the Introduction, our study addresses two research questions:

- **RQ1**–What high-level social-emotional/social-cognitive, structural, or personal factors are related to student outcomes?

- **RQ2**–How do any identified factors and associated outcomes vary by demographic group?

5.1.2 Course Context

Our study was conducted in the Fall of 2019 at a large US research-intensive university operating on a quarter system with an 11 week term (10 weeks of instruction and one week for final exams). We studied four undergraduate lower-division computer science courses: A CS1 course requiring no prior programming experience (CS1-NPE); a fast-paced CS1 for students with prior programming experience (CS1-PE); a CS2 course that includes Basic Data Structures and Object-Oriented Programming; and a Computer Organization course (CompOrg) focused on C, assembly, and basics of hardware design. CS1-NPE and CS1-PE have the same formal prerequisites, but students are advised to take one course or the other based on their prior experience. CS1-NPE is the first course in a pair of courses taken across two terms that together are equivalent to one term of CS1-PE. The second course in the pair was not included in our study because few students take this course in the fall quarter. CS1-PE, CS1-NPE, and CompOrg were taught using active learning techniques (Peer Instruction [26], worksheets, etc.) and CS1-NPE included an interactive textbook with auto-graded programming exercises. CS2 was taught using a more traditional lecture format. A breakdown of our study population can be found in Table 5.1.[2]

5.1.3 Survey Design

Our survey was designed not only for research, but also as a tool for instructors at our institution to gain insight into their students concerns, respond to student issues, and make course corrections throughout the term. This motivated designing a survey that collected actionable data throughout the term and maximized student response rate. As such, we collected data regarding a wide range of possible struggles drawn from previous literature, our own experience as educators and instructors, and student feedback. As researchers, we hoped a broad survey would allow new perspectives on these potential struggles to emerge during the analysis. Question (6) is the

[2]We acknowledge that gender extends beyond simply male and female. However, our data was collected as male/female so to remain consistent we decided to maintain this binary classification in our analysis.

Table 5.1. Summary of courses, with the number of participating students, breakdown by group, number of surveys administered, and quartile size (see Section 5.2.2). Numbers in parentheses give proportions between genders (Male vs. Female), race categories (Non-BLN+ vs. BLN+) and matriculation status categories (First-Year vs. Transfer).

Course	Total Students	Male	Female	Non-BLN+	BLN+	First-Year	Transfer	Surveys Given	Quartile Size
CS1-NPE	544	349 (0.64)	187 (0.34)	455 (0.84)	89 (0.16)	477 (0.88)	58 (0.11)	7	136
CS1-PE	436	340 (0.78)	94 (0.22)	410 (0.94)	26 (0.06)	390 (0.89)	44 (0.10)	8	109
CS2	268	187 (0.70)	79 (0.29)	241 (0.90)	27 (0.10)	198 (0.74)	67 (0.25)	10	67
CompOrg	428	318 (0.74)	109 (0.25)	380 (0.89)	48 (0.11)	365 (0.85)	61 (0.14)	4	107

one question taken directly from an existing instrument, the sense of belonging instrument used in Smith et al. [144].

We trialed our survey in CS1-PE and CS2 in the summer of 2019, both to assess whether it could be completed within about 5 minutes and to solicit other barriers to learning that we might have missed. The final survey is shown below (questions used in this study have been bolded).

1. **In the last week, approximately how many hours did you spend outside of class time working on work for this course?** (*Open text box, numeric answer required*)

2. How challenging was the work for this class this week? (*Likert scale: 1–5 where 1 was labeled as "Not at all challenging" and 5 was labeled as "Extremely challenging"*)

3. In the past week, which of the following, if any, did you seek help from/work with for this course? (*Checkboxes, select one or more*)

    (a) My peer(s) in this class (as pair programming/joint assignment submission)

    (b) My peer(s) in this class (for help completing my own version of the assignment)

    (c) My friends who are not in this class

    (d) Piazza (I posted one or more questions)

    (e) Piazza (I got help from reading responses to others' questions)

    (f) A tutor

    (g) A TA

    (h) The instructor

    (i) I did not seek help from anyone this week

    (j) Other (Please specify):

4. **In the past week, on a scale of 1 to 5, to what degree did each of the following interfere with your ability to learn and complete the work for this course? (*Likert scale for each item: 1–5 where 1 was labeled as "Not at all" and 5 was labeled as "Significantly"*)**

   (a) **Requirements for other classes**

   (b) **Illness**

   (c) **Family obligations**

   (d) **Work obligations**

   (e) **Social/personal life issues**

   (f) **Confusion specifically about the assignment**

   (g) **Confusion generally about the material**

   (h) **Getting stuck on a bug**

   (i) **Inability to get help**

   (j) **Embarrassment/discomfort asking others**

   (k) **Self-doubt/lack of confidence**

   (l) **Lack of interest in the assignment or material**

   (m) **Goofing off/procrastination**

   (n) **Other (Please specify:)**

5. **What is your current overall satisfaction with your performance in this class? (*Likert scale: 1–5 where 1 was labeled as "Extremely dissatisfied" and 5 was labeled as "Extremely Satisfied"*)**

6. **Reflecting on your experiences over the past week, to what extent do you agree with the following statements. (*Likert scale for each item: 1–5 where 1 was labeled as "Not at all" and 5 was labeled as "Completely"*)**

(a) **I feel accepted in this class**

(b) **I feel comfortable in this class**

(c) **I feel supported in this class**

(d) **I feel like I don't belong in this class**

7. **At this time, approximately how many other students in this course would you be comfortable reaching out to study with? (*Open text box, numeric answer required*)**

8. Reflecting on the last week, how stressed have you been overall? (*Likert scale for each item: 1–5 where 1 was labeled as "Not stressed at all" and 5 was labeled as "Extremely stressed"*)

9. Optional: If you want, please enter any information to expand on or explain your answers to any of the questions on this survey. (*Open text box.*)

For the study we present here, we focused on questions (1), (4), (5), (6), and (7), 18 questions in all, as question (4) is comprised of 14 questions. We chose to omit the remaining questions as they did not provide specific detail on why the student would be likely to struggle. Sentiments generated from question (2) and (8) may repeat in other questions such as questions on the assignments, material, or other personal issues. Question (9) is open-ended. Question (3) was not included because we were unclear on how to meaningfully combine this with the rest of our data.

5.1.4 Survey Administration and Data Cleaning

Students were given the survey by their instructor as part of the course's weekly or bi-weekly homework assignments, and they received a small amount of credit for completing it. Instructors maintained administrative control over the surveys in order to ease their use for responding to student issues and making course corrections on the fly. Survey responses were not anonymous to facilitate individual student support from the instructor, though all identifying information was removed before analysis according to our approved human subjects protocol, as described below.

Prior to analysis, all survey responses, as well as course performance data including overall course grade and final exam score, were sent to an external team that deidentified the data and removed any students who were under 18 years old or had opted out of the study. Using registrar data, this team also added demographic data about the students, notably race/ethnicity, gender, and matriculation status. Our university uses a binary gender classification scheme (Male/Female), which we acknowledge limits our analysis. For race/ethnicity, it uses several categories. However, for the privacy of students from groups that are represented in small numbers, the race/ethnicity data we received was aggregated into two groups: (1) White/Asian students and (2) Black, LatinX, Native American, and Pacific Islander (BLN+) students.

We performed a small amount of pre-processing on the data in order to maximize consistency. First, the answers for the sense of belonging questions (6a), (6b), and (6c) were reverse-coded, as the 1-5 range runs from negative to positive sentiments, whereas for all the other questions the 1-5 range runs from positive to negative responses. As a result, a higher number is negative for all these questions.

Second, in order to account for a different number of surveys offered per course, for each question we calculated each student's average response across all surveys they completed. Although using averages with Likert-scale data is controversial, we chose to average each student's responses instead of taking the student's median response for two reasons. First, students responded on a 5-point scale with labels on only the extremal values, making the difference between the levels appear linear. Second, median values would potentially drop important variations between a single student's responses across the quarter. For example, a student who responded with 1, 1, 1, 5, 5 on five surveys would be recorded with a response of 1 if we had used medians, completely losing the struggle the student experienced in the last part of the quarter. There were cases in which a student did not respond to any surveys during the term. These students were removed from any analysis, but were included when providing descriptive statistics of our courses, quartiles, and indirectly, response rates to surveys.

Finally, to associate survey responses with outcomes, we needed a metric to compare student performance across the four courses. We settled on using final exam score, as it is individual work, proctored, and designed to measure knowledge with no weight given to participation or effort. To address variations in difficulty and grading, we normalized the scores across the courses using z-scores.

5.1.5 Exploratory Factor Analysis

The breadth of our survey meant that it was not immediately clear which individual challenges might comprise higher-level themes. Thus, after cleaning the data, we applied Exploratory Factor Analysis (EFA) to discover underlying structure in the data [31]. EFA finds shared variance among the variables and combines them into factors, which may be non-observable. To ensure that EFA could be applied, we conducted both the Bartlett's Sphericity and Kaiser-Meyer-Olkin tests to our data. These tests determine if factor analysis can be applied to the data by testing the overall significance of the correlations within the correlation matrix and testing if relationships between variables was high respectively. Both tests passed, with Bartlett's Sphericity being significant ($\chi^2 = 27706$, p « 0.01) and KMO indicating appropriately strong relationships between variables (KMO = 0.94) [9]. In the EFA analysis itself, we assigned a question to a factor if it had a loading of 0.4 or greater [133]. We also applied an oblique rotation in our EFA process, as we assume that our questions are not independent of each other[9]. To determine the number of factors to which it would be appropriate to fit our data—balancing model parsimony and the captured variance—we included those factors that had an eigenvalue greater than 1. To validate the appropriateness of this cutoff, we examined both the last factor included and the first factor left out, including their eigenvalues and the amount of variance they explained. We also used a scree plot visualization as an additional check of sensibility. We describe the results of our factor analysis in the next section.

5.2 Results

5.2.1 Factors Derived from Exploratory Factor Analysis

Using EFA as described in the previous section, we found that four factors were the right number to best describe our data set. Running EFA with one more factor explained only slightly more variance (2%), and generated a factor composed of one survey question. The survey questions that comprise our factors are shown in Table 5.2, along with their loadings onto their factors. We were able to identify a common theme among the questions that make up each factor and hence named them: Lack of Sense of Belonging (LSoB), In-Class Confusion (ICC), Personal Obligations (PO), and Lack of Confidence (LoC). We then ran Cronbach's alpha for each of our factors, all of which demonstrated good internal consistency [148] and are above the 0.7 threshold which is considered sufficient evidence for internal reliability [53].

Some questions did not load well onto any factors. Three questions are worth discussing. First, the EFA did not capture question (6d) in LSoB. Question 6d had a factor loading of 0.37, below our threshold of 0.4. It's possible that students were inconsistent in their responses to (6d) because the preceding questions (6a)-(6c) were worded for 1 being the worst and 5 being the best, while (6d) was coded the opposite way. Second, questions (1) and (7) were not captured by our model. We are not sure why these questions did not enter into our model, as we had expected them to be important.

The rest of our analysis will be driven by observing the frequency and associations of these factors to student performance in courses and among different demographics in our study population. When using the term "factor" in the remainder of this section, we are referring to one of the factors determined using EFA. When referring to students' reporting on a factor, we mean reporting on one or more questions that our EFA linked to that factor.

Table 5.2. Survey question makeup of each EFA factor, along with factor loadings, variance explained, and Cronbach's Alpha.

Factors	Component Survey Questions	Question Loading	Variance Explained	Cronbach's Alpha
Lack of Sense of Belonging (LSoB)	5 (Performance Satisfaction)	0.54	16%	0.93
	6a (Feel accepted)	1.00		
	6b (Feel comfortable)	0.95		
	6c (Feel supported)	0.96		
In-Class Confusion (ICC)	4f (Confusion on assignment)	0.87	13%	0.89
	4g (Confusion on material)	0.83		
	4h (Getting stuck on a bug)	0.77		
Personal Obligations (PO)	4a (Other Course Requirements)	0.51	11%	0.78
	4b (Illness)	0.64		
	4c (Family obligations)	0.73		
	4d (Work obligations)	0.67		
	4e (Social/Personal life issues)	0.65		
Lack of Confidence (LoC)	4j (Embarrassment)	0.80	10%	0.86
	4k (Lack of Confidence)	0.81		
	4l (Lack of Interest)	0.59		

5.2.2 RQ1: Association between Factors and Outcomes

We explored the relationships of the factors to student performance using three sources of information: the distribution of reported factors among students at different performance levels and to highlight responses among struggling students, the combinations of reported factors at different performance levels, and the survey response rate at different performance levels.

To explore how responses differed among students across different performance levels, we partitioned students into four equal-sized quartiles based on their final exam scores, with Q1 being the top 25% and Q4 being the bottom 25%. (Quartile sizes are listed in Table 5.1).

To compute each student's score for each factor, we averaged the student's responses for the questions comprising the factor (cf. Table 5.2) [31]. For example, a student's score for the Personal Obligations factor is the sum of their (average) responses to questions 4a, 4b, 4c, 4d and 4e. As each question was a Likert score from 1–5, this score has a range from 4–20. For cross-factor comparison and visualization of factor scores, we normalized individual factor scores by dividing by the number of questions composing that factor, creating a shared range of 1–5.

Distributions of Reported Factors

We first explore how responses for each factor vary for students at different performance levels. Figure 5.1 shows the distributions (as violin plots) of scores for each of the four factors, across all courses, separated by quartile. The small white dot in the center of each violin indicates the median of each (normalized) factor score, while the rest of the violin illustrates the distribution. Two trends are clear from Figure 5.1.

First, the distributions of responses across factors are different. Students report the highest scores on ICC and LSoB, slightly lower scores for PO, and lower scores for LoC. Only students in Q4 seem to really struggle with lack of confidence. Using Kruskal-Wallis tests, we determined that overall the differences between factors are statistically significant (see Table 5.3, top row). Post-hoc tests confirmed that all pairwise comparisons are also statistically significant.

The large effect size for this difference between factors appears in the top row of Table 5.3 and was calculated with the eta-squared measure using the H-statistic generated by the Kruskal-Wallis test [152].

Second, for all factors, there is a statistically significant difference in distributions across quartiles (see Table 5.3, second row). In Figure 5.1, both the median factor score and the number of students at the high end of the distribution increases from Q1 to Q4, indicating that students report more struggle with all factors as the quartiles go from high performing to lower performing. Concerningly, some students in Q3 and Q4 still report low scores for ICC, perhaps evidence supporting the Dunning-Kruger effect [64] and showing that students may not be able to accurately self-evaluate their struggles. The effect sizes for "Quartiles" demonstrates the degree by which student struggle varies by performance quartile. For example, ICC has the largest effect size of the factors which is reasonable considering students who are lower performing should be more likely to report higher levels of confusion.

Figures 5.2, 5.3, 5.4, and 5.5 provide violin plots for each of the four factors, broken down by both course and quartile. Generally, we see the same trends that factor median scores increase and the distribution becomes more weighted toward the top as we move from Q1 to Q4, although the severity of the increase varies by factor and by course. For LSoB (Figure 5.2), CS1-PE and CS2 do not completely fit the overall trend. For CS1-PE, Q1–3 are fairly similar with only Q4 showing a marked decrease in sense of belonging. CS2 reports the lowest sense of belonging with even students in Q1 reporting a sense of belonging comparable to Q3 or Q4 for the other courses. Its unclear why the CS2 course is different; potential sources may be the courses notoriously high difficulty and the more traditional lecture format used in that course (relative to the active learning methods used in the other courses).

CS2 and CompOrg also seem to differ from CS1-NPE and CS1-PE in a couple of ways. First, students report higher rates of confusion in these courses overall, which may be an artifact of higher difficulty of these courses (Figure 5.3). Second, for LoC (Figure 5.5) CS2 and CompOrg have higher factor scores even among Q1–3, again perhaps due to the difficulty of these later

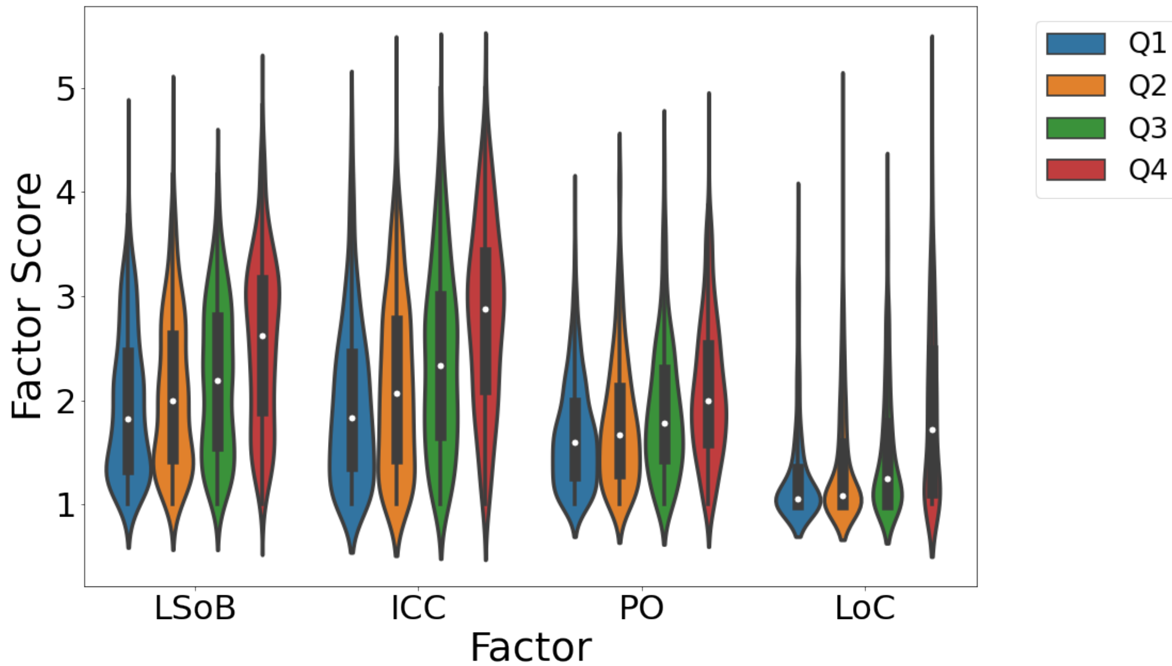


Figure 5.1. Distribution of factor scores across quartiles and factors. The distribution for each factor and quartile span across all courses.

courses in the curriculum. Finally, for PO (Figure 5.4), we see that CompOrg has higher numbers in general, even for students in Q1–Q2. This may be an artifact of students taking more courses, the addition of transfer students (who often start in CS2 or CompOrg) who tend to have more personal obligations as shown later this section, or that high performing students in CompOrg are often hired as instructional staff for the earlier courses.

Reporting of Simultaneous Factors

The previous section showed that higher factor scores are associated with lower performance. In this section, we examine how many students in each quartile are reporting high levels of of struggle with more than one factor. We defined a high level of struggle as having a factor score above the 75th percentile in comparison to the level reported by other students. (We note that we experimented with different thresholds and similar trends were present.)

Figure 5.6 shows the proportions of students from each quartile who report high levels of struggle for each factor. We find that students in the bottom quartile make up a far greater

Table 5.3. Kruskal-Wallis score and effect size across factors, courses and quartiles. A * indicates statistical significance of $\alpha < 0.05$. Interpreting the eta-squared effect size for Kruskal Wallis, 0.01 through < 0.06 is small, 0.06 through < 0.14 is moderate, and anything greater than or equal to 0.14 is large [58].

Group	EFA Factor	Kruskal-Wallis Value	p-val	Effect Size
Factors	-	1163.04	p « 0.01*	0.18
Quartile	LSB	129.28	p « 0.01*	0.08
	ICC	196.63	p « 0.01*	0.12
	PO	99.26	p « 0.01*	0.06
	LoC	165.51	p « 0.01*	0.10

proportion (approaching half in some cases) of the students who report high levels of struggle on each factor compared to students in the other three quartiles. What is missing from this figure (and the previous section) is precisely how these factors might be combining for an individual student. As such, we sought to determine whether students are reporting interference with their performance from multiple factors. If students tend to struggle with just one factor, then perhaps targeted interventions aimed at the indicated factor would help. But if there is a group of students struggling with two or more factors, then more comprehensive interventions may be necessary.

Figure 5.7 shows the proportion of students who were above the 75th percentile in reporting high amounts of four, three, two, one, or zero of the factors. Recall that because we defined a student as experiencing a particular factor as a relative percentage to other students, each factor will be equally represented. We can clearly see that the proportion of students who report a high amount of struggle on zero factors decreases by quartile. Likewise, the number of students reporting high amounts of struggle on all four factors increases notably between Q1 and Q4. There are also consistent increases in the proportion of students reporting two or three factors across the quartiles. We conducted a Pearson's correlation between the number of factors and normalized final exam score across all students and classes and found that number of factors reported as high struggle was negatively correlated with final exam score (r = -0.35, p « 0.01).

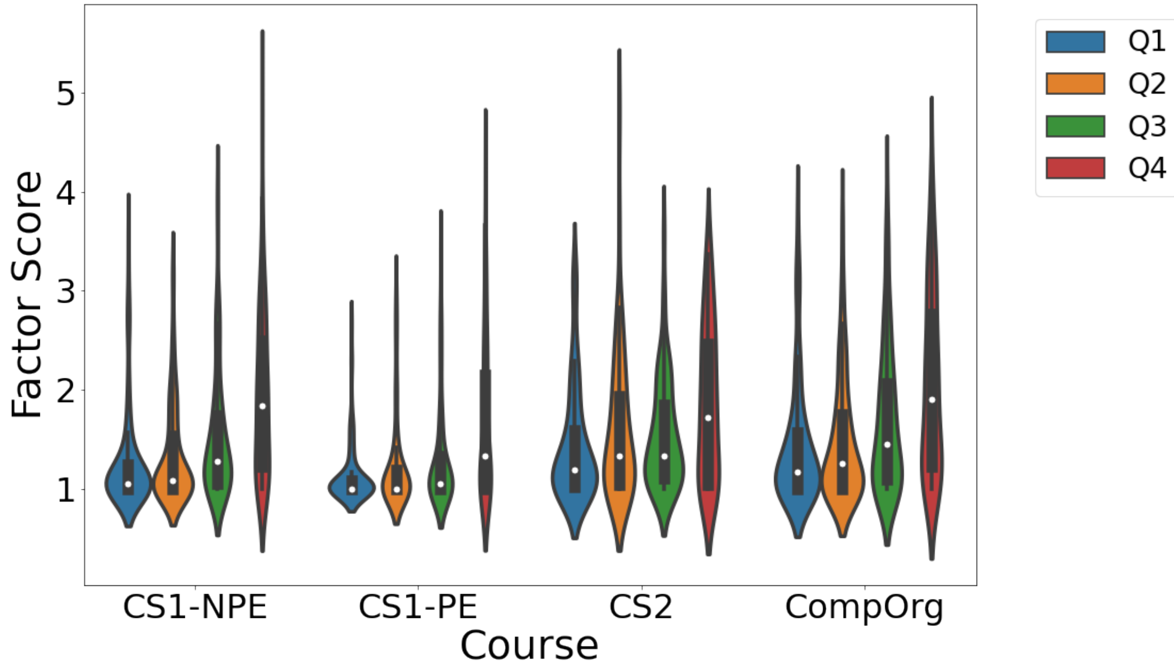


Figure 5.2. Lack of Sense of Belonging (LSoB) factor score distribution by course and quartile.

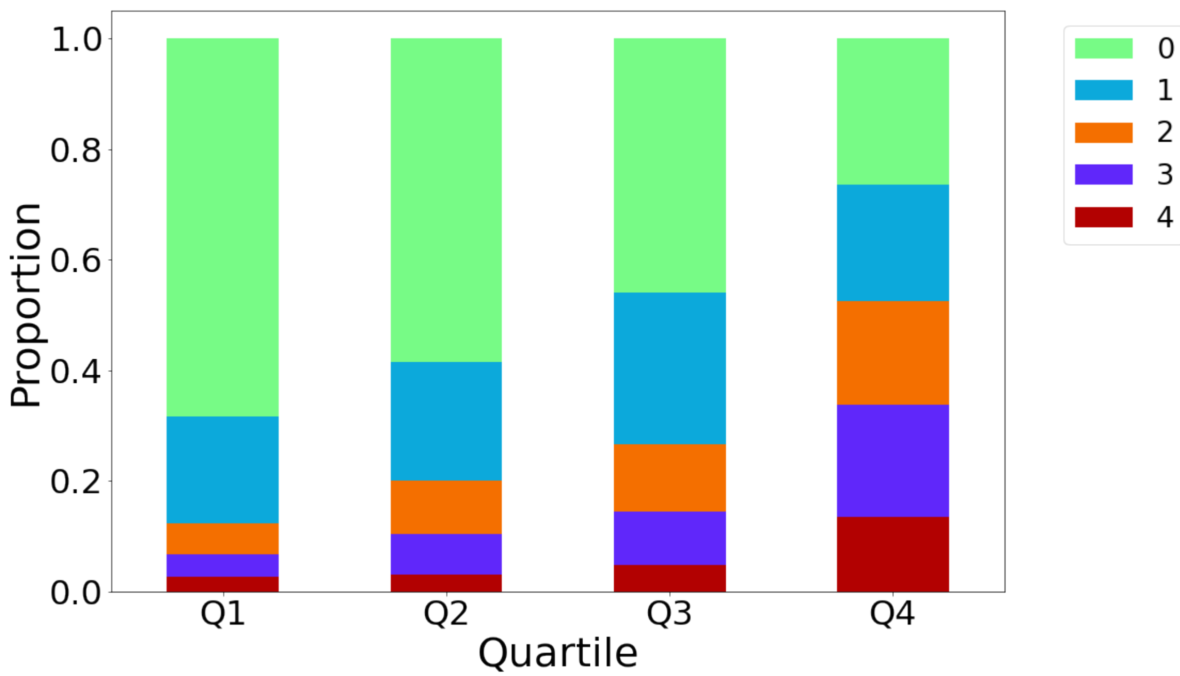


Figure 5.7. Proportion of factors students self-report above the 75th percentile per quartile. The data per quartile covers all courses.

Differences between CS1-NPE and CS1-PE. We next examined whether and how combinations

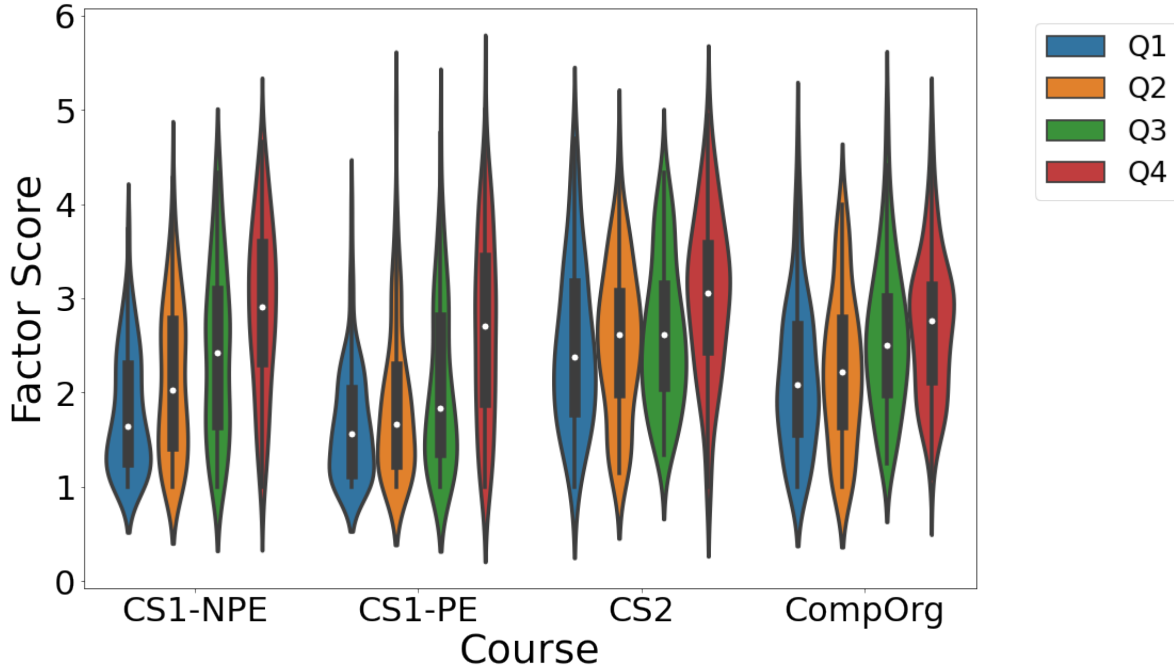


Figure 5.3. In-Class Confusion (ICC) factor score distribution by course and quartile.

of high degree of struggle with factors varied between courses. We found that for most courses the patterns were similar, except there is a notable difference in the results between CS1-NPE and CS1-PE. Figure 5.8 shows the factor breakdown by quartiles in CS1-NPE and CS1-PE, respectively. When examining the proportions of students reporting high degrees of struggle for a single course, recall that the 75th percentile threshold was set for all courses so one course may have more students above that threshold than another.

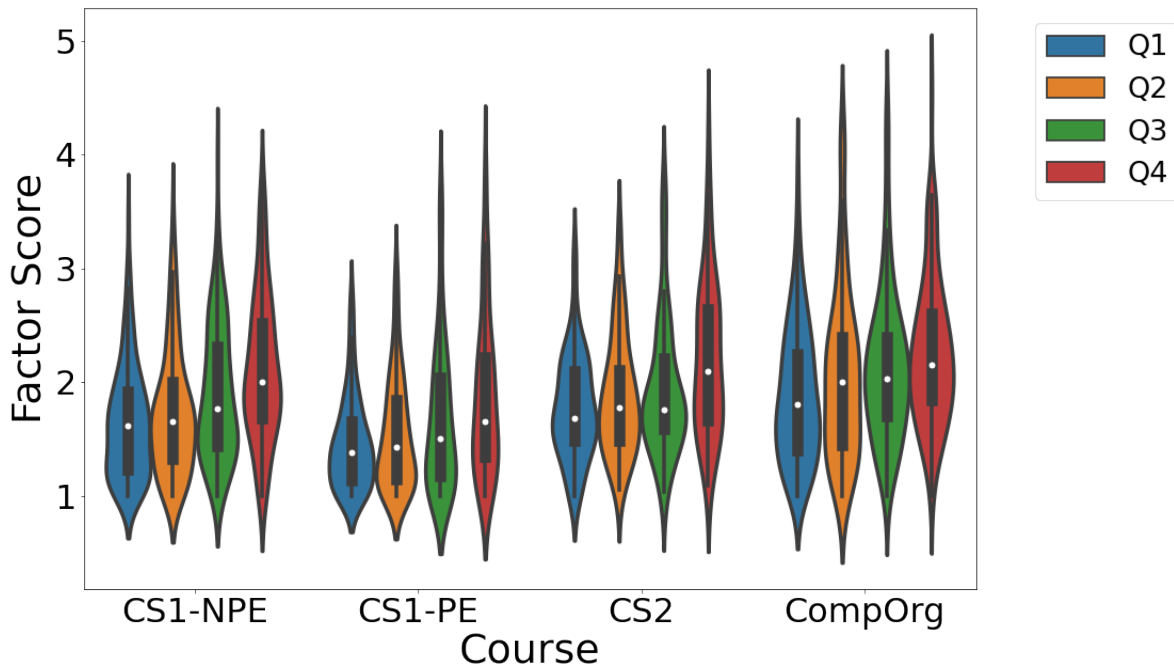


Figure 5.4. Personal Obligations (PO) factor score distribution by course and quartile.

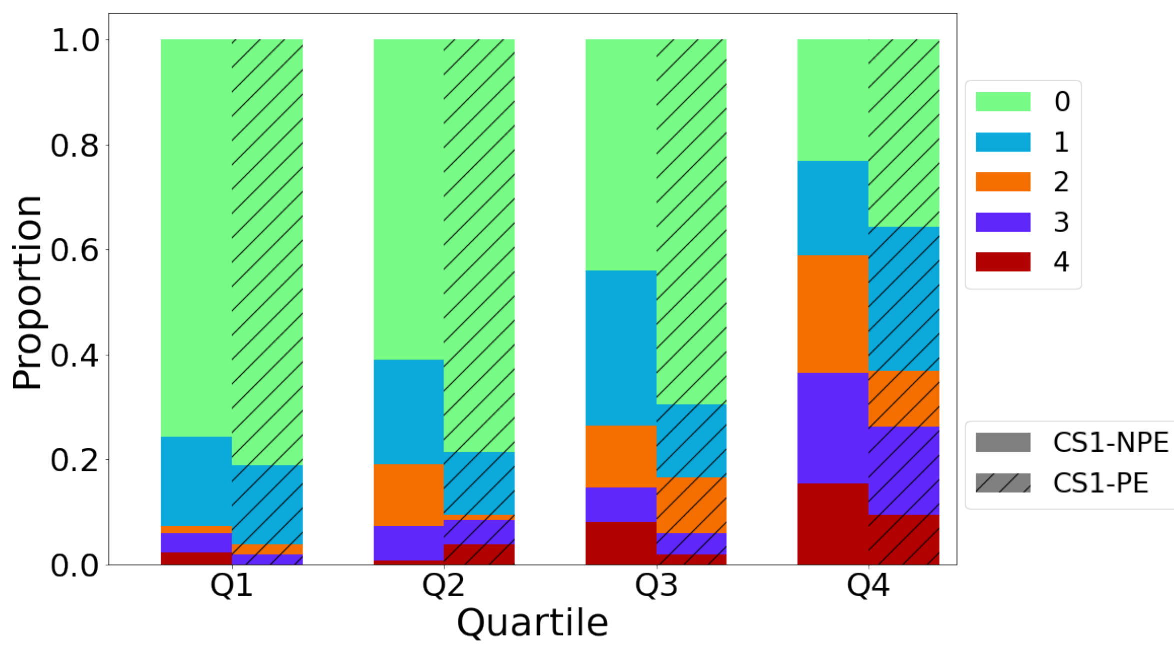


Figure 5.8. Proportion of factors students self-report above the 75th percentile, per quartile, in CS1-NPE and CS1-PE.

For all quartiles, CS1-NPE has a higher proportion of students with a high level of

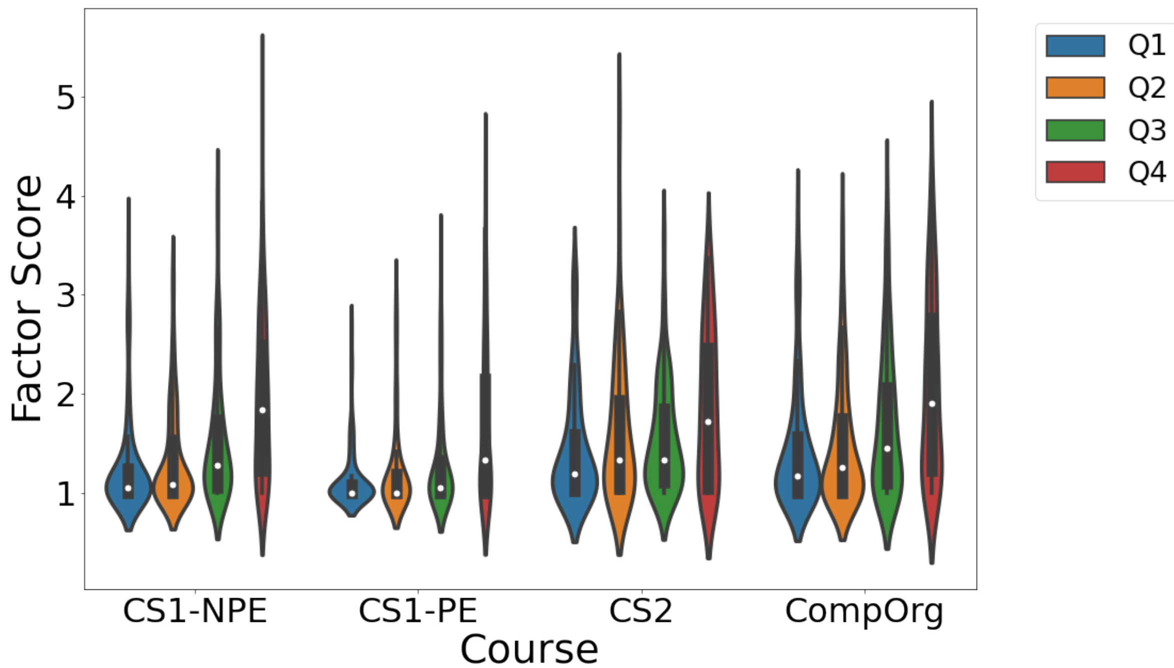


Figure 5.5. Lack of Confidence (LoC) factor score distribution by course and quartile.

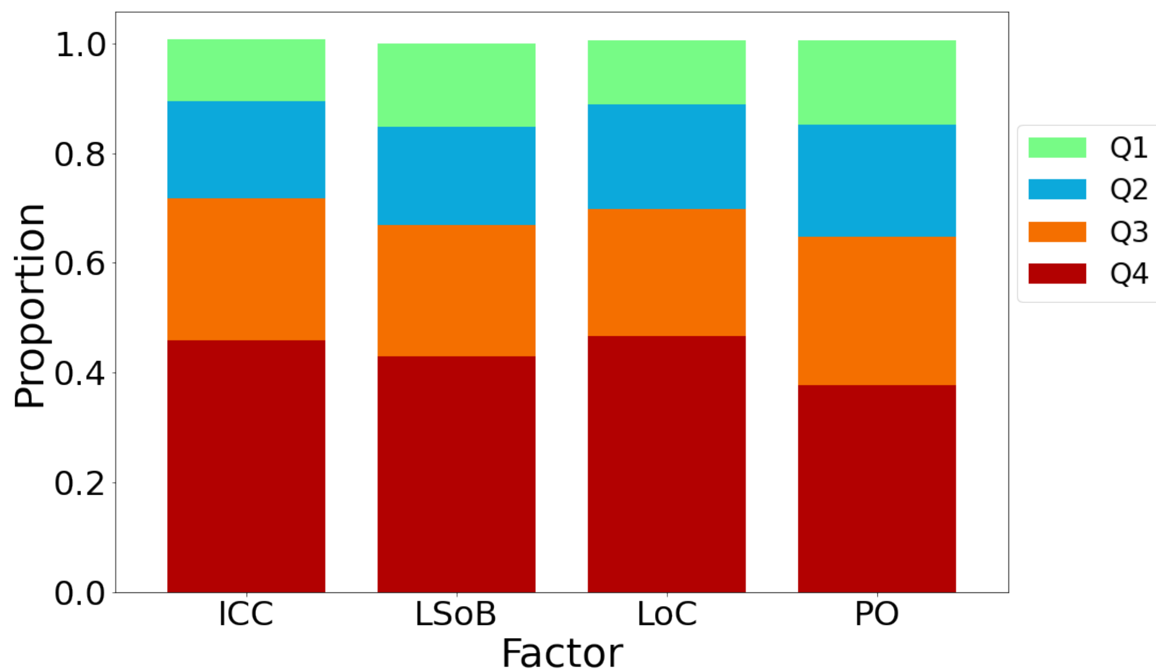


Figure 5.6. Proportion of students who report above the 75th percentile for each factor. The stacked bars represent quartiles 1 through 4.

at least one factor. Surprisingly, students in CS1-NPE in Q1 and Q2 are reporting far more struggles than CS1-PE students in those same quartiles. These findings highlight the value of a CS1-NPE that offers a more "gentle" introduction to computing, as the CS1-NPE population faces more challenges in their computing course relative to those in CS1-PE. One reason for these differences might be due to CS1-NPE attracting more students who are not majoring in, nor intending to major in, computer science. The population of CS1-NPE can also skew toward students later in their studies so they may have more personal commitments relative to students earlier in their studies. It is also possible that these differences are due to (as we will see in the next section) the more diverse population of students who take CS1-NPE. In any event, the goal of the CS1-NPE course is to be sensitive to these issues and is perhaps not fully realizing that goal.

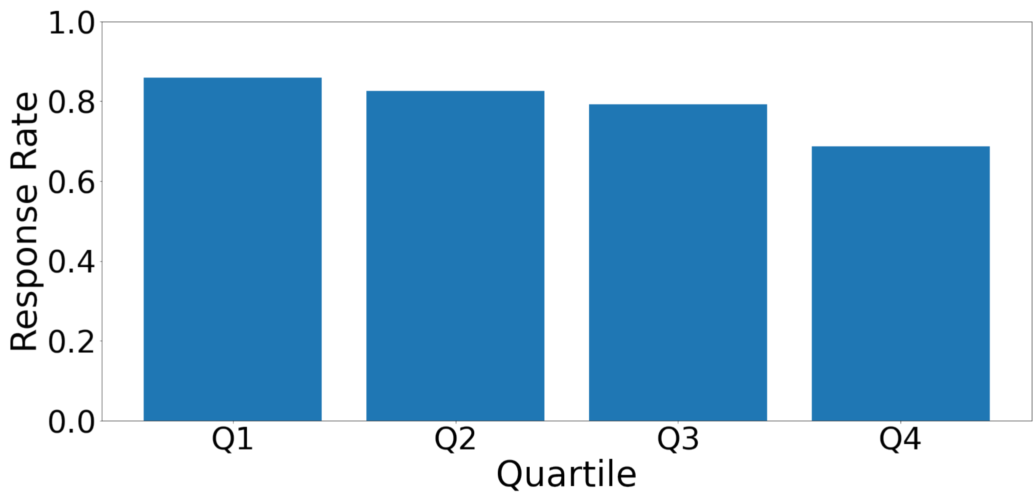


Figure 5.9. Survey response rates across all courses by quartile. Non-responders: Q1 = 3, Q2 = 2, Q3 = 8, Q4 = 27.

Response Rates

After examining some of student responses, we became concerned that survey response rates might not be uniform over quartiles, thus skewing the above results. Figure 5.9 presents the response rates, per quartile, across all courses. Consistently, lower performers completed fewer surveys, with students comprising the first quartile completing 86% of surveys and those in the fourth quartile completing 69%. Students who responded to zero surveys followed a similar, but somewhat stronger pattern across all courses: 3 in Q1, 2 in Q2, 8 in Q3, and 27 in Q4. The trends are the same at the course level. We note this both because: (1) this pattern of response rates could itself be a signal of distress (e.g., lack of time) and (2) underreporting for students in the lower-performing quartiles suggests that the results presented above may be conservative, as students who do not feel they have enough time to complete a survey might report high levels of these factors interfering with their performance.

Take-Aways for RQ1: Association between Stress Factors and Outcomes

Overall, we see that students are struggling on many factors, at times even simultaneously, and these show a consistent pattern with respect to lower outcomes. Students in our CS1-NPE experience higher levels of outside stress than those in CS1-PE, supporting the need for such a course that can be tailored to this population. Interventions on a specific factor may help some students, but many of the students who need the most support, those in Q3 and Q4, may require interventions in multiple areas that span both inside and outside the classroom.

5.2.3 RQ2: Association between Stress Factors and Outcomes, by Demographic

As cited in Section **??**, there is reason to believe that there are disparities in the challenges encountered by minority demographics. This section performs the same analysis as for RQ1, but factored by majority/minority group demographics. For demographic breakdowns across courses and quartiles, please refer to Figures 5.10, 5.14, and 5.18 in each group's respective section. Unfortunately, these figures show that for most courses women, BLN+, and transfer students are underrepresented in the higher quartiles and over-represented in the lower quartiles compared to men, non-BLN+ students and first-year students, respectively. The analysis in this section may explain some of these differences.

Gender

As shown in Figure 5.11, compared to men, women report a slightly higher incidence of most factors. Figure 5.12 reveals that women also consistently report experiencing interference from multiple factors, especially in Q2 and Q4. The differences are statistically significant as seen from our results in Table 5.4. Response rates for men and women, as shown in Figure 5.13, follow a similar pattern as seen for the whole cohort, with women overall responding at a slightly higher rate.

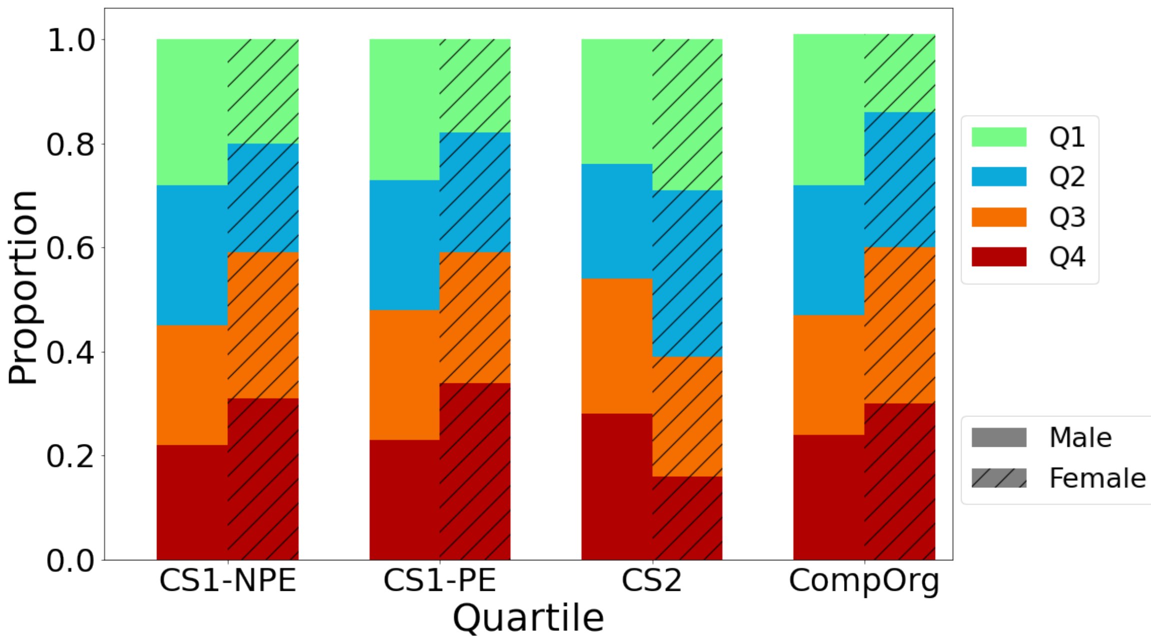


Figure 5.10. Proportion of male and female students in each quartile across courses.

Race/Ethnicity

Differences between BLN+ vs. non-BLN+ students for the four factors overall (Figure 5.15) are small and only significant for PO and ICC (Table 5.4). The stacked bar charts seen in Figure 5.16 reveal that BLN+ students report high degrees of struggle with the different factors in similar proportions to non-BLN+ students, with a few notable exceptions. First, over the first three quartiles BLN+ students report a lower incidence of multiple factors than non-BLN+ students, but in Q4 double the proportion of BLN+ students report a high degree of interference

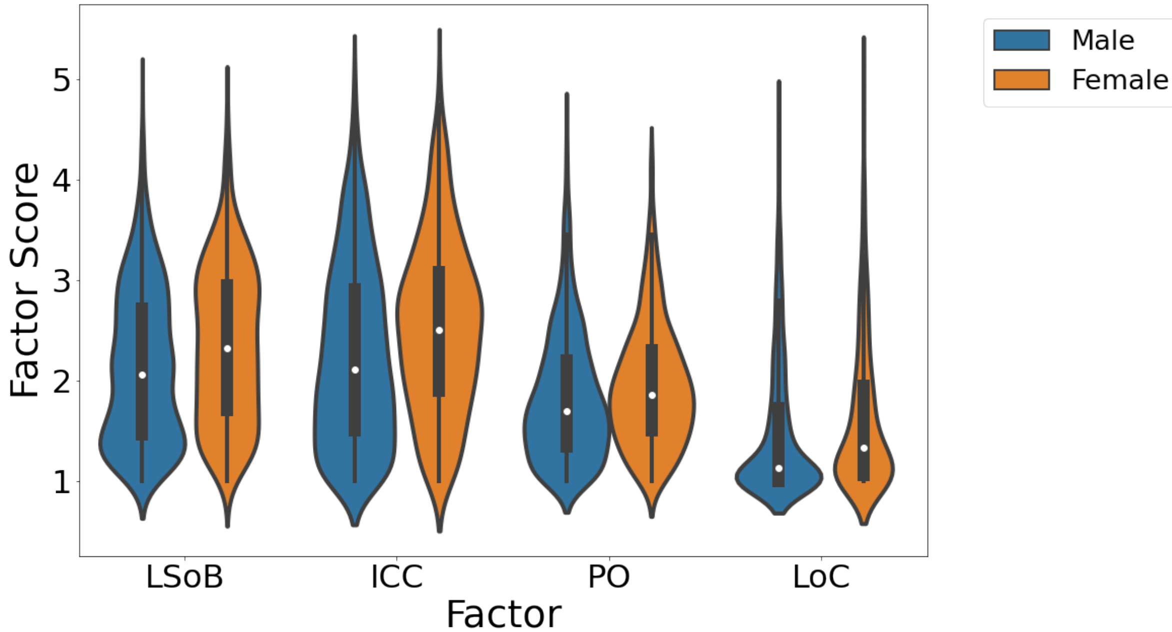


Figure 5.11. Distribution of factor scores across gender. The distributions for each factor and quartile span across all courses.

from all four factors. This suggests that the educational structures at our institution may be doing a better job of helping non-BLN+ students cope when everything is going wrong, compared to BLN+ students, which could partially explain the high representation of BLN+ students in Q4 in most courses (Figure 5.14). Based on response rates in Figure 5.17, we see the same trend of lower quartiles responding at lower rates for both BLN+ and non-BLN+ students. Apart from Q4, however, the response rates between BLN+ and non-BLN+ students are very similar.

Transfer Status

A significant minority of students arrive at our university through the transfer pathway, with the majority of students matriculating directly out of high school. We wished to study transfer students as they are generally older and perhaps have a family to support, and so may struggle from different factors than an average student who enters our university directly from high-school. Figure 5.19 shows that transfer students experience higher levels of all factors than first-years. Higher levels of interference from personal obligations is not surprising given the

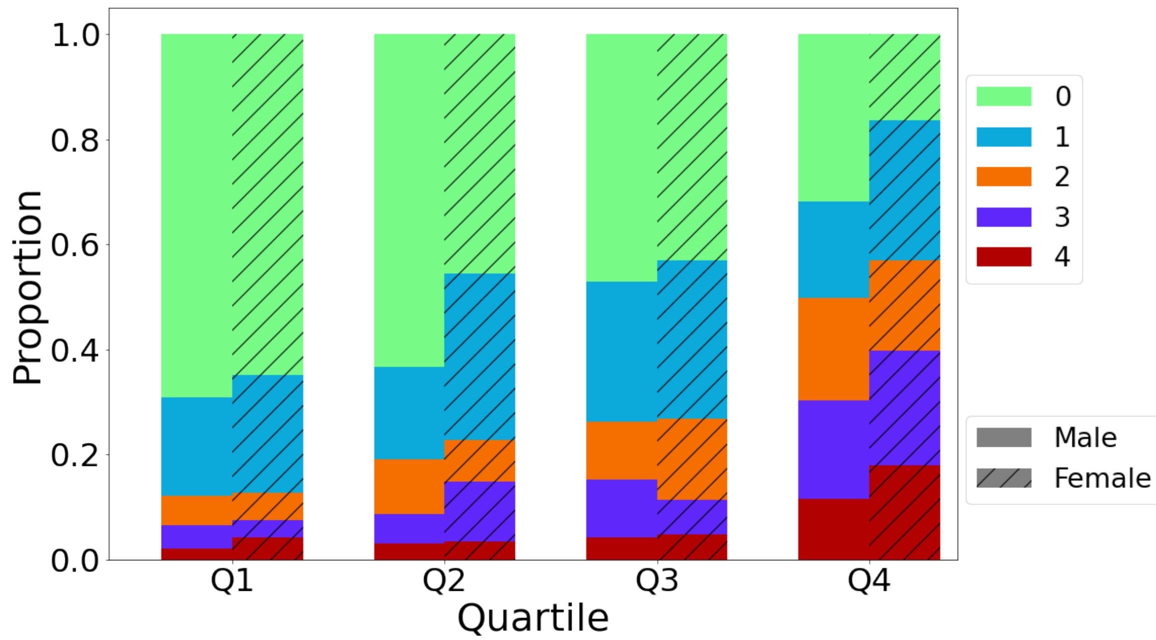


Figure 5.12. Proportion of number of factors students self-report above the 75th percentile, by gender across all courses.

aforementioned difference between transfer students and first-year students. In terms of ICC, more experienced students may be more capable of recognizing areas of confusion or they may not be as well prepared for the class as their prior classes may not align perfectly with what is taught at our institution. For sense of belonging, they may feel the stigma of the age difference (e.g., being "behind") compared to their majority first-year peers. From Figure 5.20 we can see that a larger proportion of transfer students report a high degree of interference for multiple factors than their first-year peers. In Q4, transfers report over 50% more multiple factors than their first-year counterparts, suggesting that the educational structures at our institution may be better serving their majority counterparts. In terms of response rates (Figure 5.21), we see a similar trend as with other groups.

Kruskal Wallis Tests

We ran Kruskal-Wallis tests and effect size across gender, race/ethnicity, and matriculation status. Results can be found in Table 5.4. For the majority of tests we found statistical significance

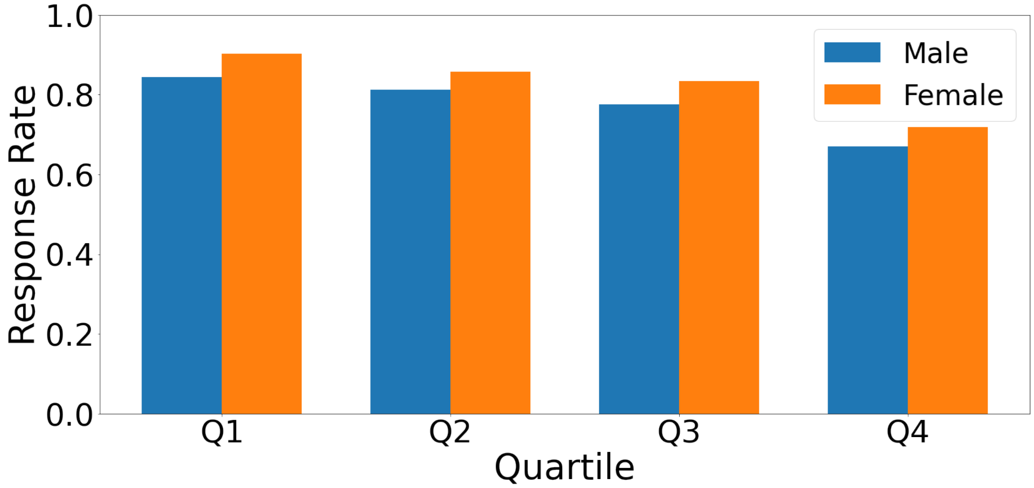


Figure 5.13. Survey response rates across all courses by gender. Male non-responders: Q1 = 3, Q2 = 2, Q3 = 5, Q4 = 20. Female non-responders: Q1 = 0, Q2 = 0, Q3 = 3, Q4 = 7.

between the groups based on their factor average. We see that comparing groups based on race/ethnicity was not significant for sense of belonging, which is interesting since prior work focuses heavily on how these groups suffer from lack of belonging in computing [131]. For those factors where the findings are significant, there are small effect sizes for gender, BLN+, and transfer students. These small effect sizes seem to indicate that the differences between the subgroups within each group is small for each individual factor. These results show us that struggles our students experience may not be substantially increased or mitigated based on their gender, race/ethnicity, or matriculation status.

5.3 Discussion

5.4 Limitations to Generalizability and Threats to Validity

5.4.1 Limits due to Studying Four Courses, in a Single Term, at a Single Institution

The surveys analyzed for this study were given in four courses during one term at one research-intensive university. The reported results may not generalize to all institutions or other courses in the major. Still, the courses surveyed are critical in recruitment and retention, and

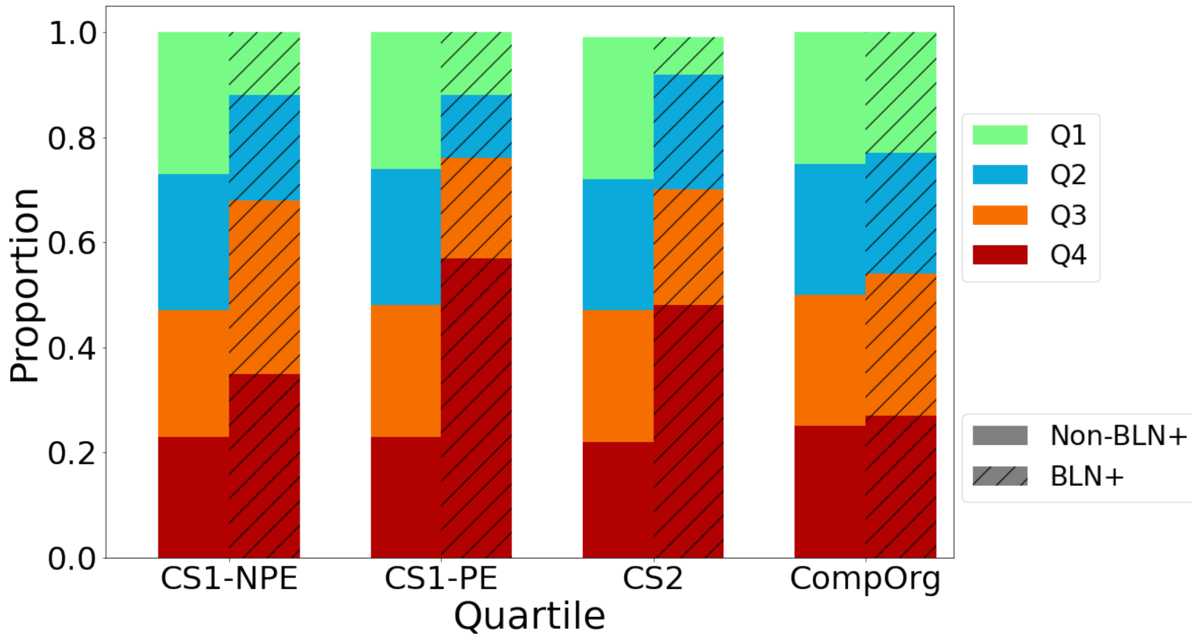


Figure 5.14. Proportion of non-BLN+ and BLN+ students in each quartile across courses.

for those who persist, play an important role in future student success as shown by recent work evaluating the role of prerequisites for courses similar to those in this study [154, 61]. Likewise, research-intensive universities graduate the majority of computing students in North America. Finally, the Fall term can be expected to present unique challenges for the many students matriculating in that term. However, that first experience is crucial for recruitment and persistence of students in computing [128].

5.4.2 Threats due to Student Self-Evaluation

Since students are required to self-evaluate to fill out the surveys, there is the potential that students may not be accurately reporting their level of stress on certain factors. For example, many students performing near the bottom of the course did not report serious issues for ICC. In terms of validity, the concern here is that students in different quartiles may exhibit different levels of accuracy.

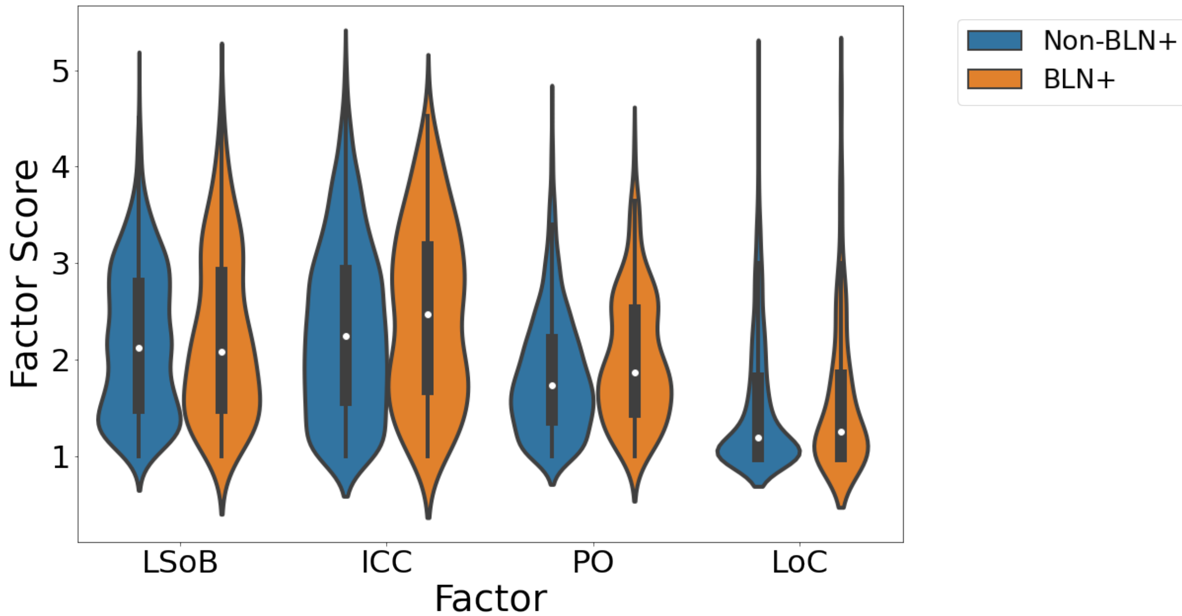


Figure 5.15. Distribution of factor scores across race. The distributions for each factor and quartile span across all courses.

5.5 Discussion

Our central research question was to examine the social-emotional /social-cognitive, structural and/or personal factors that are related to student struggle, with the goal of gaining a more holistic view of pressures that are interfering with students' learning. We found that these pressures, and students' lives, are indeed complex. Not only do different students face different struggles, but that particularly lower-performing students are struggling across multiple factors. We find that gender, race, and matriculation status are related to these factors with students from underrepresented groups generally experiencing more sources of stress. Our study confirms results from previous studies as well as offers a new perspective on student struggle from a broad scale across several lower-division and mid-level courses. These results have implications both for researchers and for instructors and departments to build educational structures that allow all students to thrive.

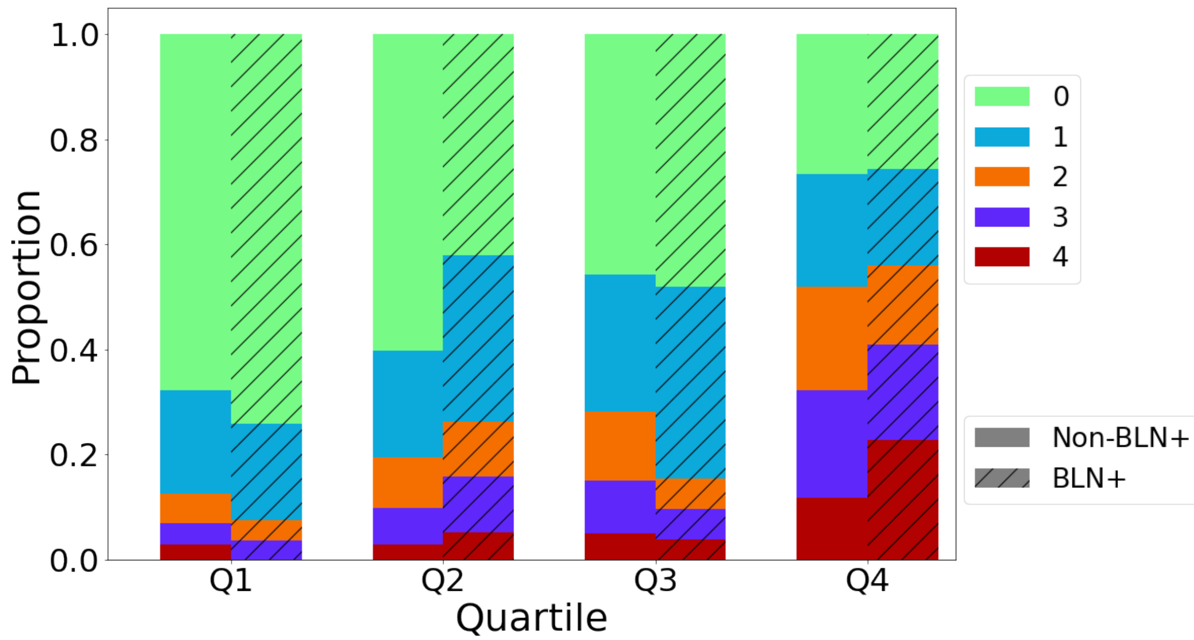


Figure 5.16. Proportion of number of factors students self-report above the 75th percentile, by race across all courses.

5.5.1 Relationship to Previous Work

We found four factors that were centrally related to students' struggle across lower-division computer science courses; two of these factors (Lack of Confidence, and Lack of Sense of Belonging) were related to students' social-emotional experience. Consistent with previous work, a lack of sense of belonging was an important factor related to struggle. However, unlike Veilleux et al., who found that sense of belonging was related to students' perception of their grades but not their actual grades [155], we found that sense of belonging was indeed related to students' actual grades. Similar to Sax et al., we found that women report slightly lower sense of belonging than men. On the other hand, while Sax et al. found that BLN+ students reported a slightly higher sense of belonging than non-BLN+ students, we found that sense of belonging levels in BLN+ and non-BLN+ students were similar. We also found that transfer students—a population not often specifically studied in CS courses—have a lower sense of belonging than non-transfer students. Based on situated cognition, these students may have difficulty feeling like full members of the community which may limit their ability to master the knowledge and

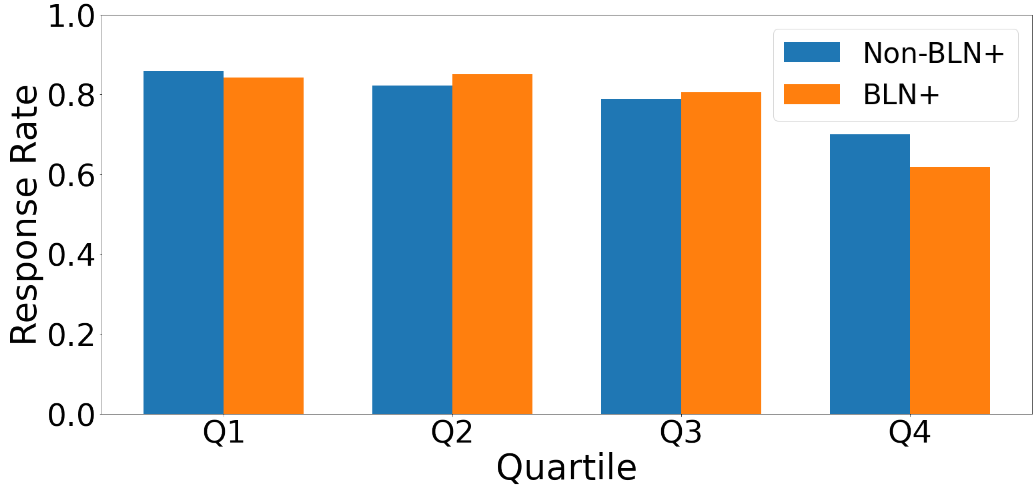


Figure 5.17. Survey response rates across all courses by students' race/ethnicity. Non-BLN+ non-responders: Q1 = 3, Q2 = 2, Q3 = 7, Q4 = 21. BLN+ non-responders: Q1 = 0, Q2 = 0, Q3 = 1, Q4 = 6.

skill of that community.

The differences in confidence among low and high performing students and by gender are consistent with prior findings in computing that self-efficacy is associated with student success [79] and that women tend to have lower self-efficacy [12, 15]. Those students with lower self-efficacy may be struggling with higher levels of anxiety and have lower interest in the field. Our study adds to this previous literature by showing that lack of confidence persists beyond CS1, into mid-level CS courses. Perhaps most concerning are those students in the top quartiles of the class (particularly in CS2 and CompOrg) who express these low levels of confidence as this may be an indication of poor metacognition or that these students are experiencing imposter phenomenon [124].

A third important factor that emerged in our study was Personal Obligations. As discussed in Section **??**, we are not aware of any studies that relate outside commitments to success in CS courses, yet it is known that in general commitments outside schoolwork may impact dropout rates [147] and may require students to develop different strategies to succeed [80]. Their ability to manage these outside commitments may impact their mental health [145]. Our results show not only that CS students who struggle in class report higher instances of interference

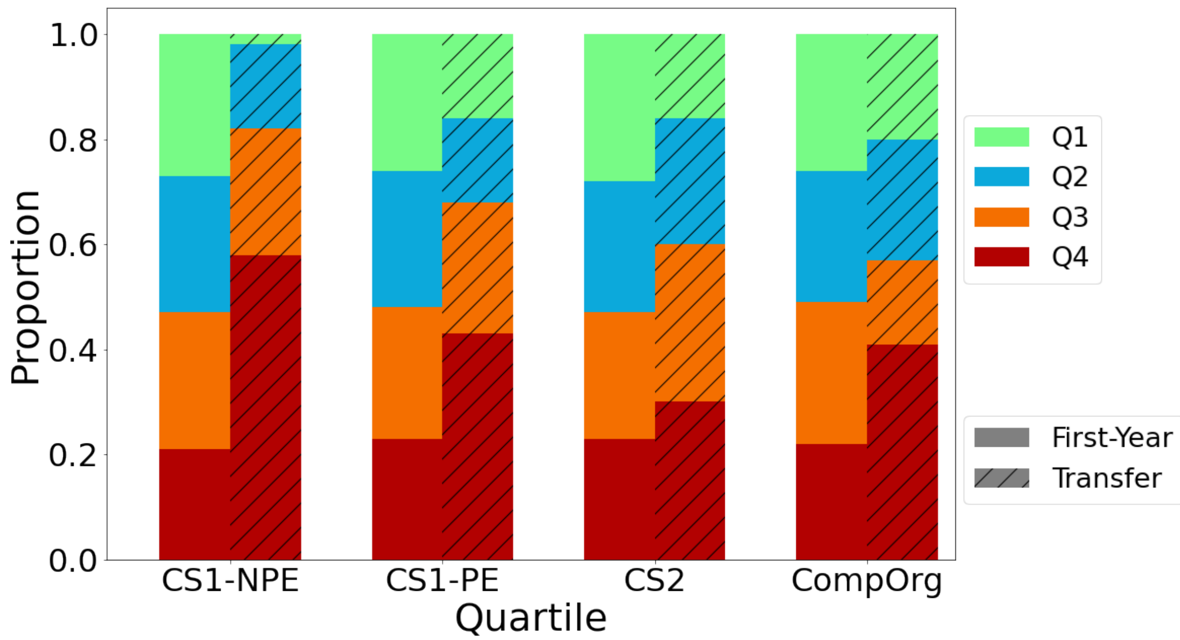


Figure 5.18. Proportion of first-year and transfer students in each quartile across courses.

from outside sources such as work and other classes, but also that these struggles are often compounded or impacted by other struggles they face. Indeed, most lower-performing students who report struggling with personal obligations also report struggling with confusion, lack of confidence, lack of belonging, or more than one of these additional factors. As we find that BLN+ students disproportionately encounter more personal obligations (albeit with a small effect size) and one can expect these challenges to be more common for students with lower socioeconomic status [163], helping students address these challenges may be essential for improving the diversity of the field.

5.5.2 Future Research Directions

Understanding Interactions Between Factors: Our work has begun to reveal a complex web of struggle that students face in early CS courses. Our EFA found four separate factors, but the factors are all significantly correlated with one another: Spearman's rho varies between 0.36 and 0.67 for all of the pairwise factor comparisons with LSoB and PO having the lowest correlation and ICC and LoC having the highest. Better understanding the relationship between these factors,

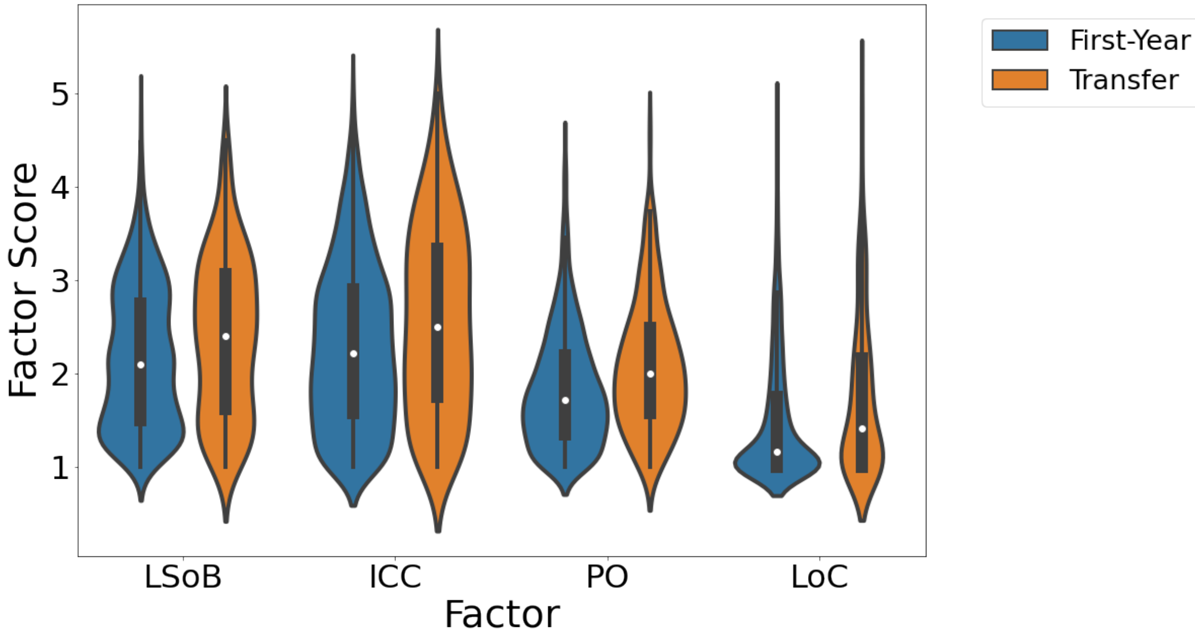


Figure 5.19. Distribution of factor scores across matriculation status. The distributions for each factor and quartile span across all courses.

as well as considering additional sources of struggle, is critical to building structures where all students can succeed. For example, when a student reports confusion with the material, and also report a low sense of belonging and high personal obligations does this mean that the student's confusion causing them to have a lower sense of belonging? Or are the student's personal obligations keeping them from feeling that they belong, which in turn is making it harder for them to learn? These different scenarios could imply very different intervention techniques.

Closing Demographic Gaps: Although it was not the central focus of our study, our analysis revealed significant (and sometimes large) exam performance differences between men and women, non-BLN+ and BLN+ students, and transfer and first-year students. It is possible (even likely) that these differences are partially due to differences in prior experience; a recent study found that prior experience is correlated with higher grades even into the upper division [3]. Still, the magnitude and consistency of these differences is troubling. The results presented here on student struggle give us new insight into the broader barriers that women, BLN+ students, and transfer students are facing which might allow us to better address the uneven playing field for

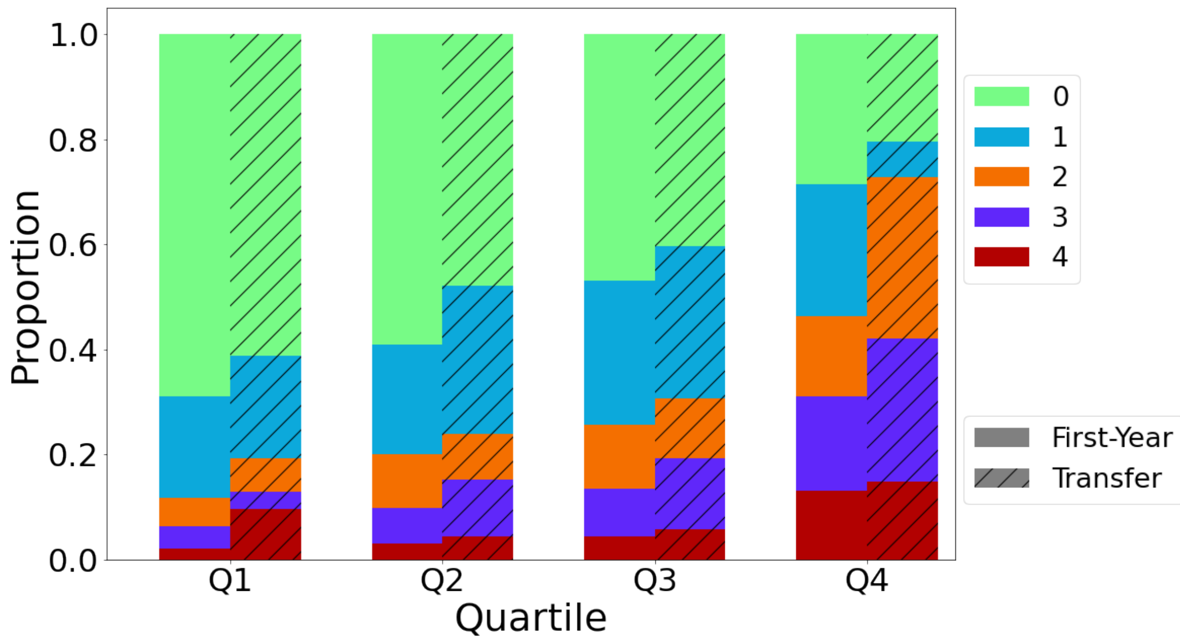


Figure 5.20. Proportion of number of factors students self-report above the 75th percentile, by matriculation status across all courses.

these students. Yet, more research is needed specifically to understand effective ways of closing these performance gaps.

Ensuring We Hear All Students: Our results on disproportionate response rates from students in different performance quartiles indicate that we are not hearing from precisely the the students that we most want to understand. We believe that this pattern is not unique to our study, but is likely to be consistent for most survey-based research. Developing mechanisms that focus on response rates from the most affected groups is important to the validity of this approach. Furthermore, combining this work with deeper and more individualistic approaches such as critical theory approaches [54] will help ensure that the voices of those we most want to learn from are not left out.

5.5.3 Student Profiles

Part of our survey allowed students to provide comments elaborating on anything on the survey. We wanted to highlight some comments presented by students that may provide context

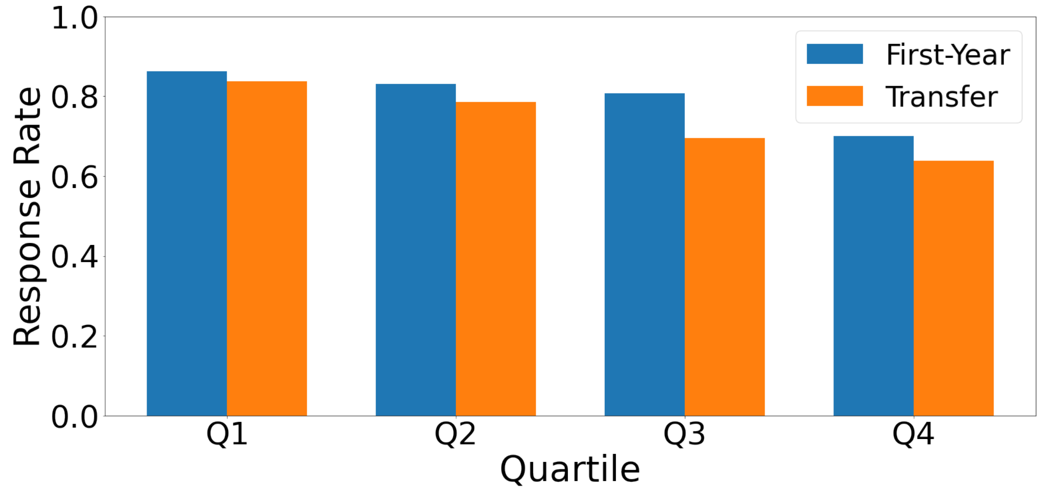


Figure 5.21. Survey response rates across all courses by whether a student is a first-year or transfer. First-year non-responders: Q1 = 3, Q2 = 2, Q3 = 3, Q4 = 16. Transfer non-responders: Q1 = 0, Q2 = 0, Q3 = 3, Q4 = 10.

to some of our findings.

One Q4 student in CS2 whose survey responses fell below our 75th percentile threshold for all four factors still mentioned that their stress, "mainly comes from commuting and how my schedule is set up" where they claim they arrive to school at 5am and do not leave until 5pm. They also mention getting difficulties getting help from instructional staff due to long wait times. For this student, it appears they are under stress and they recognized they needed help, but they perhaps under-reported their challenges or other students were facing more challenges.

Another student in CS2, but in Q1, reported high degrees of stress, enough to exceed the 75th percentile for all four factors, mentions the difficulty getting help on weekends. They then express that due to this issue, homework must be turned in, "full of bugs or spend a long while on it (with no social life)" to get full credit. The student also makes mention that they struggled due to lack of friends in the class, "my friends are all out either not in the course or done with the class" suggesting they felt working with others might have helped. For this student, these stresses did not impact their final exam performance indicating either they were able to overcome these challenges or were overstating their challenges.

Comments provided by students in Q4 whose survey responses resulted in them exceeding

Table 5.4. Kruskal Wallis test results for each factor across different groups. A * indicates statistical significance of $\alpha < 0.05$. Interpreting the eta-squared effect size for Kruskal Wallis, 0.01 through < 0.06 is small, 0.06 through < 0.14 is moderate, and anything greater than or equal to 0.14 is large [58].

Group	EFA Factor	Kruskal-Wallis Value	p-val	Effect Size
Gender	LSoB	26.84	p « 0.01*	0.02
	ICC	36.16	p « 0.01*	0.02
	PO	15.78	p « 0.01*	0.01
	LoC	32.85	p « 0.01*	0.02
Race/Ethnicity	LSB	0.29	0.59	0.00
	ICC	4.51	0.03*	0.00
	PO	10.09	p « 0.01*	0.01
	LoC	0.93	0.34	0.00
Matriculation Status	LSB	13.86	p « 0.01*	0.01
	ICC	13.94	p « 0.01*	0.01
	PO	27.36	p « 0.01*	0.02
	LoC	20.25	p « 0.01*	0.01

the 75th percentile for all four factors had both commonalities and showed the uniqueness of each students' experience. Common issues included assignment difficulty, stress from exams, stress from other courses and lack of time. One student in CS1-NPE mentions that their stress is "not because of the class specifically, but because its the first full week and I'm still adjusting." For this student, it seems it is the adjustment to college life as a new first-year student was overwhelming and difficult. Another student in CS1-NPE mentions that although they are "finding it difficult, ...feel [themselves] slowly falling in love with programming and problem solving." Although struggling, this student outputs a sense of optimism about their situation. These responses, combined with other provided by students, remind us that each student has their own unique experience in our courses and experience (and report) their stress in different ways.

5.5.4 Educators' Call to Action

Our results clearly demonstrate that students are struggling across multiple dimensions of their lives. As such, targeted interventions to address one factor, such as lack of sense of belonging or in-class confusion, may not be sufficient. Moreover, finding that a student is seriously struggling across multiple dimensions may be a good indicator that they will struggle to succeed in the course.

Furthermore, one of the main factors that emerged that may interfere with students' success, personal obligations, points to complex systemic issues that must be addressed beyond the design of a particular course. Departments, programs, and universities must consider whether the rules governing our programs are reasonable for all students, given the external pressures many students face. As just one example, students at our institution who are receiving financial aid are required to be enrolled in a minimum number of units—a full course load—which may be too much some quarters in view of their personal obligations and the way that our courses are currently designed. Given the presence of these constraints on students, there is a growing mandate to reflect on our current processes and make improvements to help our most at-risk students.

Finally, one of our notable findings was the substantial difference in response rates between high performing and low performing students. We have no way to recover the experience of students who did not respond to the surveys, but we can guess that the non-responders were experiencing even more barriers to learning than those who did respond. These surveys are meant to show us what students struggle with and if those most struggling are not responding, identifying how to help them becomes more difficult. This differential response rate is important to acknowledge in any survey-based study, and finding methods to ensure that these students are not excluded from this type of research is an important area of future work.

5.6 Summary

Through this study we have provided evidence that students in early computer science courses are struggling across different categories of stress: personal obligations, lack of belonging, lack of confidence, and in-class confusion. The lowest performing students are most likely to report struggling with the all four of these categories. The categories span from inside to outside the classroom, and suggest that interventions to help students better understand CS concepts or increase their sense of belonging must be paired with larger structural changes to address the barriers students face in their broader lives. Instructors can thus help by designing courses that will help support students who face a myriad of barriers. Lastly, our work adds support to the notion that it is time to start reflecting on institutional shortcomings and start addressing compounded struggles rather than individual ones.

5.7 Acknowledgements

This work was supported in part by NSFaward #1712508, as well as a UCSD Sloan Scholar Fellowship, a UCSD STARS Fellowship, and a Gates Millennium Scholarship. Chapter 5, in part, is a reprint of the material as it appears in the Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”. The dissertation author is the primary investigator and author of this paper.

Chapter 6

[Pedagogy] The Role and Impacts of Pedagogical Innovation in CS1

The primary goal of this study is to determine the long-term impact of a redesign of a CS1 course at UC San Diego in 2008. For context, those in the computer science field are well aware that the number of computer science majors has risen and declined over time. The rise of the dot-coms at the end of the 1990s correlated with a surge in interest in computer science and related fields, with the number of bachelor's degrees in computing peaking in the United States in 2003 [99]. Following that peak, the number of students interested in computer science suffered a precipitous decline. To respond to that decline, computer science programs began to focus on how best to attract and retain students into the computer science discipline. Within this context, our institution began a concerted effort to overhaul our CS1 course for students without prior experience in order to reduce failure rates, improve retention of students in computing majors, and to attract new students to the major [112].

The faculty driving the course revision set forth a mandate to adopt best practices from the computing education and science education communities. After reviewing possible practices, the course was redesigned to include three practices recognized within these communities. The first was to ensure computing was being taught in a context meaningful to students through the use of Media Computation [44]. The second sought to actively engage students in core course concepts during "lecture" through the use of Peer Instruction [25]. The third was to create a community among students by having them engage in Pair Programming [159]. We note that less was known about the efficacy of each of these practices in computing than is known now; Section **??** provides the context for the changes at that time.

Prior work by Porter and Simon reported on the success of this course revision—specifically showing that the course redesign significantly lowered failure rates while also resulting in more students taking courses in the major one year after taking the introductory course [110]. However, because the evaluation of the new course was done only a few years after the course restructuring, there are a number of questions that could not be answered then that are possible to answer now: Did the increase in retention in computing courses 1-year later persist to increase retention at the time of graduation? How well did students from this larger group of majors perform in later courses? Were the changes more beneficial for different demographics of students? And can we explain the improved outcomes simply by the changes over time?

In this work, we have extracted over 17 years of CS1 student data to evaluate the impact of the course redesign on passing rates, retention of majors to the point of graduation, the number of students switching into computing majors, student performance in required upper-division computing courses, and time-to-degree. Using student demographic data, we are able to evaluate how the course redesign impacted male and female students as well as students from under-represented minority groups (URM).[1] Moreover, a fortunate facet of our computer science curriculum is that the major has long had two alternative paths into the major. The first path, for students *without* prior programming experience, is the path with the course redesign. The second path, for students *with* prior programming experience, experienced no significant changes to its curriculum over the same time. By comparing student results for both paths, we are able to explore whether the benefits experienced after the course redesign are better explained by changes over time or by the redesign itself.

The contributions of this comprehensive longitudinal evaluation of incorporating Peer

[1]Similar to the Special Report on "Women, minorities, and persons with disabilities in science and engineering" [100] we define URM as Chicano, Latino, African-American, American Indian, and Alaska Native.

Instruction, Media Computation, and Pair Programming in an introductory computing course include:

- Student passing rates and retention of majors improved significantly after the course redesign.
- There is no evidence that this larger population of students suffered worse outcomes later in terms of grades in subsequent courses or time-to-degree.
- Outcomes for students from underrepresented groups in computing (women, URMs) improved, as did outcomes for students from represented groups.
- The benefits for students are better explained as a result of the course redesign rather than possible changes over time.
- Instructors who taught both before and after the revision experienced similar benefits for their students.

6.1 Study Design

Our study focuses on the following research questions:

- **RQ1**—How do student outcomes compare between the version of the course before and after the redesign to include best practices?
- **RQ2**—Can changes in student outcomes be attributed to possible changes over time?
- **RQ3**—Which groups of students benefited from the introduction of these best practices?

The first question is used to gain an overview of how the introduction of best practices changed student performance and retention in the major. Our second research question was motivated by the possibility that any significant difference in performance and retention could be explained away by changes over time. The third question addresses how the course revision

Table 6.1. Metrics evaluated in this study

Metric	Description
Fail Rate	The percentage of students who received a D, F, or Withdrawal among students enrolled in the course.
Retained	The percentage of students who graduated with a degree in computing after entering CS1-NPE or CS1-PE already majoring in computing.
Switched	The percentage of students who graduated with a degree in computing after entering CS1-NPE or CS1-PE majoring in another discipline than computing (or as undeclared).
Upper-Division GPA	Average grades received (GPA), including counting withdrawals as failures, for students who attempted at least 5 upper-division computer science courses required for the major.
Time-to-Degree	Number of years from starting at our institution until graduation for students who receive a bachelor's in a computing major.

impacts students belonging to underrepresented groups in computing. Our analysis focuses on the two introductory programming courses which are the starts of two separate paths into the CS major at our institution: **CS1-NPE (No Prior Experience)** and **CS1-PE (Prior Experience)**, both of which act as an introduction to fundamental topics and techniques of programming. CS1-NPE is the first in a two-term course that serves as this introduction. The course is designed for students with little to no programming experience and is the course that experienced the redesign to include the three best practices. CS1-PE is designed for students with prior programming experience and leverages that prior experience to teach the same learning goals in a single term, rather than two terms.

We focus on two time periods: academic years 2001–2007 and 2008–2012. Academic years will be referred to by the year in which they begin, as all academic years span across two different calendar years. The first-time period, 2001–2007 (Fall 2001–Spring 2008), refers to the time period before best practices were implemented in CS1-NPE. The second time period, 2008–2012 (Fall 2008–Spring 2013), refers to the time period when best practices were implemented in CS1-NPE. CS1-PE had no significant change in course delivery over the time period between

2001—2012 and, as such, is useful for comparison. We gathered student data from those who enrolled in CS1-NPE and CS1-PE across the twelve-year time period (including their grades and graduation outcomes after the end of that period).

The analysis ends with the Spring of 2013 as our institution began restricting students' ability to major in computing at that time. The restrictions were significant, creating a highly competitive environment for those hoping to major and still turning away many students who were interested in the major. As many of the metrics we evaluate would likely be impacted by this change, particularly for underrepresented students based on recent work by Nguyen and Lewis [102], we end our analysis at the point those changes were made. In addition, we need to allow students time to graduate after they take the CS1 course for a number of our metrics to be accurate (e.g., retention and time-to-degree).

The primary metrics used in our evaluation appear in Table 6.1. Students included in the analysis are those who took CS1-NPE or CS1-PE during the regular school year as the courses are infrequently taught in the summer. Summer terms are accelerated and are typically taught by outside instructors. In addition, students needed to have earned a letter grade or withdrew from the CS1-NPE or CS1-PE course. Students who enrolled initially but dropped before the 4th week drop deadline were not present in the data provided and are hence not part of our analysis.

The data was provided from UC San Diego's educational services office in accordance with Human Subjects approval. The characteristics of the courses and students appear in Table 6.2. One special challenge in the data was how to handle students who failed CS1-PE or CS1-NPE but then later retook either CS1-PE or CS1-NPE. For Fail Rate, the student is counted in each course they received a grade. But all other metrics are tracked based on the final attempt outcomes of each student. As such, we grouped students based on their final attempt of a course. For example, if a student took CS1-NPE and received an "F" in 2007 and then took CS1-PE in 2009 and received a "C", they are considered a CS1-PE student in the 2008-2012 time period for metrics other than Fail Rate.

Once the student data was filtered, we ensured that the data matched with results in

Table 6.2. Overall group breakdown of the classes. Unique students are students who took their *last* attempt in the given time period. Percentages are based on unique students.

	CS1-NPE		CS1-PE	
	2001–2007	2008–2012	2001–2007	2008–2012
# Enrolled	1732	2046	1513	1265
# Unique	1510	1868	1365	1179
% Female	25.4%	31.6%	15.4%	18.9%
% URM	10.5%	14.5%	5.3%	9.0%
% Comp Major	25.6%	41.7%	39.2%	53.7%

the prior evaluation of this course [110]. Specifically the Fail Rates in CS1-NPE and CS1-PE reported in the previous study matched the rates in our data.[2]

6.1.1 Data Analysis

Throughout our analysis, we note that different subsets of students are considered for different calculations, dependent on a pre-selected criteria. For example, a student who took CS1-NPE but earned a non-CS degree would not be considered in our calculation for Time-to-Degree. Rather than reporting the number of students for each metric throughout, we provide these numbers in Tables 6.3 and 6.4.

To answer our research questions, we use a combination of descriptive statistics, statistical tests for significance with significance set at $p = 0.05$, and effect sizes. To determine whether particular factors impacted outcomes for students when compared against other factors, we used a likelihood ratio test on regression models [101]. For clarity, more details regarding the particular tests performed are included along with the results.

[2]Numbers matched or were within 1% of those previously reported. Conversations with the office providing the data explained that different databases handle some students differently (e.g., students who withdrew for medical leave, etc.).

Table 6.3. CS1-NPE Sample Sizes

	All		Male		Female		Non-URM		URM	
	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012
# Enrollment	1732	2046	1289	1398	443	647	1517	1749	215	297
# Unique	1510	1868	1126	1276	384	591	1351	1597	159	271
# CS Majors Entered	739	1037	599	807	140	229	642	875	97	162
# Non-CS Majors Entered	771	831	527	469	244	362	709	722	62	109
# >= 5 Upper Div Courses	405	763	321	618	84	174	366	699	39	94
# CS Graduates	387	779	303	600	84	178	355	695	32	84

Table 6.4. CS1-PE Sample Sizes

	All		Male		Female		Non-URM		URM	
	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012
# Enrollment	1513	1265	1261	1023	252	241	1431	1140	82	125
# Unique	1365	1179	1155	955	210	224	1292	1073	73	106
# CS Majors Entered	669	730	588	615	81	115	627	664	42	66
# Non-CS Majors Entered	696	449	567	340	129	109	665	409	31	40
# >= 5 Upper Div Courses	567	664	494	562	73	102	540	613	27	51
# CS Graduates	535	633	465	532	70	101	512	590	23	43

Table 6.5. An overall summary of average values for binary metrics, including significance test (* for p<0.05) and effect size.

	CS1-NPE				CS1-PE			
	2001–2007	2008–2012	*p* (Z-test)	Relative Risk	2001–2007	2008–2012	*p* (Z-test)	Relative Risk
Fail Rate	0.236	0.099	**2.909e-31***	0.418	0.200	0.179	0.161	0.895
Retained	0.471	0.624	**9.482e-11***	0.711	0.714	0.767	**0.024***	0.816
Switched	0.051	0.159	**1.101e-12***	0.886	0.083	0.163	**3.469e-05***	0.914

Table 6.6. An overall summary of average values for continuous metrics, including significance test (* for p<0.05) and effect size.

	CS1-NPE				CS1-PE			
	2001–2007	2008–2012	*p* (T-test)	Cohen's D	2001–2007	2008–2012	*p* (T-test)	Cohen's D
Upper-Division GPA	2.708	2.842	**0.002***	0.187	2.875	2.921	0.290	0.061
Time-to-Degree	4.490	4.490	0.937	-0.002	4.070	4.196	**0.042***	0.063

6.2 Results

6.2.1 RQ1: All Students

Tables 6.5 and 6.6 summarize the average metrics across both time periods for both courses. The overall average for each metric is followed by the significance and effect size test used to analyze the results. For binary metrics, such as Fail Rates, Retained, and Switched rates, we used a Z-test to test for significance. Relative risk was used in order to analyze the effect size of the intervention in our data [1]. Relative risk conveys the risk of a negative outcome (failure, leaving the major, not joining the major) after an intervention relative to the risk before. Values less than 1 express a reduction in risk (e.g., the relative risk for Fail Rates for CS1-NPE of 0.418 means the risk of failing after the best practices intervention is 41.8% of the average risk before the intervention).

For continuous metrics, such as Upper-Division GPA and Time-to-Degree, we used a T-test for significance and Cohen's D for effect size. Since standard Cohen's D uses high variance in the data to calculate an accurate effect size, standard Cohen's D does not work well with GPA data. Hence, we used a variation of the Cohen's D appropriate for GPA analysis [164].

Looking at students in CS1-NPE in Tables 6.5 and 6.6 we see substantial positive impacts occurring in Fail Rates, Retention rates in the major, Switch rates into the major, and upper-division course performance. The reduced failure rates for CS1-NPE also led to fewer students retaking CS1 (from 8.6% of all students in the course failing and retaking it to 4.1%) whereas CS1-PE saw an rise in students retaking CS1 between the time periods (from 11% to 14.8%).

The relative risk indicates that students who took CS1-NPE with the intervention were at around 42% of the total average risk of failing the course, 71% of the total average risk of not being retained in the major, and at around 89% of the total average risk of not switching into the major. The positive effect, per Cohen's D, on Upper-Division GPA is small (where a small effect size is considered at 0.2) [164]. For all these metrics, students in the best practices time period saw statistically significantly improved outcomes with varying effect sizes. There was no perceivable impact to Time-to-Degree as the average is the same for both time periods. We note that a myriad of factors impact Time-to-Degree but this is consistent with the other findings that there is no evidence students were harmed by the best-practices course.

6.2.2 RQ2: Effect of Time Periods

Comparing CS1-NPE and CS1-PE statistics

Recall that there was no intervention implemented in CS1-PE during the time period of the intervention in CS1-NPE. Hence, our first step in answering RQ2 is to explore how students performed when taking CS1-PE in each time period. Tables 6.5 and 6.6 provide the results for CS1-PE. Here we see that CS1-PE also experienced statistically significant improvements to Retained and Switched. Unfortunately, it also saw a statistically significant increase in Time-to-Degree. For Fail Rates and Upper-Division GPA, it appears CS1-NPE uniquely benefited in a statistically significant way during the best practices time period. For Retained and Switched, both courses benefited which might be expected given the increased interest in computing nationally during the same time period. Examining the effect size for Retained and Switched, we see the impact on CS1-NPE appears larger than for CS1-PE as the relative risk is higher for CS1-PE. We hesitate to draw too large a conclusion from this, however, as CS1-PE started with better values for each metric. Lastly, although Time-to-Degree remained roughly constant between time periods for CS1-NPE, it worsened significantly for CS1-PE.

Outcomes per Year

One possible reason for the improvements for the two averages between time periods could simply be a steady improvement over time (e.g., a steady positive slope from 2001–2012 would result in lower averages in 2001-2007 than 2008–2012). To explore this possible explanation, Figures 6.1–6.5 contain the average for each metric over each academic year. The black vertical bar indicates the year when best practices were introduced into CS1-NPE.

The Fail Rates in Figure 6.1 varied year to year with no general trend (although there

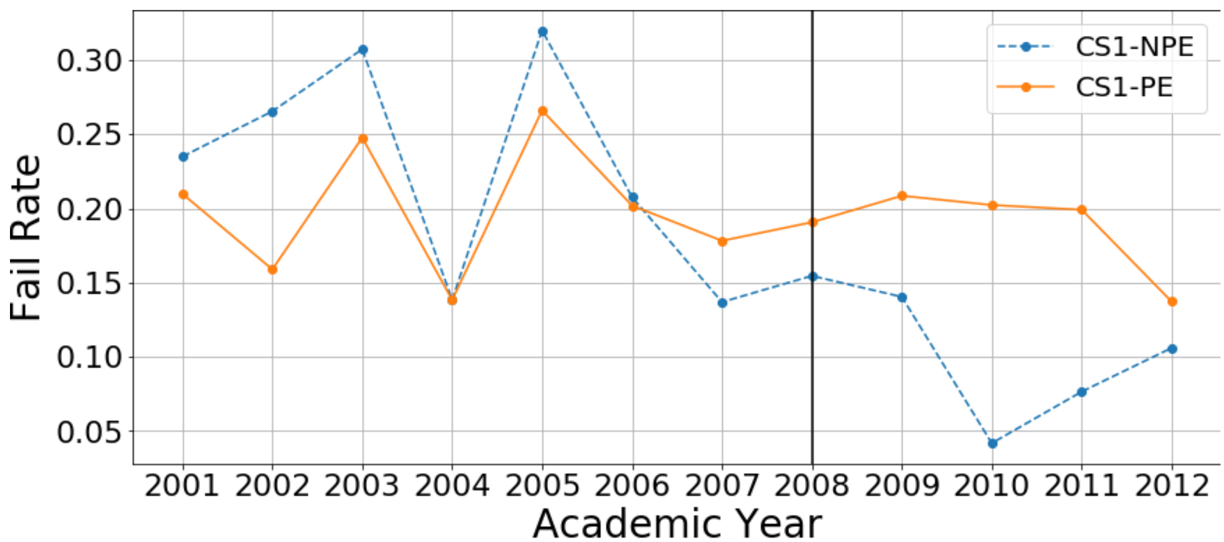


Figure 6.1. Average annual Fail Rates for CS1-NPE and CS1-PE.

might have been a negative trend for CS1-NPE starting in 2005). The decline in Fail Rates for CS1-NPE stands out relative to CS1-PE. Retained rates for CS1-PE in Figure 6.2 remain relatively constant with a slight increase in the later time period whereas CS1-NPE sees a marked increase over time (2004 is a particularly poor year for CS1-NPE). CS1-PE and CS1-NPE appear rather comparable for the rate of students switching into the major over time, in Figure 6.3, with perhaps CS1-NPE catching up to CS1-PE during the best practices time period.

Upper-Division GPA over time, found in Figure 6.4, shows that students in CS1-NPE during the earlier time frame under-performed in later courses relative to those who took CS1-PE. Encouragingly, after the change to best practices, CS1-NPE appears to close that gap. Figure 6.5 shows that Time-to-Degree over time remains mostly consistent. The fact that Time-to-Degree is longer for students in CS1-NPE is expected as CS1-NPE is the first of a two course sequence relative to a single CS1-PE course. Both courses feed into a long dependency chain of required courses for computing majors that might explain the additional quarter (approximately).

Regression Modelling for Time

Another approach to examining the impact of time on the results is to use regression modeling on CS1-NPE. Specifically, the goal is to build a regression model (logistical for binary

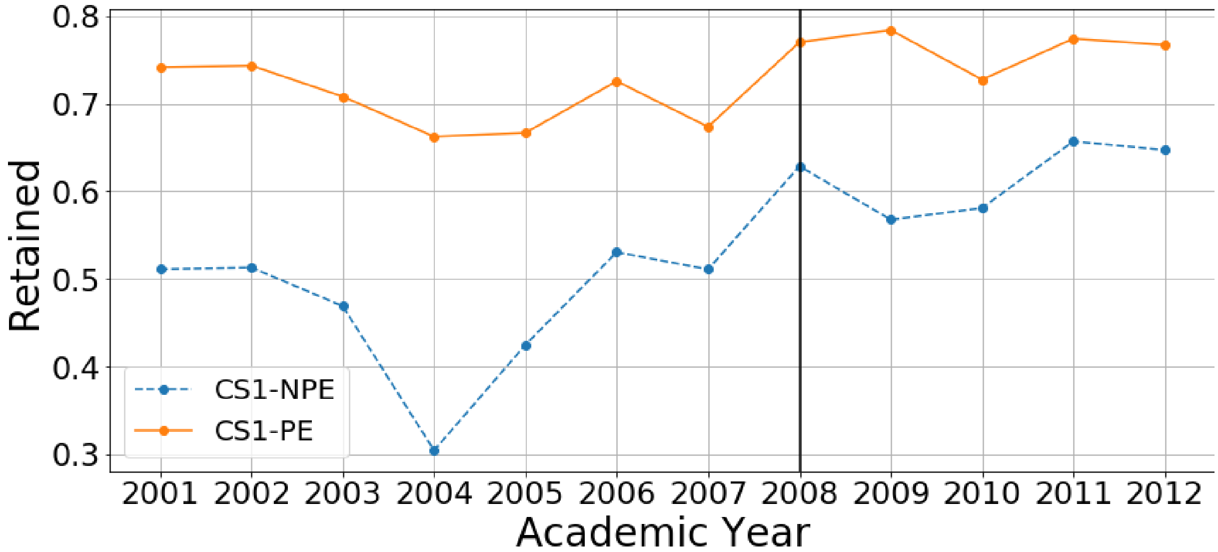


Figure 6.2. Average annual Retained rates for students who entered CS1-NPE and CS1-PE as computer science majors and graduated with a computer science degree. Y-axis begins at 0.3 to help show differences over time.

outcomes, linear for continuous outcomes) to predict student outcomes. The first model is given years as dependent variables (along with an offset) to predict the particular outcome. The second model is provided with the dependent variable of best practices (a binary variable false for 2001–2007 and true for 2008–2012) on top of the year and offset variables. The question becomes whether the model's accuracy improves by adding the best practices variable. If the model's accuracy improves significantly, this means more of the variance in the outcomes can be explained if the model knows about the course revision than without. This would provide evidence that the transition to best practices changed the outcomes for CS1-NPE students beyond changes over time.

To accomplish this, we use the likelihood ratio test [101]. Specifically, we observe the likelihood ratio between models for each metric and their corresponding chi-squared p-value, with degree of freedom 1, to see if the addition of the intervention feature is significant in predicting student outcomes. The results of these tests appear in Table 6.7. For Fail Rates, Retained, and Upper-Division GPA, including the best-practices feature better informs the model than time alone. This provides further evidence that the course-redesign impacted outcomes

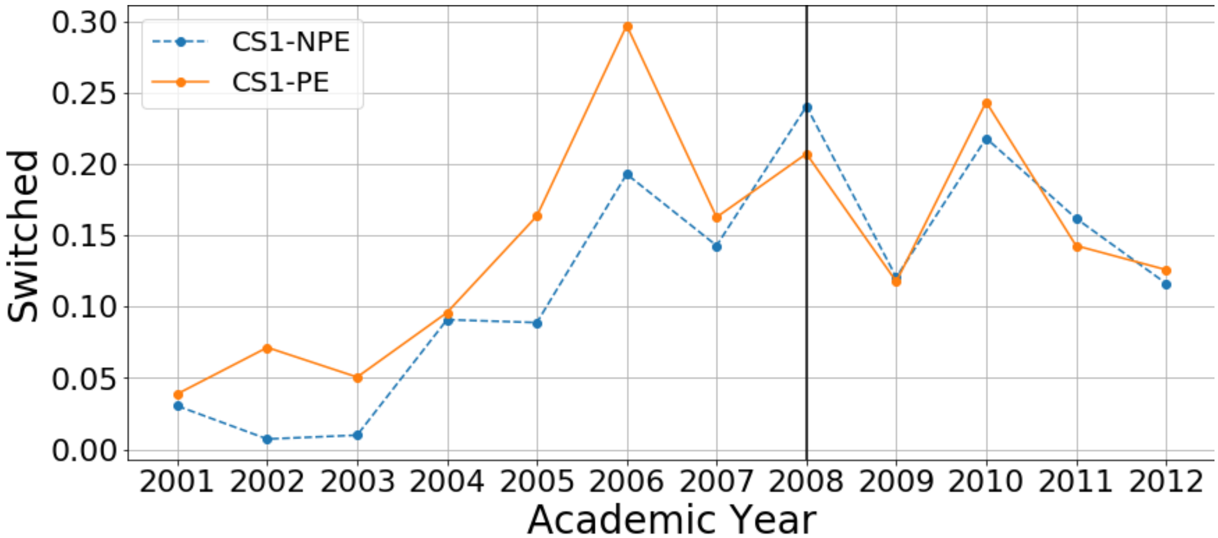


Figure 6.3. Average annual Switched rates for students who entered CS1-NPE and CS1-PE as non-computer science majors and graduated with a computer science degree.

Table 6.7. Results of likelihood ratio test and corresponding chi-squared p-values between regression models. For Fail Rates, Upper-Division GPA, and Retained, a best-practices feature significantly improves the model beyond the model that includes only time in years.

Metric	Likelihood Ratio	*p*
Fail Rates	11.021	**0.0009***
Retained	5.089	**0.024***
Switched	1.529	0.216
Upper-Division GPA	4.000	**0.046***
Time-to-Degree	2.800	0.094

beyond what one might expect from just changes over time.

6.2.3 RQ3: Underrepresented Groups

Table 6.8 provides the outcomes for male and female students across both time periods in CS1-NPE. The addition of best practices into CS1-NPE appears to have a positive benefit for both male and female students. Recall that our data set is heavily skewed towards males in each year observed in the study. Similar to our overall results, Fail Rates, Retained and Switched rates, and Upper-Division GPA improved significantly for male and female students after best practices were enacted. Examining the scale of the benefits, women appear to benefit more for

Table 6.8. CS1-NPE Results by Gender

	Male				Female			
	2001–2007	2008–2012	*p* (Z-test)	Relative Risk	2001–2007	2008–2012	*p* (Z-test)	Relative Risk
Fail Rate	0.231	0.094	**4.514-23***	0.405	0.251	0.110	**4.799e-10***	0.438
Retained	0.471	0.634	**5.660e-10***	0.691	0.471	0.585	**0.033***	0.785
Switched	0.040	0.188	**1.763e-14***	0.846	0.074	0.122	0.057	0.948
	2001–2007	2008–2012	*p* (T-test)	Cohen's D	2001–2007	2008–2012	*p* (T-test)	Cohen's D
Upper-Division GPA	2.736	2.849	**0.023***	0.156	2.599	2.824	**0.016***	0.323
Time-to-Degree	4.489	4.505	0.786	0.008	4.509	4.433	0.477	-0.038

Table 6.9. CS1-PE Results by Gender

	Male				Female			
	2001–2007	2008–2012	*p* (Z-test)	Relative Risk	2001–2007	2008–2012	*p* (Z-test)	Relative Risk
Fail Rate	0.177	0.172	0.764	0.973	0.313	0.203	**0.005***	0.649
Retained	0.707	0.771	**0.012***	0.784	0.765	0.748	0.779	1.075
Switched	0.088	0.171	**0.0002***	0.910	0.062	0.138	**0.048***	0.919
	2001–2007	2008–2012	*p* (T-test)	Cohen's D	2001–2007	2008–2012	*p* (T-test)	Cohen's D
Upper-Division GPA	2.894	2.905	0.815	0.014	2.748	3.011	**0.018***	0.366
Time-to-Degree	4.126	4.177	0.464	0.025	3.693	4.297	**1.275e-05***	0.302

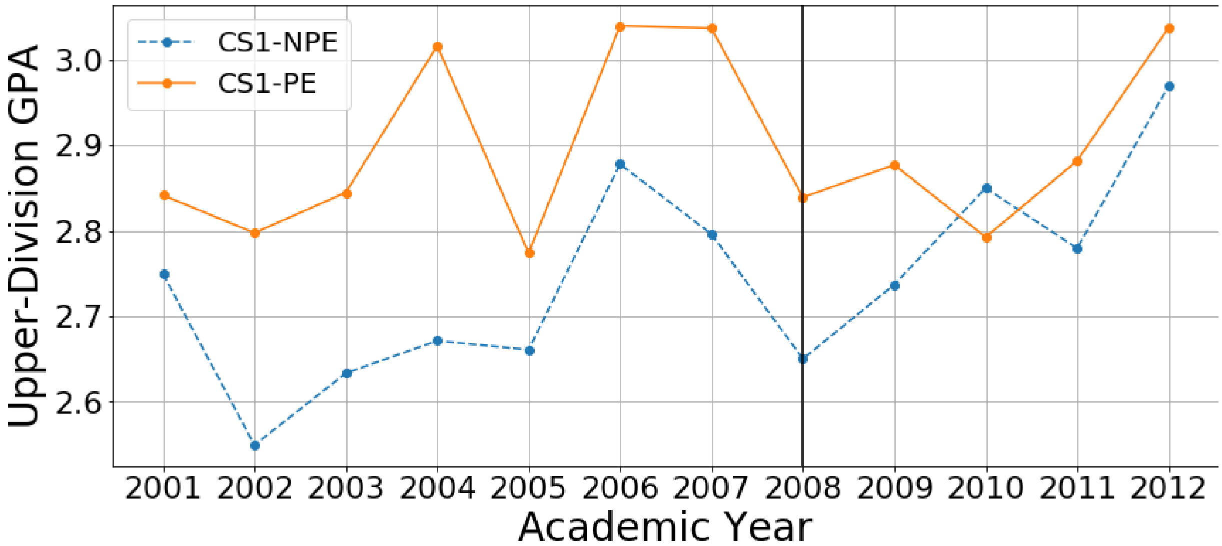


Figure 6.4. Average annual Upper-Division GPA for students who had at least five upper division course attempts. Y-axis begins at a GPA of 2.5 to help show the differences over time.

some metrics and men for others.

Table 6.9 provides the outcomes for CS1-PE where Switch rates are significant for both genders whereas Fail Rates, Upper-Division GPA, and Time-to-Degree are significant only for women. It appears that much of the benefits that we saw overall for CS1-PE were due to large improvements for women over this time period. It is unclear what may have changed between these time periods to benefit female students in CS1-PE and remains a topic for future analysis.

Similar benefits can be seen when comparing non-URM versus URM students throughout both courses and time periods in Tables 6.10 and 6.11. However, it is important to acknowledge that the number of URM students that were considered in each metric analysis was small compared to non-URM students. The first finding that stands out is simply than URM students struggle at our institution relative to non-URM students in CS1-NPE and CS1-PE. Fail Rate is particularly striking as the failure rate for URMs for CS1-NPE was nearly twice that of CS1-PE (44.7% versus 24.4%) before the transition to best practices. After the transition, CS1-NPE Fail Rate for URM dropped remarkably from 44.7% to 17.5% between time frames whereas CS1-PE saw an increase for URM students from 24.4% to 28.0%. Also striking is that the percentage

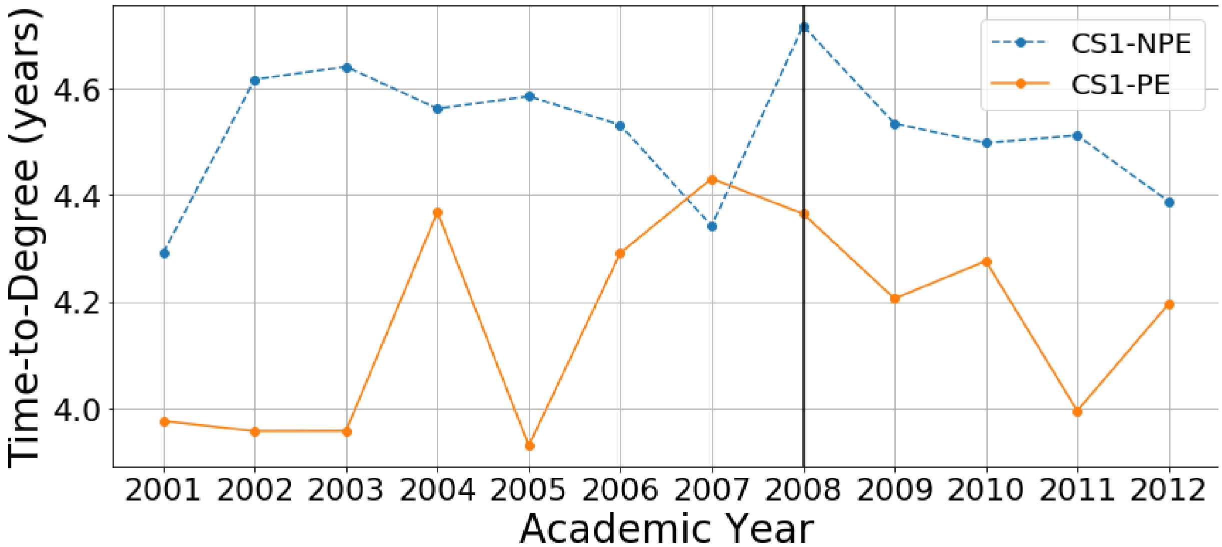


Figure 6.5. Average annual Time-to-Degree for computer science graduates. Y-axis begins at 3.9 years as expected time-to-degree is 4 years at UC San Diego.

of URM students switching into the major for CS1-NPE and CS1-PE was 0 during the 2001-2007 time period. The numbers were small (between 2001–2007, only 62 URM students took CS1-NPE as non-majors and only 31 URM students took CS1-PE), but the fact none switched remains jarring. In the best practices time period, the Switch rate for CS1-PE increased to 7.5% while CS1-NPE rose to 14.7%.

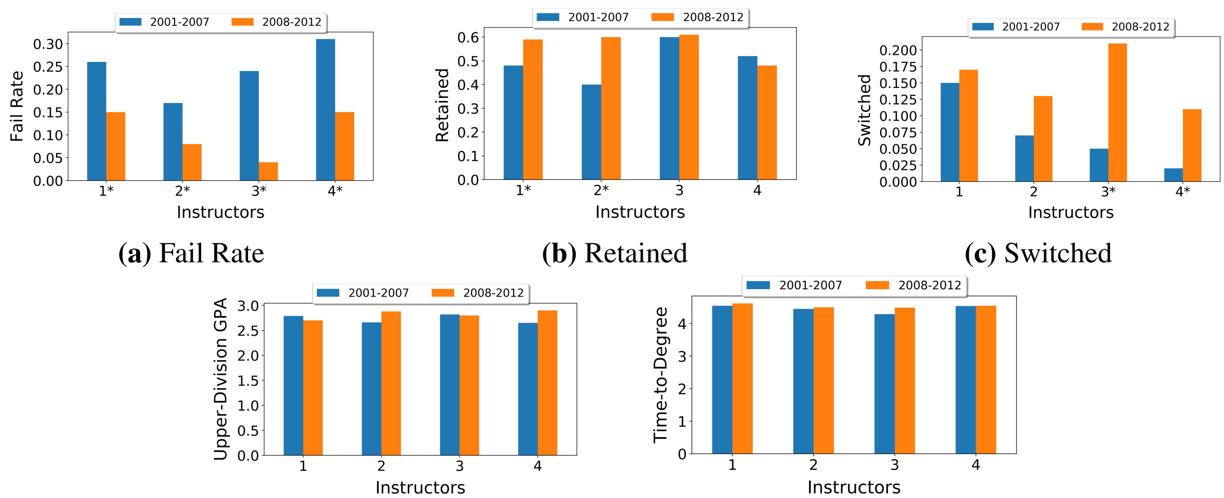
Overall, for CS1-NPE, both URM and non-URM students benefited statistically significantly for three of the five metrics. In addition, for CS1-NPE, effect sizes for improvements for URMs were comparable to non-URMs except for Upper-Division GPA where URMs experienced a considerably larger improvement than non-URM. In contrast, for CS1-PE, none of the metrics are statistically significant for URM students. This is likely due to a combination of lower effect sizes for all students and lower numbers of URM students in the course (relative to CS1-NPE).

Table 6.10. CS1-NPE Results for non-URM and URM Students

	Non-URM				URM			
	2001–2007	2008–2012	*p* (Z-test)	Relative Risk	2001–2007	2008–2012	*p* (Z-test)	Relative Risk
Fail Rate	0.206	0.086	**1.548e-23***	0.416	0.447	0.175	**2.839e-12***	0.392
Retained	0.492	0.662	**1.748e-11***	0.666	0.330	0.420	0.150	0.866
Switched	0.055	0.161	**7.005e-11***	0.888	0	0.147	**0.001***	0.853
	2001–2007	2008–2012	*p* (T-test)	Cohen's D	2001–2007	2008–2012	*p* (T-test)	Cohen's D
Upper-Division GPA	2.754	2.877	**0.007***	0.175	2.273	2.581	**0.033***	0.409
Time-to-Degree	4.461	4.453	0.884	-0.004	4.859	4.792	0.744	-0.034

Table 6.11. CS1-PE Results for non-URM and URM Students

	Non-URM				URM			
	2001–2007	2008–2012	*p* (Z-test)	Relative Risk	2001–2007	2008–2012	*p* (Z-test)	Relative Risk
Fail Rate	0.197	0.168	0.055	0.850	0.244	0.280	0.567	1.148
Retained	0.726	0.783	**0.016***	0.791	0.548	0.606	0.551	0.871
Switched	0.087	0.171	**3.303e-05***	0.908	0	0.075	0.118	0.925
	2001–2007	2008–2012	*p* (T-test)	Cohen's D	2001–2007	2008–2012	*p* (T-test)	Cohen's D
Upper-Division GPA	2.894	2.938	0.318	0.059	2.503	2.720	0.307	0.248
Time-to-Degree	4.061	4.190	**0.044***	0.064	4.261	4.279	0.947	0.009



(a) Fail Rate **(b)** Retained **(c)** Switched

(d) Upper-Division GPA **(e)** Time-to-Degree

Figure 6.6. Averages for each metric for each instructor who taught CS1-NPE in both time periods. Instructor significance for a given metric between time periods is indicated with an asterisk next to the instructor's label.

Table 6.12. CS1-NPE sample sizes for instructors who taught during both time frames.

	Instructor 1		Instructor 2		Instructor 3		Instructor 4	
	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012	2001–2007	2008–2012
# Enrolled	333	499	453	168	129	502	612	172
# Unique	280	455	404	158	113	481	522	163
# CS Majors Entered	178	260	224	67	47	271	265	69
# Non-CS Majors Entered	102	195	180	91	66	210	257	94
# >= 5 Upper Div Courses	101	195	110	53	34	209	146	44
# CS Graduates	101	187	102	52	31	210	142	43

6.3 Discussion

6.3.1 Better Instructors?

One concern for our analysis was that there might have just been an improvement in the quality of teachers in the later time period. Perhaps more dedicated teachers were willing to teach the best practices version of the course than those who taught the course previously. We examined

Table 6.13. Results of likelihood ratio test and corresponding chi-squared p-values comparing a model with time and instructor features against a model that also includes the best-practices feature. For Fail Rates and Retained, the best-practices feature significantly improves the model.

Metric	Likelihood Ratio	p
Fail Rate	62.926	**2.146e-15***
Retained	6.580	**0.010***
Switched	0.819	0.365
Upper-Division GPA	3.000	0.083
Time-to-Degree	2.000	0.157

this in two separate ways. First, we looked at four instructors who had taught both before and after the course redesign. Because sample sizes reduced when looking at individual instructors who might have taught only a single term during one of the time periods, we expected few findings to be statistically significant. However, we still found statistically significant reductions for Fail Rates, Retained, and Switched. Figure 6.6 provides the results per instructor and sample sizes for each instructor appears in Table 6.12. From Figure 6.6, we see that student outcomes improved for each instructor during the best practices time period for all significant differences in performance. Overall, it appears the transition to best practices by these instructors corresponded to improved student outcomes (Fail Rates, Retained, and Switched).

Second, we also used a likelihood ratio test to examine whether our linear or logistic model that includes offset, years, and instructor parameters would improve if given best practices improved the models' accuracy. Results can be found in Table 6.13. For both Fail Rates and Retained, the best practices feature improves the model's performance significantly, suggesting it was the adoption of best practices that resulted in the improved outcomes for those metrics.

6.3.2 Implications of Findings

Challenges from Colleagues: A motivation for this study was faculty colleagues challenging research results on Peer Instruction, Media Computation, and Pair Programming. A common refrain was that although these approaches may have succeeded in lowering failure rates or briefly improving retention, it was really just bringing in poorly prepared students who were bound to struggle and leave the program later. We are quite pleased that, for this particular redesign, the larger body of students went on to succeed in the rest of the program at the same rate as before (or better).

Factors in Success: Given the positive outcomes related to the redesign, what factors led to its success? We suspect there were at least two main contributors. The first was the mandate to adopt evidence-based practices and the selection of three practices that, although they each seemed promising at the time, have each been shown to be broadly effective in the research since. The second was a broad commitment to the redesign from the faculty who drove the changes themselves, the graduate and undergraduate instructional staff who helped enact those changes, and the other faculty who adopted the course when they later taught it.

Underrepresented Groups: A recent meta-analysis of pedagogy in STEM found that active learning causes the achievement gap for underrepresented groups (relative to represented groups) to be narrowed [150]. We found that students from both groups benefited from the course redesign and hence found no consistent reduction in the achievement gap for our metrics. However, the raw difference in the benefit was larger in many cases because the outcomes for students from underrepresented groups (particularly URMs) were worse at the start. For example, the Fail Rate for represented students in CS1-NPE dropped from 20.6% to 8.6% after the addition of best practices while the Fail Rate for URM students in CS1-NPE dropped from 44.7% to 17.5%. As raw percentages, URM students could be seen as benefiting more from the change. However, as a ratio, URM students failed CS1-NPE 2.16 times more than represented groups before the change and 2.04 times after the change. Although the improved outcomes for URM students is encouraging, the resistant disparity in outcomes for URM students is a clear problem that urgently requires more research and critical analysis.

Longitudinal Studies: This project suffered from several challenges due to its longitudinal nature, including gaining approval to collect sensitive data and inconsistent data reporting within internal databases. Despite these challenges, we strongly encourage similar studies so that the community can improve understanding of the long-term impact of pedagogical changes.

6.3.3 Threats to Validity

Change in Time: Our comparisons of the CS1-PE and CS1-NPE courses, as well as regression modeling of course outcomes, both suggest that the redesign of the course explains the improvement in Fail Rates and Retained. However, it is impossible to know if there were changes in the perception of the computing field at the time of the redesign—particularly given the increases in enrollments in CS between 2008 and 2013 at UC San Diego and nationally [99]. In addition, changes elsewhere in the major may have impacted student outcomes. However, there were no significant systemic changes to our major over the examined time period.

Better Teachers: Although the improved Fail Rates and Retained for the four instructors who taught before and after the course redesign suggest that the improvement in the course are associated with the better outcomes found more broadly, there were still differences in instructors between time frames. For example, one instructor who had slightly higher Fail Rates in general taught the course more during the earlier time frame than the later. As such, it is difficult to know definitely how large an impact the variation in instructors had on the outcomes of this study.

Combining Best Practices: One challenge in interpreting these results is that by combining multiple best practices in one course, we cannot distinguish which were more important or if they were needed in combination. As such, we can only conclude the combination appears to have been successful.

6.3.4 Call to Action

Jobs in the computing industry are plentiful and pay well. Unfortunately, diversity in the computing industry continues to languish. We adopted practices in one single course that increased the number of successful students graduating with computing degrees. To be candid, the implementation of the course itself was significantly less challenging (and more fun and rewarding) than conducting this longitudinal analysis. Given the extensive evidence showing these best practices provide substantial benefit to our students, why are more institutions not replicating or implementing similar changes? We do not know for sure, but can posit several factors. First, we recognize faculty change can be hard [52]. Second, we suspect institutions feel the need to distinguish themselves. In particular, computer scientists may feel the need to build their own solutions as the computing field favors the "new" (albeit un- or less-tested) and denigrates anything "old" (where old is a handful years?). Are these self-centered factors worth the cost of keeping the discipline less accessible to students who could succeed?

We also recognize US society is biased and does not provide a level playing field for many subgroups, but especially people of color. The computing profession is infamous for its lack of diversity and for limited results in efforts to make change in this area. While URM students benefited substantially from the best practices we implemented, at the end they were still *2 times* more likely to fail the course than majority students. We as a community should find this deeply concerning. The time has come for CS instructors to adopt evidence-based instructional practices to improve outcomes for their underrepresented students and for the CER community to prioritize finding further solutions to help URM students.

6.4 Summary

In this longitudinal study that spans two decades, we examined the impact of redesigning a CS1 course for students without prior programming experience to include Peer Instruction, Media Computation, and Pair Programming. We find that over the five years after the redesign, students in the course had lower failure rates and computing majors had a higher chance of being retained through graduation. These improvements appear connected to the course revisions, even when examining changes in other courses (without changes) and examining changes over time. Moreover, instructors who taught before and after the revision saw similar improvements to student outcomes after the course revision.

Despite this larger group of students succeeding and progressing into later computer science courses, we find no evidence that their outcomes were worse than those of students from before the redesign. Although there are encouraging signs that students from underrepresented and represented groups benefited from the changes and that the magnitude of those benefits were higher for URM students, there remains a large gap between URM students and represented students that deserves further study. Given the multiple research studies documenting the success of each of these instructional practices and the evidence from this study that these improved outcomes are sustained over many years, we hope CS1 instructors will be further encouraged to adopt these evidence-based instructional practices for the betterment of their students.

6.5 Acknowledgements

The authors thank Rita Keil, Laura Kertz, Carolyn Sandoval, and Ying Xiong for their assistance as well as the reviewers for their helpful feedback. This work was supported in part by NSF award #1712508 as well as a UCSD Sloan Scholar Fellowship, a UCSD STARS Fellowship, and a Gates Millennium Scholarship. Chapter 6, in part, is a reprint of the material as it appears in the Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. "A Longitudinal Evaluation of a Best Practices CS1". The dissertation author is the primary investigator and author of this paper.

Chapter 7

Discussion

This dissertation began with the following research questions

- **Exploration:** What factors related to CS1 student success have been explored in the literature?
- **Holistic:** What factors do students in early computer science courses, including CS1, report through course experience surveys as being sources of struggle and barriers to their success and experience in their courses?
- **Pedagogy:** How does implementing multiple best practice pedagogical approaches in a CS1 course improve outcomes?

7.1 Exploration

What factors related to CS1 student success have been explored in the literature?

7.1.1 Instructor and Student Factors

Our literature review uncovered a large amount of papers related to the factors tied to the student experience and the instructional setting. Many of these factors were studied in relation to some sort of student outcome such as final CS1 course grade or retention in the major. This is expected as student learning is tied to the student itself and the learning environment that they are placed in. In addition, researchers are able to better control the instructional setting through pedagogical techniques and interventions, thus making it a prime area for research in student success.

7.1.2 The Shortage of Multi-Factor Studies

This literature review highlights how incredibly vast the CS1 student success research space is. Across over 300 papers, we found a wide range of factors that have been studied in regards to a variety of CS1 outcomes. Furthermore, the factors themselves ranged from cognitive abilities, gender, race/ethnicity, sense of belonging, self-efficacy, prior programming experience, and many others. It is also important to note that some of these studies have either inconclusive or conflicting results. This may be due to the nature and sample of the studies not being similar, but this should provide motivation to work towards replicating these findings in other situations to gain a better understanding of how certain factors may be influencing students.

Many studies also focused on a single or two factors in relation to CS1 student success. There is nothing inherently wrong about this nor does it lower the value these studies bring to the research community. However, with so many factors being studied, our work indicates that it is perhaps time to begin focusing on multi-factor studies. The field has clearly demonstrated that CS1 student success is possibly tied to a variety of factors and these factors students are likely experiencing

7.2 Holistic

What factors do students in early computer science courses report as being sources of struggle and barriers to their success and experience in their courses?

7.2.1 Sources of Struggle

We find that students are reporting struggle in their early computer science courses across four prominent areas: In-Class Confusion, Lack of Sense of Belonging, Lack of Confidence, and Personal Obligations. These four areas span different aspects of the student, learning environment, and situations outside of the classroom. This diversity in sources of struggle indicates that different students have different needs and must be helped in a way that specifically targets their need. If they are confused about the class material, perhaps additional tutoring or resources can help them improve. If a student lacks confidence or feels like they don't belong in computing, creating a more supportive classroom environment or creating peer support programs within the computing department can be an effective intervention. Finally, and perhaps the most difficult to address, is if a student's outside personal obligations are impacting their ability to learn. Our study did not ask students specifically what type of personal obligation was causing their stress but one can think of issues such as financial stress, familial or partner issues, and commuting/transportation barriers that may be impacting their ability to participate and engage fully in their courses.

We observed a pressing trend among students across courses, particularly CS1-NPE and CS1-PE, of the lower performing students reporting higher scores on their responses, indicating a higher level of stress or at the least a higher perceived impact of this stress source on their ability to perform. Figure 7.1 highlights this trend as we can see an increase in factor score distributions as we see from the higher performers from quartile 1 to the lower performers in quartile 4.

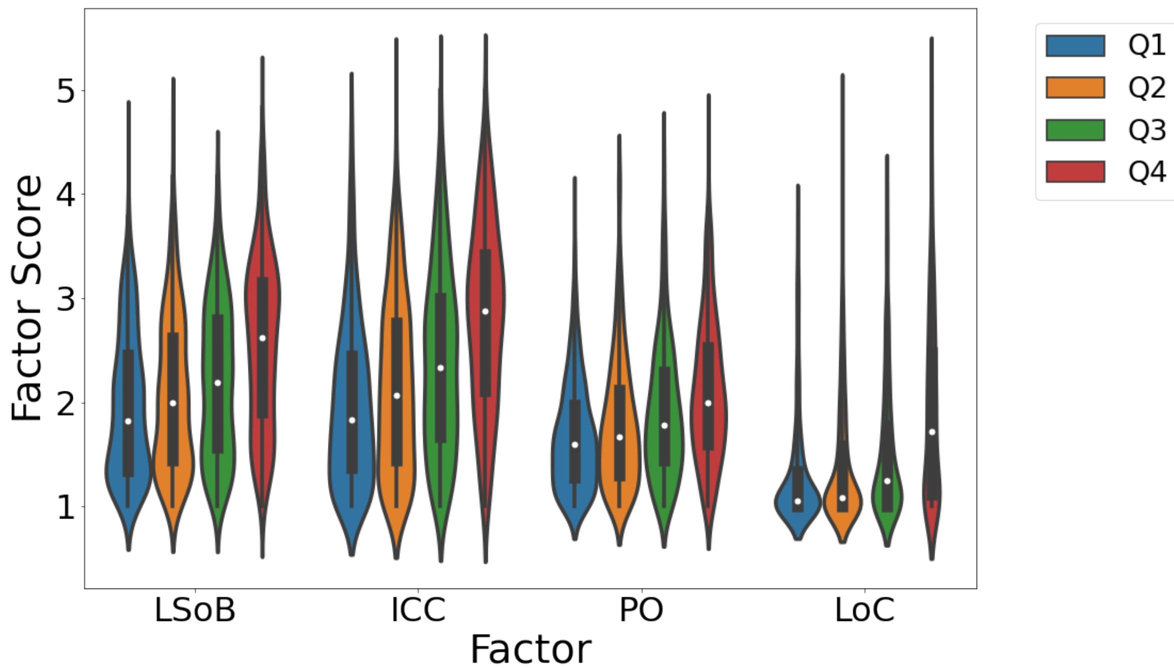


Figure 7.1. Distribution of factor scores across quartiles and factors. The distribution for each factor and quartile span across all courses.

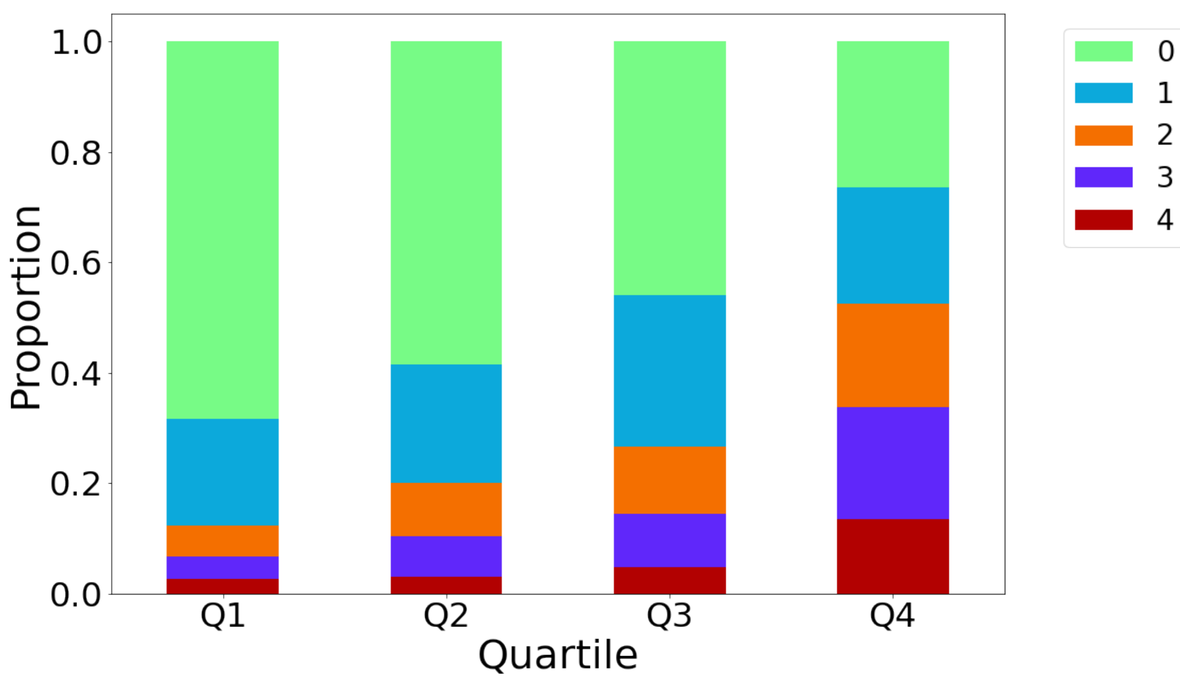


Figure 7.2. Proportion of factors students self-report above the 75th percentile per quartile. The data per quartile covers all courses.

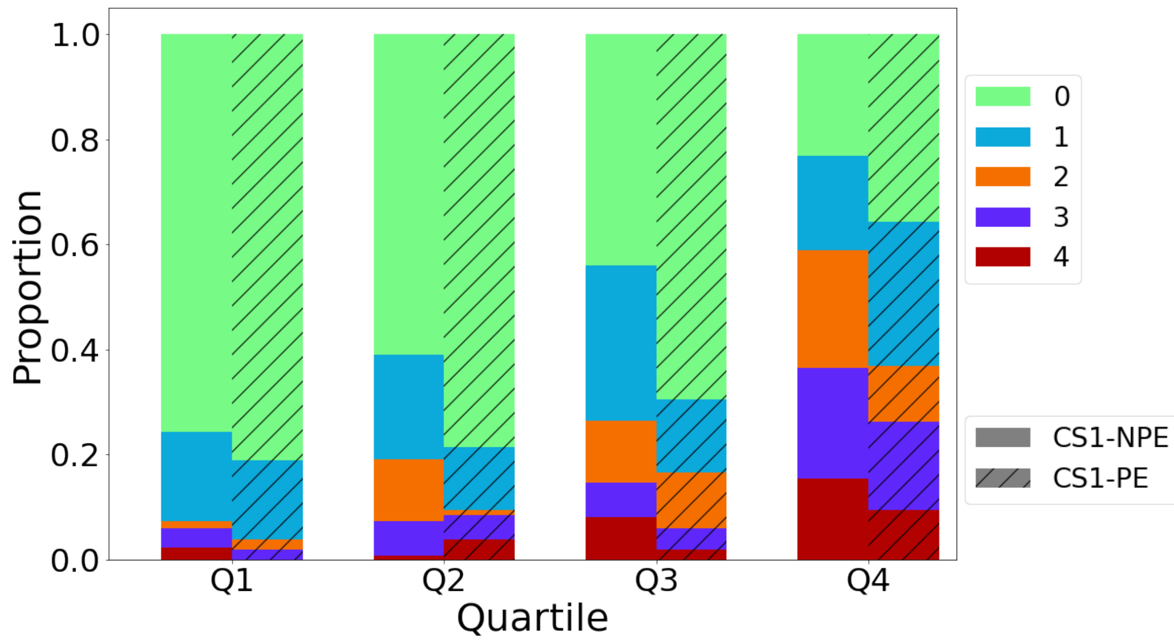


Figure 7.3. Proportion of factors students self-report above the 75th percentile, per quartile, in CS1-NPE and CS1-PE.

7.2.2 Impact of Multiple Struggles

One important trend we found in our results is how lower performing students were reporting higher struggles across multiple sources of struggle. As you can see in Figures 7.2 and 7.3, higher performing students are not reporting struggles at the same rates and across the same number of sources as lower performers. It is important to note that survey response rates were lower for students in the bottom quartiles than in the upper ones, but this can be more indicative of some struggling students not speaking up. It is possible that other students were facing stress from these sources but simply did not report it.

This further provides evidence that the CS1 student experience is multi-faceted, with students facing issues from a potential variety of sources. If instructors and institutions aim to better support their students, they must understand the areas students require additional assistance with. If a student is struggling with issues unrelated to the course material, a pedagogical change may not be the most beneficial for them. Similarly, if a student is feeling a low sense of confidence or outside personal obligations are interfering with their learning, a mindset workshop will likely provide no benefit. Furthermore, we must also aim to understand potentially compounding effects of these factors. A key follow up question raised from these results is how factors are potentially interacting and influencing each other. Is a student struggling with the course material because their personal obligations are too much? Is a student feeling a low sense of belonging because they are not doing well with the material? These interactions should be focuses of future research in this area.

7.2.3 Underrepresented Groups

We found similar trends among women, BLNPI, and transfer students in our results. These groups of students were reporting higher levels of stress across the different sources uncovered in the study. It is also important to note that these groups also tended to make more of the lower performing students, coinciding with our results explained above. This is concerning as these are the groups CS has historically had difficulty attracting and retaining. These results further highlight the importance of seeking to holistically understand how students are experiencing CS1 and other CS courses in their program.

7.3 Pedagogy

How does implementing multiple best practice pedagogical approaches in a CS1 course improve outcomes?

7.3.1 Incorporating Multiple Best Practices

We find that the use of peer instruction, media computation, and pair programming improved CS1 student outcomes at University of California San Diego. After the implementation of this trio of best practices, we saw student fail rates drop across the board. Students were performing better in their introductory programming courses, particularly in the version that implemented this new pedagogy. Students were also being retained at higher rates within the major across both courses, particularly in the updated version. Our results also indicated that the change in pedagogy improved pass rates by simply making the course easier. Students who took CS1-NPE after the changes continued on to earn a higher average GPA in their upper-division courses compared to students in CS1-NPE prior to the changes. Time to degree also remained unchanged indicating this course change had no impact on a student's timeline to earning a degree.

7.3.2 Effect of Time

When analyzing these initial results, we were concerned about the potential impact of time on the improved outcomes. We wished to explore the possibility of these improvements being explained by the fact that students in the latter years were perhaps better prepared or that time frame was the only significant difference between the two. However, our likelihood ratio tests provided evidence indicating that this was not the case. In fact, the presence of the best practice pedagogies were significant in the improved outcomes.

7.3.3 Choice of Instructor

Another concern regarding these results were that these improvements could be likely due to the presence of great instructors. Perhaps instructors willing to adopt these best practices were more dedicated and willing to work with these practices. However, our results showed improved student outcomes, such as fail rates and retention rates, for each instructor implementing these best practices. Our likelihood ratio tests also further suggested that these pedagogies were likely responsible for the improvements and not solely explained by having good instructors.

7.4 Theoretical Foundation

The work presented in this dissertation has demonstrated how the student experience in introductory programming courses are influenced by a myriad of factors. The community has explored many factors, but mostly focusing on a single factor. This has made it difficult to combine and compare studies despite the setting of these studies all being the same, introductory programming courses. Our work has begun looking at the student experience more holistically, an approach that should be considered for future studies in this area.

Our results show the importance of two areas, student factors and instructor factors. Many studies, including our own, have observed factors that are internal to the student such as sense of belonging and confidence. These are not the only ones as other studies have focused on cognitive ability, self-efficacy, and interest in computing. However, our holistic approach hints that a student's experience can be influenced by multiple factors. Lower performing students self-reported their struggles across in-class confusion, sense of belonging, personal obligations, and lack of confidence at higher rates than their better performing peers. A future theoretical framework must focus on understanding how such factors are influencing each other and aim to uncover any potential compounding effects.

Institutions, departments, and educators have more control on instructor factors, such as pedagogy and student learning environment, in their ability to goal of improving student learning. A future theory can use pedagogy as a tool to influence factors and minimize the impact of negative compound effects caused by these factors. Where student factors can give us an understanding of the needs of our students, pedagogy and other instructor factors can be a tool to address these concerns.

Issues related to personal obligations are trickier to influence but should remain in a future theoretical framework. A student is a human being with responsibilities and obligations outside of their studies. A balance must be achieved between supporting students but also not being too lenient such that all can easily pass the course without learning the material. Currently there are not many studies within the computing education space focusing on factors such as familial responsibilities, financial obligations, or other personal issues that can impact student learning. However, this is likely due to the personal nature of such obligations. Asking students to share their experiences on such topics can lead to issues of privacy. In addition, the student may also not be comfortable in sharing such experiences with their instructor or a research team.

The extensive amount of factors also raises the importance of conducting future research in a manner to tease out these interrelationships. Future studies in this area, particularly those focused on understanding which factors are impacting student outcomes, should be designed in a manner of developing correlations and relationships. Ideally we would want to begin uncovering causal relationships where possible. We must also be cautious about exploring too many factors simultaneously. Factors selection for studies should be guided by academic reasoning otherwise we risk further complicating this already difficult to navigate field.

To summarize, the results presented in this dissertation highlight the importance of looking at the student CS1 experience more holistically. Student learning is being hindered by multiple factors that are likely interacting while pedagogy can be used as a tool support learning and mitigate the negative impacts of these factors. Future work in this area must begin looking at multiple factors, not one or two. We demonstrate how there are potential underlying relationships between factors and we must begin understanding these in order to begin developing a theoretical framework.

7.5 Acknowledgements

Chapter 7, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Chapter 8

Conclusion

There are currently high failure rates and low retention rates in computer science, especially among women and BLNPI students. As CS1 is the first course students are exposed to and is foundational to future CS courses, it is incredibly important for students to be successful in this course. The "Learning Edge Momentum (LEM)" hypothesis provides further evidence of how crucial student learning is in CS1 as students are likely to struggle in future CS courses if their CS1 knowledge is lacking and filled with misconceptions.

Researchers have explored a variety of factors related to CS1 student success and have studied the impacts of pedagogical and institutional interventions. However, most of these studies only focus on one or two factors of student success despite the community declaring multiple areas are impacting student CS1 learning. This nature has created pockets of research within the CS1 space lacking a connective framework. We hypothesized that the student experience is composed of multiple factors potentially interacting and compounding on each other. Therefore it is important that future CS1 research, particularly those focused on exploring and improving the student experience, focuses on multiple factors related to student success.

Throughout this dissertation we focused on answering questions related to the exploration of the CS1 literature space, the holistic experience of students, and the pedagogical impacts instructors can have. More specifically, we explored the following research questions: **(1) exploration:** What factors related to CS1 student success have been explored in the literature?

(2) impacts: What factors do students in early computer science courses, including CS1, report as being sources of struggle and barriers to their success and experience in their courses? **(3) pedagogy:** How does implementing multiple best practice pedagogical approaches in a CS1 course improve outcomes?

In Chapter 4, we explored over 300 research papers focusing on factors related to CS1 student success and outcomes. We organized the results of these papers using the Biggs' 3P model of student learning to gain an understanding of where the community is focusing. We found evidence of the community focusing heavily on the impact student and instructional factors and their direct correlation with an outcome in CS1 or the CS major. We also found fewer papers looking at multiple factors and focusing on the techniques students are employing while learning. The learning process is a critical piece of the Biggs' 3P model that appears to be understudied in the CS1 space. Future work should shift focus into these spaces to gain a deeper understanding of how students are experiencing and learning in their CS1 courses.

In Chapter 5, we took a holistic approach to the student experience in early computer science courses. We found evidence that lower performing students across these courses, including CS1, were reporting higher levels of stress across a range of factors. Meanwhile, their higher performing peers were reporting little to no stress on very few or even zero factors. Although we did not explore how the influence of these factors are related, if they even are, future work should begin to explore these potential compounding relationships that are impacting students.

In Chapter 6, we explored the benefits instructors can have on CS1 student learning through the implementation of a trio of best practice pedagogies. We found that fail rates for CS1 dropped and retention in the major improved across all groups, regardless of instructor. Our likelihood ratio tests also indicated that this improvement was likely caused by the pedagogical changes and not simply a difference in students within the time frames. However, despite these improved outcomes, women and BLNPI students were still performing below the rest of their peers. This result, along with our results from Chapter 5, indicate that these students are likely struggling from areas that pedagogical innovations do not address such as sense of belonging or outside personal obligations.

Based on the work presented in this dissertation, we have found evidence indicating how complex the CS1 student experience is. Our literature review uncovered a variety of different types of factors that have been correlated with CS1 outcomes. Such a range of factors is evidence that a single factor is not responsible for students underperforming in their CS1 course. Although three best practice pedagogies improved performance in CS1 and retention in the major, many students continued to struggle.

In summary, CS1 is a critical course for all students aspiring to earn a CS degree at any higher education institution. The concepts presented in this course will only be built upon in future courses, making it important that students leave with a strong foundational understanding. In our exploration we uncovered pockets of research work focusing on a variety of factors, particularly those related to students and the instructional setting. In our holistic study, we found that the students who are struggling and underperforming the most across early computer science courses, especially CS1, were facing high stress from a different sources of struggle that appear to impact their ability to perform. In our pedagogy study, we found that although these techniques of teaching CS1 content did yield benefits, there were still many students, particularly women and BLNPI students, who were still failing and leaving the CS major. This indicates that pedagogical revamps to CS1 may not be sufficient to support all students in CS1 courses, likely due to other factors impacting their learning that pedagogy does not impact.

In the future, we plan to continue exploring the relationships between some of these sources of struggle, particularly at R1 institutions similar to University of California San Diego. We aim to uncover potential compounding effects these issues may have on students to better guide support programs, interventions, and pedagogical practices. Furthermore, we aim to begin consolidating this vast amount of work into a theoretical framework to explain student experiences and guide research in the field.

8.1 Acknowledgements

Chapter 8, in part, is a reprint of material as they appear in three publications: 1) Proceedings of of the 2020 ACM Conference on International Computing Education Research (ICER 2020). Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. “A Longitudinal Evaluation of a Best Practices CS1”, 2) Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021). Adrian Salguero, William G. Griswold, Christine Alvarado, and Leo Porter. “Understanding Sources of Student Struggle in Early Computer Science Courses”, 3) Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Adrian Salguero, Ismael Villegas Molina, Lauren Elizabeth Margulieux, Quintin Cutts, and Leo Porter. “Applying CS0/CS1 Student Success Factors and Outcomes to Biggs’ 3P Educational Model”. The dissertation author was the primary investigator and author of this paper.

Bibliography

[1] Douglas G Altman. *Practical statistics for medical research*. CRC press, 1990.

[2] Christine Alvarado, Cynthia Bailey Lee, and Gary Gillespie. New cs1 pedagogies and curriculum, the same success factors? In *Proceedings of the 45th ACM technical symposium on computer science education*, pages 379–384, 2014.

[3] Christine Alvarado, Gustavo Umbelino, and Mia Minnes. The persistent effect of pre-college computing experience on college cs course grades. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, page 876–881, 2018.

[4] Ana Paula Ambrósio, Fábio Moreira Costa, Leandro Almeida, Amanda Franco, and Joaquim Macedo. Identifying cognitive abilities to improve cs1 outcome. In *2011 Frontiers in Education Conference (FIE)*, pages F3G–1. IEEE, 2011.

[5] American Association of University Women. Educational Foundation. Commission on Technology and Gender and Teacher Education. *Tech-savvy: Educating girls in the new computer age*. American Association of University Women, 2000.

[6] Albert Bandura. Self-efficacy mechanism in human agency. *American psychologist*, 37 (2):122, 1982.

[7] Lecia J Barker and Kathy Garvin-Doxas. Making visible the behaviors that influence learning environment: A qualitative exploration of computer science classrooms. *Computer Science Education*, 14(2):119–145, 2004.

[8] Lecia Jane Barker, Kathy Garvin-Doxas, and Michele Jackson. Defensive climate in the computer science classroom. In *Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education*, pages 43–47, 2002.

[9] Amy S Beavers, John W Lounsbury, Jennifer K Richards, Schuyler W Huck, Gary J Skolits, and Shelley L Esquivel. Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, 18(1): 6, 2013.

[10] Brett A Becker and Keith Quille. 50 years of cs1 at sigcse: A review of the evolution of introductory programming education research. In *Proceedings of the 50th acm technical symposium on computer science education*, pages 338–344, 2019.

[11] Jens Bennedsen and Michael E Caspersen. Failure rates in introductory programming. *ACM SIGCSE Bulletin*, 39:32–36, 2007.

[12] Sylvia Beyer. 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, 2014.

[13] Maureen Biggers, Anne Brauer, and Tuba Yilmaz. Student perceptions of computer science: A retention study comparing graduating seniors with cs leavers. In *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education*, page 402–406, 2008.

[14] Lisa S. Blackwell, Kali H. Trzesniewski, and Carol S. Dweck. Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, 78(1):246–263, 2007.

[15] Jennifer M Blaney and Jane G Stout. Examining the relationship between introductory computing course experiences, self-efficacy, and belonging among first-generation college women. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, pages 69–74, 2017.

[16] Ryan Bockmon, Stephen Cooper, Jonathan Gratch, Jian Zhang, and Mohsen Dorodchi. Can students' spatial skills predict their programming abilities? In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, pages 446–451. ACM, 2020.

[17] Dennis Bouvier, Ellie Lovellette, John Matta, Jing Bai, Jacqueline Chetty, Stan Kurkovsky, and Jia Wan. Factors affecting the adoption of peer instruction in computing courses. In *Proceedings of the Working Group Reports on Global Computing Education*, pages 1–25, 2019.

[18] Nicholas A Bowman, Lindsay Jarratt, KC Culver, and Alberto Maria Segre. How prior programming experience affects students' pair programming experiences and outcomes. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, pages 170–175, 2019.

[19] John Seely Brown, Allan Collins, and Paul Duguid. Situated cognition and the culture of learning. *Educational researcher*, 18(1):32–42, 1989.

[20] Jeni Burnette, Crystal Hoyt, V Russell, Barry Lawson, Carol S. Dweck, and Eli Finkel.

A growth mind-set intervention improves interest but not academic performance in the field of computer science. *Social Psychological and Personality Science*, page 194855061984163, 2019.

[21] Nicole A Buzzetto-More, Ojiabo Ukoha, and Narendra Rustagi. Unlocking the barriers to women and minorities in computer science and information systems studies: Results from a multi-methodolical study conducted at two minority serving institutions. *Journal of Information Technology Education: Research*, 9(1):115–131, 2010.

[22] Lori Carter. Why students with an apparent aptitude for computer science don't choose to major in computer science. In *Proceedings of the 37th SIGCSE technical symposium on Computer Science Education*, pages 27–31, 2006.

[23] Michelene TH Chi and Ruth Wylie. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, 49(4):219–243, 2014.

[24] Stephen Cooper, Karen Wang, Maya Israni, and Sheryl Sorby. Spatial skills training in introductory computing. In *Proceedings of the eleventh annual International Conference on International Computing Education Research*, pages 13–20. ACM, 2015.

[25] Catherine H. Crouch and Eric Mazur. Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69, 2001.

[26] Catherine H Crouch and Eric Mazur. Peer instruction: Ten years of experience and results. *American journal of physics*, 69(9):970–977, 2001.

[27] Quintin Cutts, Sarah Esper, Marlena Fecho, Stephen R Foster, and Beth Simon. The abstraction transition taxonomy: Developing desired learning outcomes through the lens of situated cognition. In *Proceedings of the ninth annual international conference on International computing education research*, pages 63–70, 2012.

[28] Pranita Deshpande, Cynthia B Lee, and Irfan Ahmed. Evaluation of peer instruction for cybersecurity education. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 720–725, 2019.

[29] Amanda B Diekman, Emily K Clark, Amanda M Johnston, Elizabeth R Brown, and Mia Steinberg. Malleability in communal goals and beliefs influences attraction to stem careers: evidence for a goal congruity perspective. *Journal of personality and social psychology*, 101(5):902, 2011.

[30] Amanda B Diekman, Mia Steinberg, Elizabeth R Brown, Aimee L Belanger, and Emily K Clark. A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in stem gender gaps. *Personality and Social Psychology Review*, 21(2):142–175, 2017.

[31] Christine DiStefano, Min Zhu, and Diana Mindrila. Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation*, 14(1):20, 2009.

[32] Eileen Doyle, Ioanna Stamouli, and Meriel Huggard. Computer anxiety, self-efficacy, computer experience: An investigation throughout a computer science degree. In *Proceedings Frontiers in Education 35th Annual Conference*, pages S2H–3, 2005.

[33] Alberto Esquinca, Erika Mein, and Helena Mucino. Latinx students' sense of belonging in engineering/computer science at an hsi. *Proceedings of the Collaborative Network for Engineering and Computing Diversity*, 2021.

[34] Sally Fincher, Anthony Robins, Bob Baker, Ilona Box, Quintin Cutts, Michael de Raadt, Patricia Haden, John Hamer, Margaret Hamilton, Raymond Lister, et al. Predictors of success in a first programming course. In *Proceedings of the 8th Australasian Computing Education Conference (ACE 2006)*, volume 52, pages 189–196. Australian Computer Society Inc., 2006.

[35] National Center for Education Statistics. Degrees in computer and information sciences conferred by postsecondary institutions, by level of degree and sex of student: Academic years 1964-65 through 2021-22. https://nces.ed.gov/programs/digest/d23/tables/dt23_325.35.asp, 2024.

[36] National Center for Education Statistics. Bachelor's degrees conferred by postsecondary institutions, by race/ethnicity and field of study: Academic years 2020-21 and 2021-22. https://nces.ed.gov/programs/digest/d23/tables/dt23_322.30.asp, 2024.

[37] Andrea Forte and Mark Guzdial. Computers for communication, not calculation: Media as a motivation and context for learning. In *37th Annual Hawaii International Conference on System Sciences*, pages 10–pp, 2004.

[38] NSF (National Science Foundation). Science and engineering indicators. *NSF*, 2018.

[39] Kathy Garvin-Doxas and Lecia J Barker. Communication in computer science classrooms: Understanding defensive climates as a means of creating supportive behaviors. *Journal on Educational Resources in Computing (JERIC)*, 4(1):2–es, 2004.

[40] Michail N Giannakos, Ilias O Pappas, Letizia Jaccheri, and Demetrios G Sampson. Understanding student retention in computer science education: The role of environment, gains, barriers and usefulness. *Education and Information Technologies*, 22(5):2365–2382, 2017.

[41] Catherine Good, Aneeta Rattan, and Carol S Dweck. Why do women opt out? sense of belonging and women's representation in mathematics. *Journal of personality and social*

psychology, 102(4):700, 2012.

[42] Jamie Gorson and Eleanor O'Rourke. Why do cs1 students think they're bad at programming? investigating self-efficacy and self-assessments at three universities. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, pages 170–181, 2020.

[43] J.G. Greeno, D.R. Smith, and J.L. Moore. Transfer of situated learning. In D.K. Detterman and R.J. Sternberg, editors, *Transfer on trial: intelligence, cognition, and instructionh*, pages 99–167. 1993.

[44] Mark Guzdial. A media computation course for non-majors. In *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, pages 104–108, 2003.

[45] Mark Guzdial. Exploring hypotheses about media computation. In *Proceedings of the 9th Annual ACM Conference on International Computing Education Research*, pages 19–26, 2013.

[46] Mark Guzdial. Computing for other disciplines. In Sally A Fincher and Anthony V Robins, editors, *The Cambridge Handbook of Computing Education Research*, chapter 19, pages 584–605. Cambridge University Press, 2019.

[47] Mark Guzdial and Barbara Ericson. *Introduction to computing & programming in Java: a multimedia approach*. Pearson Prentice Hall, 2007.

[48] Mark Guzdial and Andrea Forte. Design process for a non-majors computing course. In *Proceedings of the 36th ACM Technical Symposium on Computer Science Education*, pages 361–365, 2005.

[49] Dianne Hagan and Selby Markham. Does it help to have some programming experience before beginning a computing degree program? In *Proceedings of the 5th annual SIGCSE/SIGCUE ITiCSEconference on Innovation and technology in computer science education*, pages 25–28, 2000.

[50] Bonnie M Hagerty and Arthur Williams. The effects of sense of belonging, social support, conflict, and loneliness on depression. *Nursing research*, 48(4):215–219, 1999.

[51] Brian Hanks, Sue Fitzgerald, Renée McCauley, Laurie Murphy, and Carol Zander. Pair programming in education: A literature review. *Computer Science Education*, 21(2): 135–173, 2011.

[52] Charles Henderson and Melissa H Dancy. Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical*

Review Special Topics-Physics Education Research, 3(2):020102, 2007.

[53] Geoffrey L Herman, Craig Zilles, and Michael C Loui. A psychometric evaluation of the digital logic concept inventory. *Computer Science Education*, 24(4):277–303, 2014.

[54] Aleata Hubbard Cheuoua. *Confronting Inequities in Computer Science Education: A Case for Critical Theory*, page 425–430. Association for Computing Machinery, New York, NY, USA, 2021. ISBN 9781450380621. URL https://doi.org/10.1145/3408877.3432453.

[55] Janet Hughes and D Ramanee Peiris. Assisting cs1 students to learn: learning approaches and object-oriented programming. *ACM Sigcse Bulletin*, 38(3):275–279, 2006.

[56] Sue Jones and Gary Burnett. Spatial ability and learning to program. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 4(1):47–61, 2008. ISSN 1795-6889. doi: 10.17011/ht/urn.200804151352.

[57] Rhody Kaner and Eitan Frachtenberg. Experience and representation of gender minorities in undergraduate computer science. Technical report, Technical Report. EasyChair, 2020.

[58] Alboukadel Kassambara. Kruskal-wallis effect size. https://rpkgs.datanovia.com/rstatix/reference/kruskal_effsize.html, 2021. Accessed: 2021-03-23.

[59] Päivi Kinnunen and Beth Simon. Cs majors' self-efficacy perceptions in cs1: results in light of social cognitive theory. In *Proceedings of the seventh international workshop on Computing education research*, pages 19–26, 2011.

[60] Sophia Krause-Levy, Leo Porter, Beth Simon, and Christine Alvarado. Investigating the impact of employing multiple interventions in a cs1 course. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, page 1082–1088, 2020.

[61] Sophia Krause-Levy, Sander Valstar, Leo Porter, and William G. Griswold. Exploring the link between prerequisites and performance in advanced data structures. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, page 386–392, 2020.

[62] Sophia Krause-Levy, William G Griswold, Leo Porter, and Christine Alvarado. The relationship between sense of belonging and student outcomes in cs1 and beyond. In *Proceedings of the 17th ACM Conference on International Computing Education Research*, pages 29–41, 2021.

[63] Sophia Krause-Levy, William Griswold G., Leo Porter, and Christine Alvarado. The relationship between sense of belonging and student outcomes in cs1 and beyond. In *Proceedings of the 2021 ACM Conference on International Computing Education Research*, 2021.

[64] Justin Kruger and David Dunning. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6):1121, 1999.

[65] Daehan Kwak, Patricia Morreale, Sarah T Hug, Yulia Kumar, Jean Chu, Ching-Yu Huang, J Jenny Li, and Paoline Wang. Evaluation of the use of growth mindset in the cs classroom. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1*, pages 878–884, 2022.

[66] Wayne W. LaMorte. The social cognitive theory. https://sphweb.bumc.bu.edu/otlt/mph-modules/sb/behavioralchangetheories/behavioralchangetheories5.html, 2019. Accessed: 2021-03-20.

[67] Celine Latulipe, Audrey Rorrer, and Bruce Long. Longitudinal data on flipped class effects on performance in cs1 and retention after cs1. In *Proceedings of the 49th ACM technical symposium on computer science education*, pages 411–416, 2018.

[68] Jean Lave and Etienne Wenger. *Situated learning: Legitimate peripheral participation*. Cambridge university press, 1991.

[69] Jean Lave, Etienne Wenger, et al. *Situated learning: Legitimate peripheral participation*. Cambridge university press, 1991.

[70] Cynthia Bailey Lee, Saturnino Garcia, and Leo Porter. Can peer instruction be effective in upper-division computer science courses? *Transactions on Computing Education*, 13(3), August 2013.

[71] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. Alignment of goals and perceptions of computing predicts students' sense of belonging in computing. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, pages 11–19, 2019.

[72] Colleen M Lewis and Niral Shah. How equity and inequity can emerge in pair programming. In *Proceedings of the 11th Annual ACM Conference on International Computing Education Research*, pages 41–50, 2015.

[73] Soohyun Nam Liao, Daniel Zingaro, Michael A Laurenzano, William G Griswold, and Leo Porter. Lightweight, early identification of at-risk cs1 students. In *Proceedings of the 12th Annual ACM Conference on International Computing Education Research*, pages 123–131, 2016.

[74] Soohyun Nam Liao, Sander Valstar, Kevin Thai, Christine Alvarado, Daniel Zingaro, William G Griswold, and Leo Porter. Behaviors of higher and lower performing students in cs1. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in*

Computer Science Education, pages 196–202, 2019.

[75] Soohyun Nam Liao, Daniel Zingaro, Christine Alvarado, William G Griswold, and Leo Porter. Exploring the value of different data sources for predicting student performance in multiple cs courses. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 112–118, 2019.

[76] Soohyun Nam Liao, Daniel Zingaro, Kevin Thai, Christine Alvarado, William G Griswold, and Leo Porter. A robust machine learning technique to predict low-performing students. *Transactions on Computing Education*, 19(3):1–19, 2019.

[77] Alex Lishinski and Joshua Rosenberg. All the pieces matter: The relationship of momentary self-efficacy and affective experiences with cs1 achievement and interest in computing. In *Proceedings of the 17th ACM Conference on International Computing Education Research*, pages 252–265, 2021.

[78] Alex Lishinski, Aman Yadav, Richard Enbody, and Jon Good. The influence of problem solving abilities on students' performance on different assessment tasks in cs1. In *Proceedings of the 47th ACM technical symposium on computing science education*, pages 329–334, 2016.

[79] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proceedings of the 2016 ACM Conference on International Computing Education Research*, pages 211–220, 2016.

[80] Janet Lowe and Vernon Gayle. Exploring the work/life/study balance: the experience of higher education students in a scottish further education college. *Journal of further and Higher Education*, 31(3):225–238, 2007.

[81] Andrew Luxton-Reilly. Learning to program is easy. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education*, pages 284–289, 2016.

[82] Andrew Luxton-Reilly, Ibrahim Albluwi, Brett A Becker, Michail Giannakos, Amruth N Kumar, Linda Ott, James Paterson, Michael James Scott, Judy Sheard, and Claudia Szabo. Introductory programming: a systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, pages 55–106, 2018.

[83] Andrew Luxton-Reilly, Vangel V Ajanovski, Eric Fouh, Christabel Gonsalvez, Juho Leinonen, Jack Parkinson, Matthew Poole, and Neena Thota. Pass rates in introductory programming and in other stem disciplines. In *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education*, pages 53–71. 2019.

[84] Anna Ly, Jack Parkinson, Quintin Cutts, Michael Liut, and Andrew Petersen. Spatial skills and demographic factors in CS1. In *Koli Calling*. ACM, 2021. doi: 10.1145/3488042.3488049.

[85] Lauri Malmi, Judy Sheard, Päivi Kinnunen, and Jane Sinclair. Computing education theories: What are they and how are they used? In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, pages 187–197, 2019.

[86] Jane Margolis and Allan Fisher. *Unlocking the clubhouse: Women in computing*. MIT press, 2002.

[87] Jane Margolis, Rachel Estrella, Joanna Goode, Jennifer Jellison Holme, and Kim Nao. *Stuck in the shallow end: Education, race, and computing*. MIT press, 2017.

[88] Lauren E. Margulieux. Spatial encoding strategy theory: The relationship between spatial skill and stem achievement. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, ICER '19, page 81–90, New York, NY, USA, 2019. Association for Computing Machinery.

[89] Allison Master and Andrew N Meltzoff. Cultural stereotypes and sense of belonging contribute to gender gaps in stem. *International Journal of Gender, Science and Technology*, 12(1):152–198, 2020.

[90] Allison Master, Sapna Cheryan, and Andrew N Meltzoff. Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of educational psychology*, 108(3):424, 2016.

[91] Charlie McDowell, Linda Werner, Heather E Bullock, and Julian Fernald. Pair programming improves student retention, confidence, and program quality. *Communications of the ACM*, 49(8):90–95, 2006.

[92] Rodrigo Pessoa Medeiros, Geber Lisboa Ramalho, and Taciana Pontual Falcão. A systematic literature review on teaching and learning introductory programming in higher education. *IEEE Transactions on Education*, 62(2):77–90, 2018.

[93] Irene T Miura. The relationship of computer self-efficacy expectations to computer interest and course enrollment in college. *Sex roles*, 16(5-6):303–311, 1987.

[94] Catherine Mooney, Anna Antoniadi, Ioannis Karvelas, Lána Salmon, and Brett A Becker. Exploring sense of belonging in computer science students. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, pages 563–563, 2020.

[95] Sukanya Kannan Moudgalya, Chris Mayfield, Aman Yadav, Helen H Hu, and Clif Kuss-

maul. Measuring students' sense of belonging in introductory cs courses. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, pages 445–451, 2021.

[96] Jonathan P Munson and Joshua P Zitovsky. Models for early identification of struggling novice programmers. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, pages 699–704, 2018.

[97] Laurie Murphy and Lynda Thomas. Dangers of a fixed mindset: implications of self-theories research for computer science education. In *Proceedings of the 13th annual conference on Innovation and technology in computer science education*, pages 271–275, 2008.

[98] Nachiappan Nagappan, Laurie Williams, Miriam Ferzli, Eric Wiebe, Kai Yang, Carol Miller, and Suzanne Balik. Improving the cs1 experience with pair programming. In *Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '03, page 359–362, New York, NY, USA, 2003. Association for Computing Machinery.

[99] National Academies of Sciences, Engineering, and Medicine and others. *Assessing and responding to the growth of computer science undergraduate enrollments*. National Academies Press, 2018.

[100] National Center for Science and Engineering Statistics. Women, minorities, and persons with disabilities in science and engineering: Special report NSF 19-340. 2019.

[101] Mary Natrella. NIST/SEMATECH e-handbook of statistical methods. http://www.itl.nist.gov/div898/handbook, 2010.

[102] An Nguyen and Colleen M Lewis. Competitive enrollment policies in computing departments negatively predict first-year students' sense of belonging, self-efficacy, and perception of department. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pages 685–691, 2020.

[103] Mogens Allan Niss. The concept and role of theory in mathematics education: Plenary presentation. In *Relating Practice and Research in Mathematics Education: Proceedings of NORMA 05, Fourth Nordic Conference on Mathematics Education*, pages 97–110. TAPIR Akademisk Forlag, 2007.

[104] University of California San Diego Computer Science and Engineering (CSE) Department. Computer Science and Engineering (CSE) Course Catalog. https://catalog.ucsd.edu/courses/CSE.html, 2024.

[105] U.S. Bureau of Labor Statistics. Occupational Outlook Handbook - Com-

puter and Information Technology Occupations. https://www.bls.gov/ooh/computer-and-information-technology/home.htm, 2024.

[106] Ilias O Pappas, Michail N Giannakos, Letizia Jaccheri, and Demetrios G Sampson. Assessing student behavior in computer science education with an fsqca approach: The role of gains and barriers. *ACM Transactions on Computing Education (TOCE)*, 17(2): 1–23, 2017.

[107] Jack Parkinson and Quintin Cutts. Investigating the relationship between spatial skills and computer science. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, ICER '18, page 106–114, New York, NY, USA, 2018. Association for Computing Machinery.

[108] Arnold Pears, Stephen Seidman, Lauri Malmi, Linda Mannila, Elizabeth Adams, Jens Bennedsen, Marie Devlin, and James Paterson. A survey of literature on the teaching of introductory programming. *Working group reports on ITiCSE on Innovation and technology in computer science education*, pages 204–223, 2007.

[109] Andrew Petersen, Michelle Craig, Jennifer Campbell, and Anya Tafliovich. Revisiting why students drop cs1. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research*, pages 71–80, 2016.

[110] Leo Porter and Beth Simon. Retaining nearly one-third more majors with a trio of instructional best practices in CS1. In *Proceedings of the 44th Special Interest Group on Computer Science Education Technical Symposium*, 2013.

[111] Leo Porter and Beth Simon. Retaining nearly one-third more majors with a trio of instructional best practices in cs1. In *Proceeding of the 44th ACM technical symposium on Computer science education*, pages 165–170, 2013.

[112] Leo Porter and Beth Simon. A Case Study of Peer Instruction: From University of California, San Diego to the Computer Science Community. In Sally A Fincher and Anthony V Robins, editors, *The Cambridge Handbook of Computing Education Research*, chapter 30, pages 861–874. Cambridge University Press, 2019.

[113] Leo Porter, Cynthia Bailey Lee, Beth Simon, and Daniel Zingaro. Peer instruction: Do students really learn from peer discussion in computing? In *Proceedings of the 7th Annual ACM Conference on International Computing Education Research*, 2011.

[114] Leo Porter, Saturnino Garcia, John Glick, Andrew Matusiewicz, and Cynthia Taylor. Peer instruction in computer science at small liberal arts colleges. In *Proceedings of the 18th Annual Conference on Innovation and Technology in Computer Science Education*, 2013.

[115] Leo Porter, Cynthia Bailey Lee, and Beth Simon. Halving fail rates using peer instruction:

A study of four computer science courses. In *Proceedings of the 44th Special Interest Group on Computer Science Education Technical Symposium*, 2013.

[116] Leo Porter, Daniel Zingaro, and Raymond Lister. Predicting student success using fine grain clicker data. In *Proceedings of the 10th Annual ACM Conference on International Computing Education Research*, pages 51–58, 2014.

[117] Leo Porter, Dennis Bouvier, Quintin Cutts, Scott Grissom, Cynthia Lee, Robert McCartney, Daniel Zingaro, and Beth Simon. A multi-institutional study of peer instruction in introductory computing. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pages 358–363, 2016.

[118] Yizhou Qian and James Lehman. Students' misconceptions and other difficulties in introductory programming: A literature review. *ACM Transactions on Computing Education (TOCE)*, 18(1):1–24, 2017.

[119] Keith Quille and Susan Bergin. Cs1: how will they do? how can we help? a decade of research and practice. *Computer Science Education*, 29(2-3):254–282, 2019.

[120] Keith Quille and Susan Bergin. Promoting a growth mindset in cs1: Does one size fit all? a pilot study. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, pages 12–18, 2020.

[121] Lauren Rich, Heather Perry, and Mark Guzdial. A cs1 course designed to address interests of women. *Acm sigcse bulletin*, 36(1):190–194, 2004.

[122] Anthony Robins. Learning edge momentum: A new account of outcomes. *Computer Science Education*, 20(1):37–71, 2010.

[123] Sarah L Rodriguez and Jennifer M Blaney. "we're the unicorns in stem": Understanding how academic and social experiences influence sense of belonging for latina undergraduate students. *Journal of Diversity in Higher Education*, 14(3):441, 2021.

[124] Adam Rosenstein, Aishma Raghu, and Leo Porter. Identifying the prevalence of the impostor phenomenon among computer science students. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pages 30–36, 2020.

[125] Nathan Rountree, Janet Rountree, and Anthony Robins. Predictors of success and failure in a cs1 course. *ACM SIGCSE Bulletin*, 34(4):121–124, 2002.

[126] Nathan Rountree, Janet Rountree, Anthony Robins, and Robert Hannah. Interacting factors that predict success and failure in a cs1 course. *ACM SIGCSE Bulletin*, 36(4): 101–104, 2004.

[127] Merilin Säde, Reelika Suviste, Piret Luik, Eno Tõnisson, and Marina Lepp. Factors that influence students' motivation and perception of studying computer science. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 873–878, 2019.

[128] Adrian Salguero, Julian McAuley, Beth Simon, and Leo Porter. A longitudinal evaluation of a best practices cs1. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, pages 182–193, 2020.

[129] Adrian Salguero, William G Griswold, Christine Alvarado, and Leo Porter. Understanding sources of student struggle in early computer science courses. In *Proceedings of the 17th ACM Conference on International Computing Education Research*, pages 319–333, 2021.

[130] Norsaremah Salleh, Emilia Mendes, and John Grundy. Empirical studies of pair programming for CS/SE teaching in higher education: a systematic literature review. *IEEE Transactions on Software Engineering*, 37(4):509–525, 2010.

[131] Linda J Sax, Jennifer M Blaney, Kathleen J Lehman, Sarah L Rodriguez, Kari L George, and Christina Zavala. Sense of belonging in computing: The role of introductory courses for women and underrepresented minority students. *Social Sciences*, 7(8):122, 2018.

[132] Carsten Schulte and Maria Knobelsdorf. Attitudes towards computer science-computing experiences as a starting point and barrier to computer science. In *Proceedings of the third international workshop on Computing education research*, pages 27–38, 2007.

[133] Walter R Schuman. Applied multivariate statistics for the social sciences. *The American Statistician*, 47(2):155–157, 1993.

[134] Stanley T Schuyler, Robert Joseph Skovira, and AJ Grant. The role of learning technique on student performance in cs1 courses. *Issues in information Systems*, pages 60–67, 2008.

[135] Duane F Shell, Leen-Kiat Soh, Abraham E Flanigan, and Markeya S Peteranetz. Students' initial course motivation and their achievement and retention in college cs1 courses. In *Proceedings of the 47th ACM technical symposium on computing science education*, pages 639–644, 2016.

[136] Beth Simon, Päivi Kinnunen, Leo Porter, and Dov Zazkis. Experience report: Cs1 for majors with media computation. In *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education*, pages 214–218, 2010.

[137] Beth Simon, Sarah Esper, Leo Porter, and Quintin Cutts. Student experience in a student-centered peer instruction classroom. In *Proceedings of the 9th Annual ACM Conference on International Computing Education Research*, 2013.

[138] Beth Simon, Julian Parris, and Jaime Spacco. How we teach impacts learning: peer instruction vs. lecture in CS0. In *Proceedings of the 44th ACM Technical Symposium on Computer Science Education*, 2013.

[139] Beth Simon, Julian Parris, and Jaime Spacco. How we teach impacts student learning: Peer instruction vs. lecture in cs0. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, SIGCSE '13, page 41–46, New York, NY, USA, 2013. Association for Computing Machinery.

[140] Beth Simon, Christopher Hundhausen, Charlie McDowell, Linda Werner, Helen Hu, , and Clif Kussmaul. Students as teachers and communicators. In Sally A Fincher and Anthony V Robins, editors, *The Cambridge Handbook of Computing Education Research*, chapter 29, pages 827–857. Cambridge University Press, 2019.

[141] Simon Simon, Sally Fincher, Anthony Robins, Bob Baker, IIona Box, Quintin Cutts, Michael De Raadt, Patricia Haden, John Hamer, Margaret Hamilton, et al. Predictors of success in a first programming course. In *Conferences in Research and Practice in Information Technology*, volume 52, pages 189–196. Australian Computer Society, 2006.

[142] Robert H Sloan and Patrick Troy. CS 0.5: a better approach to introductory computer science for majors. In *Proceedings of the 39th ACM Technical Symposium on Computer Science Education*, pages 271–275, 2008.

[143] Jessi L Smith, Erin Cech, Anneke Metz, Meghan Huntoon, and Christina Moyer. Giving back or giving up: Native american student experiences in science and engineering. *Cultural Diversity and Ethnic Minority Psychology*, 20(3):413, 2014.

[144] Tamara Floyd Smith, Denise Wilson, Diane Carlson Jones, Melani Plett, Rebecca A. Bates, and Nanette M Veilleux. Investigation of belonging for engineering and science undergraduates by year in school. In *2012 ASEE Annual Conference & Exposition*, 2012.

[145] Justin M Sprung and Anna Rogers. Work-life balance as a predictor of college student anxiety and depression. *Journal of American college health*, pages 1–8, 2020.

[146] Ben Stephenson, Michelle Craig, Daniel Zingaro, Diane Horton, Danny Heap, and Elaine Huynh. Exam wrappers: Not a silver bullet. In *Proceedings of the 48th ACM Technical Symposium on Computer Science Education*, pages 573–578, 2017.

[147] Leslie S Stratton, Dennis M O'Toole, and James N Wetzel. A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, 27(3):319–331, 2008.

[148] Mohsen Tavakol and Reg Dennick. Making sense of cronbach's alpha. *International Journal of Medical Education*, 2:53, 2011.

[149] Allison Elliott Tew, Charles Fowler, and Mark Guzdial. Tracking an innovation in introductory cs education from a research university to a two-year college. In *Proceedings of the 36th ACM Technical Symposium on Computer Science Education*, pages 416–420, 2005.

[150] Elli J Theobald, Mariah J Hill, Elisa Tran, Sweta Agrawal, E Nicole Arroyo, Shawn Behling, Nyasha Chambwe, Dianne Laboy Cintrón, Jacob D Cooper, Gideon Dunster, et al. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12):6476–6483, 2020.

[151] Dustin B Thoman, Elizabeth R Brown, Andrew Z Mason, Allen G Harmsen, and Jessi L Smith. The role of altruistic values in motivating underrepresented minority students for biomedicine. *BioScience*, 65(2):183–188, 2015.

[152] Maciej Tomczak and Ewa Tomczak. The need to report effect size estimates revisited. an overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 2014.

[153] Karthikeyan Umapathy and Albert D Ritzhaupt. A meta-analysis of pair-programming in computer programming courses: Implications for educational practice. *ACM Transactions on Computing Education*, 17(4):1–13, 2017.

[154] Sander Valstar, William G. Griswold, and Leo Porter. The relationship between prerequisite proficiency and student performance in an upper-division computing course. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, page 794–800, 2019.

[155] Nanette Veilleux, Rebecca Bates, Cheryl Allendoerfer, Diane Jones, Joyous Crawford, and Tamara Floyd Smith. The relationship between belonging and ability in computer science. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, pages 65–70, 2013.

[156] Lev Semenovich Vygotsky. *Mind in society: The development of higher psychological processes*. Harvard University Press, 1980.

[157] Christopher Watson and Frederick WB Li. Failure rates in introductory programming revisited. In *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education*, pages 39–44, 2014.

[158] Chris Wilcox and Albert Lionelle. Quantifying the benefits of prior programming experience in an introductory computer science course. In *Proceedings of the 49th acm technical symposium on computer science education*, pages 80–85, 2018.

[159] Laurie Williams and Robert Kessler. *Pair programming illuminated*. Addison-Wesley Longman Publishing Co., Inc., 2002.

[160] Brenda Cantwell Wilson. A study of factors promoting success in computer science including gender differences. *Computer Science Education*, 12(1-2):141–164, 2002.

[161] Brenda Cantwell Wilson and Sharon Shrock. Contributing to success in an introductory computer science course: a study of twelve factors. In *Proceedings of the 32nd SIGCSE Technical Symposium on Computer Science Education*, pages 184–188, 2001.

[162] T. D. Wilson, M. Damiani, and N. Shelton. Improving the academic performance of college students with brief attributional interventions. In J. Aronson, editor, *Improving Academic Achievement: Impact of Psychological Factors on Education*, chapter 5, pages 91–110. Academic Press, 2002.

[163] Melissa R Witkow, Virginia Huynh, and Andrew J Fuligni. Understanding differences in college persistence: A longitudinal examination of financial circumstances, family obligations, and discrimination in an ethnically diverse sample. *Applied Developmental Science*, 19(1):4–18, 2015.

[164] K.L. Wuensch. Standardized effect size esitmation: Why and how? http://core.ecu.edu/psyc/wuenschk/StatHelp/Effect%20Size%20Estimation.pdf, 2015. Accessed: 2020-04-02.

[165] Hanna Yakymova, Yoann Monteiro, and Daniel Zingaro. Study strategies and exam grades in cs1. In *Proceedings of the 21st Western Canadian Conference on Computing Education*, pages 1–3, 2016.

[166] Hans Yuan and Yingjun Cao. Hybrid pair programming-a promising alternative to standard pair programming. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 1046–1052, 2019.

[167] Daniel Zingaro. Peer instruction contributes to self-efficacy in CS1. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 373–378, 2014.

[168] Daniel Zingaro and Leo Porter. Peer instruction in computing: The value of instructor intervention. *Computers and Education*, 71, 2014.

[169] Daniel Zingaro and Leo Porter. Impact of student achievement goals on cs1 outcomes. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pages 279–296, 2016.

[170] Daniel Zingaro, Cynthia Bailey Lee, and Leo Porter. Peer instruction in computing: the role of reading quizzes. In *Proceeding of the 44th ACM Technical Symposium on*

Computer Science Education, pages 47–52, 2013.

[171] Daniel Zingaro, Michelle Craig, Leo Porter, Brett A Becker, Yingjun Cao, Phill Conrad, Diana Cukierman, Arto Hellas, Dastyni Loksa, and Neena Thota. Achievement goals in cs1: Replication and extension. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, pages 687–692, 2018.

[172] Stuart Zweben. Enrollment and retention in u.s. computer science bachelor's programs in 2016-17. *ACM Inroads*, 10(4):47–59, November 2019.