SAIL: A System for Adaptive Interest-based Learning in STEM Education

by

Karen Elizabeth Aguar

(Under the Direction of Hamid Arabnia and Juan Gutierrez)

Abstract

The aim of this research is to alleviate many challenges faced in STEM education through the creation of a scalable, adaptive learning framework that supports interest-based learning (IBL) in multiple domains. Adaptive Learning is the idea that software and material should "adapt" to individual student's needs, typically based on previous knowledge, pace, or learning style. This research takes a less explored approach by adapting content and practice problems based on a student's interests. Interest-based learning (IBL) has been shown to improve intrinsic motivation, leading to better learning and achievements, but no solution currently exists to facilitate and promote IBL across multiple domains. This work presents the design and pilot of SAIL, a System for Adaptive Interest-based Learning, to easily facilitate IBL in an adaptive and scalable platform. SAIL is not limited by domain, but was designed with STEM subjects in mind due to their high applicability in other fields. With SAIL, one student in an introductory programming course could practice loops through sports-themed examples while another could learn through music or science. SAIL was designed to help alleviate many of the concerns in STEM education by providing a competent and compelling curriculum delivering individualized instruction to help increase motivation, performance and fill the gaps in STEM education. SAIL showcases the interconnectivity of STEM subjects with other fields, combatting misperceptions and increasing motivation to help attract and retain a larger and more diverse population of students. With SAIL, students become active participants in their learning experience as they utilize an interactive map to traverse their unique path through interest-based course material. A large pilot study (N=307) in the context of introductory programming (Java) was conducted comparing a class using SAIL to three other classes with varying control conditions. This study resulted in new quantitative and qualitative knowledge about how SAIL can impact introductory Computer Science (CS) as well as assessing viability for other STEM fields, including K-12 STEM education. Via SAIL, we raise the standard of education, increase enjoyment, remedy gender disparities, and aid in encouraging more students to continue their CS education.

Index words: Interest-based Learning, STEM Education, Computer Science Education, Individualized Instruction, Personalized Context, Adaptive Learning, eLearning, Community Knowledge Sharing

SAIL: A System for Adaptive Interest-based Learning in STEM Education

by

Karen Elizabeth Aguar

B.S., Georgia College & State University, 2012

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2017

SAIL: A System for Adaptive Interest-based Learning in STEM Education

by

Karen Elizabeth Aguar

Approved:

Major Professors: Hamid Arabnia
Juan Gutierrez

Committee: Walter Potter
Thiab Taha

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

This dissertation is dedicated to my family. This achievement would not be possible without your love and support.

I cannot begin to explain the love, support, and encouragement that has brought me to this point. In the darkest parts of this journey, someone would always uplift me, guide me, cheer me on, and help me persevere. Thank you, to each and every one of you who have inspired and encouraged me along the way.

Thank you to my wonderful siblings, nieces and nephews, and in-laws who have poured out so much love and encouragement along the way. Thank you for excusing me from the many absences during this journey, for supporting me, celebrating with me, listening to me, and loving me.

To my incredible parents - you instilled in me a confidence that I could do anything I set my mind to. Thank you for believing in me, crying with me, laughing with me, and being my biggest cheerleaders.

Lastly, to my amazing husband - You've wiped my tears, shared in my excitement, and been my rock. Thank you for never doubting me, even when I doubted myself. I could not have made it here without you.

Acknowledgments

First and foremost, I have to thank my incredible advisor, Hamid Arabnia. You have been there with me through every step of my Ph.D. journey, and I truly believe I would not have made it through graduate school without your encouragement and support. You are one of the kindest and most selfless people I have ever known - consistently putting other's needs before your own, and I am honored to have been your Ph.D. student.

Juan Guttierez, you have played a pivotal role in my research path. Thank you for taking a chance on me, and for all of your generosity, encouragement, and support in the development of SAIL.

Walter (Don) Potter - Aside from Dr. Arabnia, who I had met previously, you were the first professor I met during my graduate studies who made me feel like I belonged and could accomplish whatever I set my mind to. In studying how many women are deterred from pursuing STEM degrees by a lack of confidence and not fitting in, it is support from those like you that boost confidence and allow students to achieve their full potential. Thank you for taking an interest in me, for supporting me, and for helping shape my research path.

Thiab Taha - you have been not only a supportive committee member, but also an incredible boss during my time at UGA. Thank you for giving me the opportunity to teach, as that is what inspires my research. The experience I have gained while teaching has been invaluable, and I cannot wait until I can teach again.

Brad Barnes - you may not be my advisor, but you have made this journey possible in so many ways. Thank you for all of your advice along the way, for helping me learn how to be a new instructor, for always having my back, and most recently, for helping me implement these experiments. You have been a great ally and friend throughout my pursuit of research and time teaching.

Gita Phelps - Having you as my very first computer science professor has had such an enormous impact on my life. I truly believe that it was your passion and conveyance of the usefulness and beauty of CS that made me choose CS as my major and embark on my current path. Thank you for being an excellent and inspiring educator.

Special thanks to my peers and friends in the department who have supported me in various ways throughout my time here. Special thanks to Saeid, Delaram, Mehdi, and Hao. You have each played a huge role in the development of this research and the outcomes of this work. Without your time and support, this work would not have been possible.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Low enrollment, low retention, and a lack of diversity are among the many problems faced by Science, Technology, Engineering and Mathematics (STEM) education. With high growth rates in STEM careers and numerous unfilled positions, attracting and retaining a larger and more diverse population of students to study STEM subjects has become an important area of research. Computer Science (CS) is at the center of this trend, accounting for 71% of new STEM jobs, but only 8% of STEM graduates [1]. These trends have sparked efforts worldwide to improve the way introductory computer science is taught.

This research aims to alleviate many challenges faced by STEM fields such as CS through the creation of a scalable, adaptive framework that supports interest-based individualized instruction. Individualized (or personalized) instruction is the concept that instruction and/or materials should be customized to the unique needs of each student. Personalizing instruction manually has been an age-old practice in education – for example, if a student was struggling with addition, the teacher may assign extra homework problems. In the last several decades, utilizing technology to individualize instruction via adaptive learning systems (ALS) has been a widely studied research area. Adaptive learning systems have been shown to improve student performance, usually personalizing instruction based on a student's previous knowledge, pace, or learning style [2]. This research takes an alternative, less-explored approach by enabling the adaptation of content, practice problems, and examples based on a student's interests.

The incorporation of personal interest into learning has been shown to increase intrinsic motivation and provide positive learning outcomes [3], with most studies implemented manually. Initial studies using adaptive learning systems as a medium for interest-based learning (IBL) have indicated tremendous potential [4], with most work done in the field of mathematics and all systems limited by domain. Little research has been done to support how IBL can impact other fields, such as computer science, and there are very few technologies developed to help facilitate personalization based on interest. Though personalizing STEM education based on interest may provide many benefits such as attracting a larger and more diverse population of students, the lack of technologies to empower IBL combined with the enormous effort required to create adaptive content hinder success. To address these limitations, we propose SAIL: a System for Adaptive Interest-based Learning in STEM education.

1.2 Our Solution

SAIL, a System for Adaptive Interest-based Learning, is a novel solution to address many of the issues in widespread STEM education. SAIL is a scalable, web-based adaptive learning framework designed to empower interest-based learning (IBL) to help achieve better attraction, diversity, retention, and a standard of competency in STEM education. SAIL is not limited by domain, but was designed with STEM subjects in mind due to their high applicability in other fields. The vision is to provide a framework where educators from multiple domains can contribute to and access adaptive content available to the community to provide students with an improved, individualized learning experience. IBL via SAIL could showcase the interconnectivity of STEM subjects with other fields, helping to combat misperceptions, increase motivation, elevate performance, and fill the gaps in STEM education.

SAIL was inspired by ALICE (Adaptive Learning for Interdisciplinary Collaborative Environments), an adaptive learning system designed for interdisciplinary instruction and currently in testing at the University of Georgia (Chapter 3 describes the design of ALICE and overviews the ongoing pilot studies). ALICE helps bring together students from a variety of backgrounds to work together on complex, interdisciplinary problems. SAIL harnesses this interdisciplinary approach to instruction while significantly expanding the adaptive framework and vision to alleviate challenges in a broader context of STEM education through IBL. SAIL showcases the interconnectivity of STEM fields with the world around us by grounding often abstract concepts in contexts motivating and interesting to each student. SAIL transforms students into active participants in their learning experience as they utilize an interactive knowledge map to traverse a unique path through interest-based course material. With SAIL, one student could practice math skills through sports-themed examples, while another could receive practice problems in a context more aligned with their own interests. This combines the anchoring of problems in real-world domains, with the inspiration of intrinsically motivating contexts to help increase performance, correct misperceptions, and encourage more students to take an interest in STEM studies. Though SAIL was designed with STEM subjects in mind, the overall construct can be applied to any discipline with broad applicability in other domains.

To begin testing the viability and impact of SAIL, we explore SAIL's impact in introductory level computer science (CS). Computer science suffers from many of the challenges seen throughout STEM education such as low-enrollment, a lack of diversity, a misperception that CS is "nerdy" and solely about technology, and a lack of qualified instructors. However, many studies have shown that adjusting the way introductory-level CS courses are taught can help increase diversity, attraction, and retention. Early exposure to research projects and breadth-first approaches have been among the most successful, as they advertise CS's ability to impact other disciplines.

We hypothesize that SAIL can help accentuate the high applicability of CS and, as evidenced by these earlier studies, attract a larger and more diverse population of students. In addition to advertising the applicability of CS, we hypothesize that the incorporation of interest via SAIL into introductory programming could lead to both increased performance and increased enjoyment in the course by helping influence intrinsically motivating factors.

To test this hypothesis, we designed the initial SAIL framework and piloted it in an introduction to programming (Java I) course at the University of Georgia. This six-week pilot involved 307 students from four classes. The treatment group, utilizing SAIL, was compared to three other groups with varying control conditions. The three controls allowed us to compare the treatment group with (1) a control with the same seasoned instructor and same flipped-classroom instructional design and (2) two control groups with separate instructors in a traditional lecture-style classroom. Professional quality modules and interest-based exercises were created for the pilot study and began the community knowledge sharing of content in SAIL for Computer Science (SAIL-CS). Quantitative, qualitative, and demographic data were collected to pursue the following research questions:

- **RQ 1 - How does the use of SAIL impact performance measures?** Based on prior studies, we hypothesize that interest-based learning via SAIL would enhance performance when compared to traditional instruction.

- **RQ 2 - How does the use of SAIL impact perceived learning and confidence in the course material?**

  In addition to assessing a student's performance measures, we look at student perceived learning to gauge how students feel about course content. This is important as studies have shown that a lack of confidence has been a hindrance to women and other minorities pursuing STEM fields. By grounding abstract concepts in personally interesting contexts, we hypothesize that SAIL will have a positive effect on student confidence in understanding and capabilities and in turn may positively affect who pursues the field.

- **RQ 3 - How does the use of SAIL influence students' attitudes and perception towards computer science (or STEM fields)?**

We hypothesize that IBL via SAIL can increase student enjoyment and interest in Computer Science. Based on past studies of minority attraction and retention in CS, we believe that SAIL can help showcase the interconnectivity of CS (or other STEM subjects) with the world around us to help remedy the misperception associated with many STEM fields and eventually, lead to better attraction and retention rates.

- **RQ 4 - How did students perceive the overall experience of SAIL when compared to the traditional mode of instruction?**

  We believe students will enjoy the active role assumed by interacting with SAIL and selecting their own interest. Additionally, we believe these interest-based exercises can help provide grounded contexts that motivate the student to learn the subject. Overall, we believe SAIL can provide the student with a positive and impactful learning experience.

- **RQ 5 - Does the impact of SAIL on (1)performance, (2)confidence, and (3)perception differ based on diversity factors such as gender, ethnicity, or prior exposure to CS?**

  As minority attraction and retention are astounding issues faced by the CS and STEM communities, many of these research questions were pursued to investigate how these measures differ among various groups and minorities. We hypothesize that SAIL would provide overall increased learning outcomes regardless of race or gender. We, however, hypothesize that SAIL would help inspire confidence in female and minority students as well as students with no prior exposure to CS through IBL. We also believe SAIL can help promote enjoyment and interest in CS among these underrepresented groups.

- **RQ 6 - Does SAIL demonstrate a potential to impact attraction and recruitment to STEM disciplines?** We hypothesize that highlighting the interconnectivity of STEM subjects with other fields can help correct misperceptions associated with those who pursue STEM careers and motivate a larger and more diverse population of students to in continue their STEM education.

1.3 Contributions

A System for Adaptive Interest-based Learning (SAIL) results in several contributions to the computer science education research community among other STEM education research communities. Specifically, this work contributes to the following issues in the CS and STEM education communities:

- Remedying gender disparities in confidence and achievement
- Elevating overall performance
- Increasing enjoyment and motivation
- Increasing attraction and retention
- Raising the universal standard of STEM education

Additionally, the design of SAIL demonstrates many novel contributions to the facilitation of interest-based learning in adaptive learning systems, including:

- A framework for IBL applicable in multiple domains
- Addressing the bottle-neck of adaptive content creation through community knowledge sharing
- A scalable design not limited by class size or instructional design

1.4 Overview of Dissertation

- Chapter 2 provides a literature review and motivation for our research.
- Chapter 3 describes ALICE - Adaptive Learning for Interdisciplinary Collaborative Environments from which SAIL evolved. We discuss the overview of the ALICE system, what we learned from our initial pilot studies, and how this knowledge inspired the creation of SAIL.
- Chapter 4 introduces SAIL, a System for Adaptive Interest-based Learning. In this chapter we describe in-detail the missions that SAIL aims to accomplish and the design decisions for an the initial SAIL framework.
- Chapter 5 outlines the need for SAIL in computer science (SAIL-CS). In this chapter we present background information about the issues with widespread computer science education and detail how SAIL could help alleviate these issues. As SAIL was created to help alleviate issues in STEM education, SAIL-CS can be seen as an isolated study of SAIL's impact in a specific STEM field, computer science, where these STEM issues are prevalent.

- Chapter 6 describes the SAIL-CS Pilot study. We discuss here the method, materials, and procedures for piloting SAIL in an introductory level computer science course at the University of Georgia.

- Chapter 7 presents and interprets both quantitative and qualitative results from the SAIL-CS Pilot study.

- Chapter 8 provides a detailed discussion of the results and relates the knowledge gained to existing literature. We end our discussion by returning to our initial research questions to provide answers based on what we have learned.

- Chapter 9 reflects on the overview of our work, summarizing our contributions and discussing possible future directions.

2 Literature Review

SAIL was designed to fill many of the gaps we identified in STEM education and existing adaptive learning tools. In this section, we review the literature motivating the creation of SAIL, including problems seen in CS education, existing adaptive learning tools and their limitations, and an overview of personalized context and the benefits of interest-based learning (IBL).

2.1 Making CS Inclusive [1]

A 2012 report from the Bureau of Labor and Statistics projected that by 2020, the computer science and mathematics job market will increase by 22%. This is significantly higher than the projected growth rates of many other occupations, including other STEM occupations [5]. This trend makes sense as computer science seems to be everywhere – yet even with a growing importance in our daily lives, the field remains underpopulated, misunderstood, and under-representative of women and minorities. Many research en-

[1] This section uses significant portions of textual materials from: Aguar et al. 2016 "Making CS Inclusive" ©2016 IEEE.

deavors seek to understand why CS is underpopulated and how to expand and diversify the field.

A diverse representation in CS is important to avoid gender-biased products, better cultivate innovation, and produce better results [6]. Additionally, more females can help address the growing demand for computer scientists. We review studies of the discouragement and encouragement of women pursuing CS and the recent efforts to remedy this divide at the University level. Much of the literature claims that CS education is needed earlier (K-12) to further promote diversity and expansion. We overview why this is the case and study the successful efforts being made.

2.1.1 Women in CS

Numerous females have played a huge role in pioneering the field that is now computer science. Some of the most notable include Ada Lovelace, Grace Hopper, and Adele Goldberg. Though many women were prevalent in the beginning of computer science, it is no secret that women have been under-represented in this field for years. The Computing Research Association's Taulbee Survey indicated that in 2016, only 17.9% of all computer science Bachelor's degrees were awarded to females [7]. Though these numbers are up (only 14% female reported in 2014 [8]), low female enrollment in CS has been an ongoing trend, with numerous efforts being made to recruit and retain women in the growing technology field. In order to attract more women into the field, one must understand what has encouraged and hindered female pursuit of CS.

The image currently associated with computer science has been shown to influence who pursues this field. The stereotype of a computer scientist is often described as someone who has a singular interest in computers with no outside hobbies or interpersonal skills [9,10]. There also exists an implied level of intelligence among computer scientists, with constant media portrayal as "geniuses" or a "nerds" [11]. Many studies in recent years have shown how these unrelatable stereotypes associated with computer science can lead women to question whether they belong in CS and whether they have the skills required to even consider the field [12]. The most recent is a 2016 study published in the Journal of Educational Psychology, showing that, with a study of 269 high school students, the stereotypes associated with CS still undermine female interest, confidence, sense of belonging, and pursuit of the field [13]. Their findings concluded that "providing [females] with an educational environment that does not fit current computer science stereotypes increases their interest in computer science courses and could provide grounds for interventions to help reduce gender disparities in computer science enrollment." [13]

To combat this misperception about who can and should pursue CS, many efforts have been made to showcase female role models and provide support to women involved in CS and STEM fields. The media has tried to bridge the disconnect of how STEM fields are advertised to young girls with shows like PBS's *SciGirls* (pbskids.org/scigirls). Organizations within the CS community such as CodeEd (codeed.org) and Geek Girl (geekgirlcamp.com) have been launched to provide support and resources for women interested in the computing field. The yearly Grace Hopper Conference (ghc.anitaborg.org) seeks to advertise the female role models existing in computer science and provide networking opportunities and encouragement for women.

Many colleges have re-designed their introductory CS courses to be more appealing to a larger audience by implementing breadth-first approaches to show students the functionality and interconnectivity of CS with other fields. This broad introduction allows students to get involved in interesting projects sooner and has been proven to help eliminate the singular-focused stereotype associated with CS, aiding in attraction and recruitment of a larger, more diverse set of students [14,15]. Harvey Mudd College is the most famous success story, drastically increasing their enrollment of female students in CS by implementing a breadth-first introductory CS course, getting students involved in interdisciplinary research projects, and sending female students to the Grace Hopper Conference [14]. In May, 2016, Harvey Mudd College saw more female CS graduates than males, with a record-breaking 54% of CS Bachelor's degrees awarded to females [16].

Since computer science now spans across almost any discipline and has a growing importance in our daily lives, its evolving nature can serve as an advantage in attracting women to CS. The literature shows that the recruitment of more women to computer science can be facilitated by adjusting the way CS is introduced – by demonstrating, early on, the interconnectivity of CS and how it can be used to impact the world [17].

Currently, these successes are isolated and the stereotypes keeping females away from CS still prevail. Studies have shown that an individual's attitude and perception towards STEM majors and careers are formed as early as middle school [18]. This argues that to recruit more computer scientists, regardless of race, gender, or background, a reformation of the CS image is needed before students reach college.

2.1.2 Early Exposure

Students entering college usually have no formal experience with computer science concepts as most primary and secondary schools do not include CS in their curriculum [19].

Most schools that do incorporate a computer science curriculum do so at the high school level, as an elective that does not count towards graduation [20].

CollegeBoard (collegeboard.org) reported that being introduced to subjects in high school and taking AP exams heavily influenced major selection in college [21]. With the low number of computer science classes available at U.S. high schools nationwide, the number of students taking the AP CS exam is well below the national numbers for comparable subjects. It was found that students who do have the opportunity to take the AP computer science exam were eight times more likely to select CS as a major [21].

With these compelling statistics, organizations such as The Computer Science Teachers Association (CSTA - csta.acm.org) and Code.org have been pushing for a change. A result of their efforts is the new "AP Computer Science Principles" (APCSP) course which launched in the Fall of 2016 [22]. This course focuses on computational thinking and how computing affects the world. Computational thinking (CT) is defined as "encompassing a set of concepts and thought processes from CS that aid in formulating problems and their solutions in different fields" [20]. The Principles AP course is not meant to replace, but to complement the existing AP Computer Science course (discussed further in Chapter 5) - which has now been renamed to AP Computer Science A (APCS-A) [23]. The APCS-P course is designed as a breadth-first approach, hoping to change stereotypes and attract more students to the field. these two courses can be taken independently, in any order; however, AP Computer Science A naturally succeeds AP Computer Science Principles for students wanting to continue their education in CS as it teaches essential syntax, debugging, and logic in a Java based environment. Unfortunately, the AP Computer Science A course has far fewer resources and many challenges preventing success [24]. The new APCS-Principles course has a high potential to successfully attract a more diverse population of K-12 students to the CS field, though measures will need to be taken to ensure that students are retained and remain interested beyond this initial introduction. More effort is needed to ensure the curriculum in subsequent courses, such as APCS-A and introductory level college courses include both a competent and compelling curriculum.

2.2 Interest-based Learning [2]

Many studies have demonstrated that self-reference and context personalization have influenced student memory and learning [25]. The Self-Reference Effect [26] has been extensively studied and has shown that the customization of information to relate to the self or someone closely associated with the self can lead to learning improvement including better recall, transfer, and retention of information [27–29]. Recent studies evolving from the SRE have demonstrated that the the customization of content to include students' familiarities and interests (personalized context) have also shown significant benefits in student learning [4, 30] and that stimulating interest can lead students to continue their education in the subject [31, 32].

In studying how interest affects learning, we must first understand that there are two distinct types of interests: situational and personal (individual). Situational interest is described as a fleeting interest that can be easily invoked by the one's environment,

[2]This section uses significant portions of textual materials from:
Aguar et al. 2017 "Towards Interest-based Adaptive Learning and Community Knowledge Sharing" ©2017 International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'17)

while personal (or individual) interest is slowly developed and longer lasting [33]. Both have been shown to positively affect learning [34], but most educational interventions target situational interest as it is much easier to stimulate for all students [35]. "The Science of Interest" published in August 2017 overviews the role of interest in learning, and concludes that "The study of situational interest would profit from the explicit introduction of instructional events that have the potential to arouse it. Without aroused situational interest, one cannot expect learning to be affected." [35].

Most of the studies to this date have been manually conducted, with initial studies using adaptive technology as a medium for personalized context demonstrating positive learning outcomes [25]. Most notable is a 2013 study that personalized word problems in one unit of an Algebra course with an Intelligent Tutoring System (ITS) as the educational medium. Students who received word problems with personalized context had better and faster performance on the affected unit of instruction and demonstrated positive outcomes in abilities to transfer knowledge and retain information [4].

While results have been overwhelmingly positive for the incorporation of student interest in education, widespread incorporation is difficult. Most studies have been conducted in the field of mathematics where key words can be easily substituted into generic word-problems [4, 25, 30]. This approach limits the domains in which personalized context can be implemented and also the depth of which problems can relate to student interests. Many STEM fields such as computer science, engineering, and mathematics are highly applicable with other disciplines and the world around us. Utilizing the interconnectivity of these subjects with the world around us could be immensely beneficial in correcting the low enrollment, retention, and diversity issues that these STEM fields often suffer in addition to helping to improve learning. Imagine a world where a student athlete could learn how to program based on sports related examples while another student in the same class could learn programming through science examples. We believe that this interest-based learning could help with diversity issues in a similar way that the breadth-first introductory programming classes could. In both scenarios, the main goal is to introduce students to computer science by showcasing how applicable it can be in other areas, thus eliminating the stereotype often associated with CS and encouraging more students to pursue its studies. In addition to this, interest-based learning has the potential to not only help attract a larger and diverse student population, but also to improve learning by arousing situational interest.

However, no system exists to easily facilitate interest-based learning through technology in a wide range of subjects. This coupled with the enormous effort required to create customized content have immensely limited the widespread incorporation of IBL in STEM education.

2.2.1 Community Knowledge Sharing

We propose a framework that supports community knowledge sharing (crowd-sourcing) of adaptive materials to address these limitations. Creating and encouraging the use of online knowledge sharing communities for educational resources is an effort being explored by many top universities. In the computer science community, efforts such as Stanford's Nifty Assignments project seek to collect 'interesting' CS assignments for reuse to help improve CS education (nifty.stanford.edu). Similarly, UC Berkeley's Ensemble project seeks to establish a digital library for computing education, with current research in participation encouragement through ranks/badges (computingportal.org). These movements show that educators are making efforts to provide students with "interesting" material to improve education - though we note that any "interesting" problem can only interest a subset of students, and recognize the need for adaptive technologies to help facilitate customized instruction based on interest. As we move towards the sharing of educational resources in online communities, more work is needed to encourage participation, organization, and optimize utilization of these materials.

2.3 Adaptive Learning

2.3.1 What is Adaptive Learning?

Adaptive learning is the notion of using computers as interactive teaching devices to adapt to the user's individual needs. It combines the fields of Computer Science, Education, Psychology, etc. The computer adapts the way it presents material or decides what the next question will be based on its interactions with the students - via observing the student and/or analyzing their responses. Adaptive Learning is a broad term that encompasses many Adaptive Learning System (ALS) varieties such as Adaptive Educational Hypermedia (AEH), Intelligent Tutoring Systems (ITS), Adaptive eLearning, and others. It is primarily used in educational settings such as classrooms and business training [**?**, 2]. Our work is best classified as adaptive educational hypermedia (AEH) - as it adapts the multimedia content presented based on a student's individual preferences. Like many AEH systems, our work lends itself nicely to online, adaptive eLearning approaches to instruction.

Adaptive Educational Hypermedia Hypermedia is a term evolved from the word 'hypertext' referencing a display of content that the user can interact with such as links, videos, or pictures. Hypermedia differs from multimedia in that the user is actively involved with hypermedia, such as clicking links to navigate through webpages, whereas multimedia interaction is passive, such as watching a video or listening to music [36]. Adaptive hypermedia is an extension of hypermedia where the content displayed is not the same for all users, but "adapts" to a personalized display based on information it collects about the user. Adaptive hypermedia has many areas of implementation, but lends itself well to web-based technologies. For example, websites such as Amazon and many others can use one's purchasing history to adapt recommendations of other items the user may be interested in [37]. Our focus is on adaptive hypermedia where it applies to education, i.e. adaptive educational hypermedia.

2.3.2 Adaptive Learning Efforts

Research on Adaptive Educational Hypermedia (AEH) has gained significant interest in the last two decades. Adaptive hypermedia systems build a model of the goals, preferences, and knowledge of individual users to adapt to their specific needs [38, 39]. Such systems modify learning experiences on the basis of the system's ability to identify the learner's needs (i.e. adaptivity) and the possibility for learners to make decisions on their own (i.e. adaptability). The majority of systems to date have addressed adaptivity based on learning styles or cognitive models [40]. Frameworks such as the Felder-Silverman learning style dimensions [41], Keefe's classification of learning styles [42], and cognitive styles models [43,44] have guided the design and implementation efforts of AEH systems.

Evidence suggests that AEH systems are effective at tailoring instruction for heterogeneous groups of students both in higher education [45] and in K-12 settings [4, 46].

Many successful adaptive tools are currently being created. One notable adaptive eLearning system, Smart Sparrow [47], has generated much excitement including a recent Forbes report claiming that it is "leading the way" [48] in changing education. It is currently being used in a handful of institutions world-wide, including Arizona State University. Smart Sparrow gained widespread attention when a 2011 study in a mechanics course reported that adaptive eLearning reduced failing grades by 24% [49].

The most research on adaptive learning systems in recent years have been in the area of Intelligent Tutoring Systems (ITS). An intelligent tutoring system is one of the most powerful forms of adaptive systems as it uses advanced AI techniques to not only determine if a student answered incorrectly, but also why that answer was incorrect. These systems act as a personalized tutor to students, customizing which problems to show based on interactions with the student, and enabling customization based on pace and prior knowledge. We overview the some of the most notable systems below:

McGraw-Hill Education has released its own adaptive learning platform and incorporated adaptive learning technology into their e-books, Smartbooks [50].

Pearson & Knewton are two major education-based companies, who joined forces to launch adaptive learning tools, Pearson's MyLab and Mastering, focusing on many K-12 common core topics [50]. For example, students in elementary mathematics can complete sample problems after learning a subject and the computer prints homework specifically focused on their weaknesses or more advanced homework if they have successfully mastered the material.

Others : Desire2Learn, Scootpad, and many other similar tools have been created to implement adaptive learning techniques, mostly focusing on the common core K-12 curriculum. Prices vary based on features and some even come equipped with accompanying smartphone/iPad apps [50].

MATHia : Carnegie Learning developed MATHia(R) and Cognitive Tutor intelligent tutoring softwares to help improve middle-school and high-school mathematics skills to meet the common core standards (https://www.carnegielearning.com/learning-solutions/software/).

Like those described above, the most successful adaptive learning tools are industry-grade, featuring user friendly designs, but are expensive and target only one subject - implementations targeting core K-12 subjects such as mathematics have been the trend. With the high expenses needed to create an adaptive learning system, most "good" implementations can only be found in "high density domains (e.g., high school algebra) where the expense to build an ITS could be offset by the large number of learners who might pay to use the tutor" [51]. This limits the availability of powerful adaptive technology to a narrow subset of subjects and those who can afford access.

Many introductory CS prototypes or initial ITSs have been proposed to tackle various aspects of introductory CS. These domains include teaching language specific syntax and semantics [52] [53] [54], learning to debug [55], and practicing algorithm design [56]. Most of these prototypes or initial systems have ambitious goals to implement the actual intelligent system in the future. With the expense needed to create an industry-grade ITS, the systems created in domains such as CS are usually fairly limited compared to the full potential of an ITS, with some still requiring testing and evaluation, and others listing numerous future improvements.

2.3.3 Current Trends and Future

The most recent research for ITS have been moving towards integrating affective adaptation - adapting based on a student's emotional state or their learning style. A recent example is a 2017 system that uses a webcam to take pictures of students' faces to detect boredom, frustration, excitement and engagement. Based on the emotion detected, it recommends problems or displays encouragement customized by the teacher [57]. The trend seems that while research endeavors are tackling different angles (new emotion detection strategies, adapting based on different learning style models, etc.), all of these implementations are singularly focused and not applicable in more than one domain. The most popular adaptive learning systems described above are all commercial products, as it is incredibly expensive to create both the intelligent system and the adapted content. In all of these designs and in all of the research we've described, all efforts in creating adaptive systems and content are then limited to that narrow domain for which it is created. While the research bank for adaptive technologies is quickly growing, we see very little growth in widespread incorporation of these adaptive technologies.

The use of authoring tools allowing customization of an ALS to fit multiple related domains has been researched to address these limitations, but there is no "one-size-fits-all" tool [58]. GIFT, Generalized Intelligent Framework for Tutoring, is among the most advanced, and is in active development at the US Army Research Lab. GIFT is attempting the creation of a generic, open-source intelligent tutoring system that makes authoring easy [59]. In the future research and creation of successful authoring tools, one must considered how to lower the skills required to author the system and the usability and acceptance of the technology for widespread use [51, 60].

2.3.4 Challenges

Significant research is currently being done in the field of Adaptive Learning Systems, but progress is slow. Creating a successful ALS is difficult, time-consuming, expensive, and usually geared towards one particular domain.

While the technology behind adaptive learning is advancing at great rates, these systems are very narrowly focused, targeting only one subject, with ease of use often lacking - limiting success. The most successful tools are industry-grade, feature user friendly designs, but are expensive and target only one subject.

The use of authoring tools allowing customization of an ALS to fit multiple related domains is being researched with no current widespread success.

2.4 Our Vision

We feel many levels of STEM education could benefit from the widespread incorporation of interest-based learning in their curriculum, but have found no tools available to the community that easily facilitate this adaptation based on interest in a user-friendly way, applicable in multiple domains. While most ALS adapt based on a student's previous knowledge, pace, or learning style [2], this research takes an alternative, less-explored approach by adapting content, practice problems, and examples based on a student's interests. Based on the studies showing that adjusting the way introductory CS classes are taught can recruit a larger and more diverse subset of students, we feel an adaptive learning system that facilitates easy adaptive learning based on interest can accomplish similar goals. In addition to utilizing adaptive, interest-based learning to highlight the interconnectivity of STEM fields with the world around us, evidence shows that interest-based learning can aid in improving student learning and achievements. We propose SAIL - an adaptive learning system that adapts based on interest, is not limited by domain, designed with the user (both student and instructor) in mind, and available for widespread use. However, as adaptive systems by nature require more effort to create content than non-adaptive courses, a task that would quickly bottle-neck if created in isolation [4, 51], we design a system that lowers the skill required to curate adaptive content, laying the groundwork for a system that can facilitate a future of community knowledge sharing of educational resources.

3 | ALICE

This chapter introduces ALICE (Adaptive Learning for Interdisciplinary Collaborative Environments), an adaptive learning system designed for interdisciplinary instruction that inspired the creation of SAIL.

3.1 ALICE Motivation

The dominating paradigm in interdisciplinary STEM education is inherently inefficient particularly for students from various disciplinary backgrounds attempting graduate studies. It consists of essentially teaching the same knowledge base to each student within the classroom; however: (i) students in these settings usually come from different disciplines, thus having different (often non-overlapping) backgrounds, and (ii) curricula in interdisciplinary fields are comprised by subject matter drawn from different (often traditionally disconnected) areas. Case in point, systems biology; in this area, students need to master a biological problem, know the theory of dynamical systems (continuous and discrete), probability, statistics, and be able to program, just to mention a few subjects. Students who take this interdisciplinary class at the senior undergraduate and junior graduate levels generally major in genetics, biochemistry, horticulture, mathematics, computer science, statistics, physics, engineering, etc. As a result of these multi-disciplinary skill requirements and the inherent diversity of student backgrounds in an interdisciplinary class, some students in the classroom have expertise in some areas and deficits in others, and these strengths and weaknesses are unique for each student. While it is possible to require all students in these settings to master simultaneously a collection of disciplinary content areas, the delivery of instruction in which all learn the same at equal pace poses uneven and unreasonable demands on students. Ideally, each student should strengthen her/his specifics weaknesses, and broaden the scope of their strengths within the same time frame allotted for the class.

We developed an open-source Web-based cyber-learning tool that allows any team of instructors spanning several scientific disciplines to curate a constellation of interdisciplinary learning resources for the purpose of creating individualized or small group learning progressions for developing prerequisite competencies and responsive education to all students. The personalization of the learning plan or syllabus for each student depends on previous knowledge and individual learning goals. This customization is achieved through an information system called ALICE (Adaptive Learning for Interdisciplinary Collaborative Environments), which connects a series of atomic units of knowledge (termed lexias) though a dynamic path and presents it to the student for the purpose of acquiring a set of competencies. The metaphor of the tree is replaced in ALICE by a dense rhizome-like network that does not privilege a particular path, but instead offers a milieu for traversal. In practice, it is the student during the learning process who makes an abstract knowledge network come to a unique realization. ALICE was initially designed for graduate and senior undergraduate learners in the subject matter of Systems Biology. Based on task analyses and dynamic assessments of individual learners, each learning progression was designed to take the learner from individual baselines to desired levels of competence.

ALICE personalizes education by: (1) creating a knowledge map of course material that is unique for each student (i.e, a personalized syllabus); (2) suggesting individualized paths across the knowledge map based on student competencies/accomplishments; (3) providing accessible Web-based interfaces for students and instructors for storing and presenting class materials, for assessment, and for recording student paths. We have created ALICE (Adaptive Learning for Interdisciplinary Collaborative Environments), an open-source web-based cyber-learning tool that allows personalization of the learning plan or syllabus for each student depending on previous knowledge and individual learning goals. ALICE is an ideal system for interdisciplinary settings where students come from a variety of backgrounds, as it eliminates redundant information and allows each student to strengthen her/his specific weaknesses, and broaden the scope of their strengths within the same time frame allotted for the class. Within ALICE, instructors spanning several scientific disciplines can curate a constellation of interdisciplinary learning resources for the purpose of creating individualized or small group learning progressions for developing prerequisite competencies and responsive education to all students. ALICE significantly impacts the role of the instructor and design of the course.

3.2 ALICE Design

The architecture of ALICE is based on the Literatronica system [61,62]. Figure 3.1 shows the workflow that permits adaptive behavior. The flow of information in ALICE has three interconnected domains: the student, the instructor, and the information system. ALICE plays a fundamental and autonomous role in guiding the students through the material. This optimization process is achieved through the real-time solution of a multi-terminal network flow and maximal network flow on a dynamic graph. In ALICE each competency has a determined finish point. Each time a learner interacts with the system, ALICE reconfigures links to have different destinations, leading every time to a personalized and potentially unique learning experience. ALICE offers an adaptive behavior that ranges from maximal flow (i.e. completion of the track with the maximum number of lexias) to minimal path (i.e. completion of the track with minimum number of lexias).

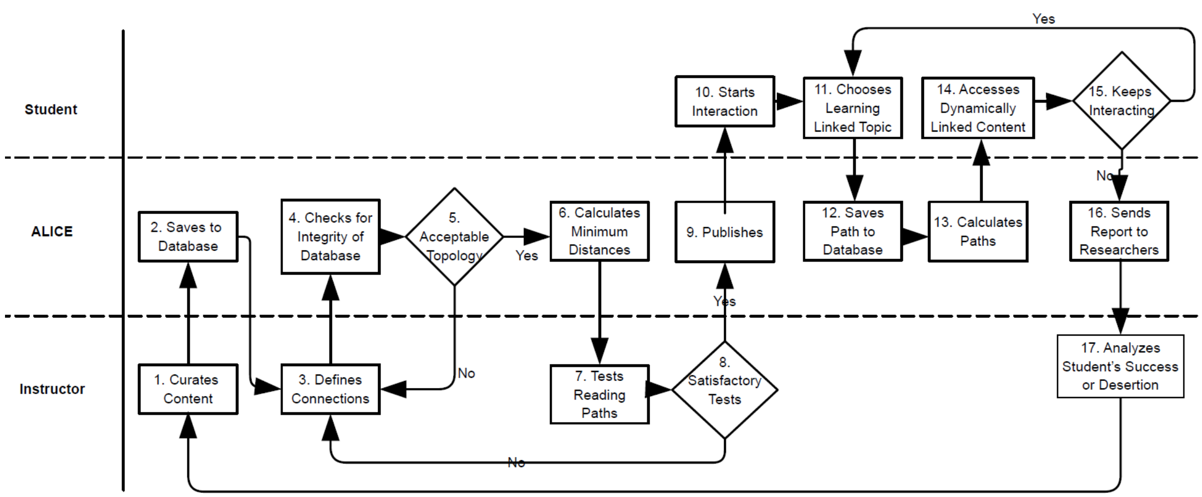

Figure 3.1: ALICE - The Flow of Information

ALICE presents a knowledge map to the students - as illustrated in Figure 3.2. Colors identify the main area of each lesson - statistics, computer science, mathematics. Square shapes represent lessons related to the central theme of the course, while circles represent pre-requisites and triangles represent the five capstone experiences (final projects). Once students select a capstone experience, their first lesson is identified with a star. Figure 3.3 shows an example of what a student's Individual Development Path (IDP) might look like to reach their selected capstone experience. Items in gray are not included in the student's IDP.

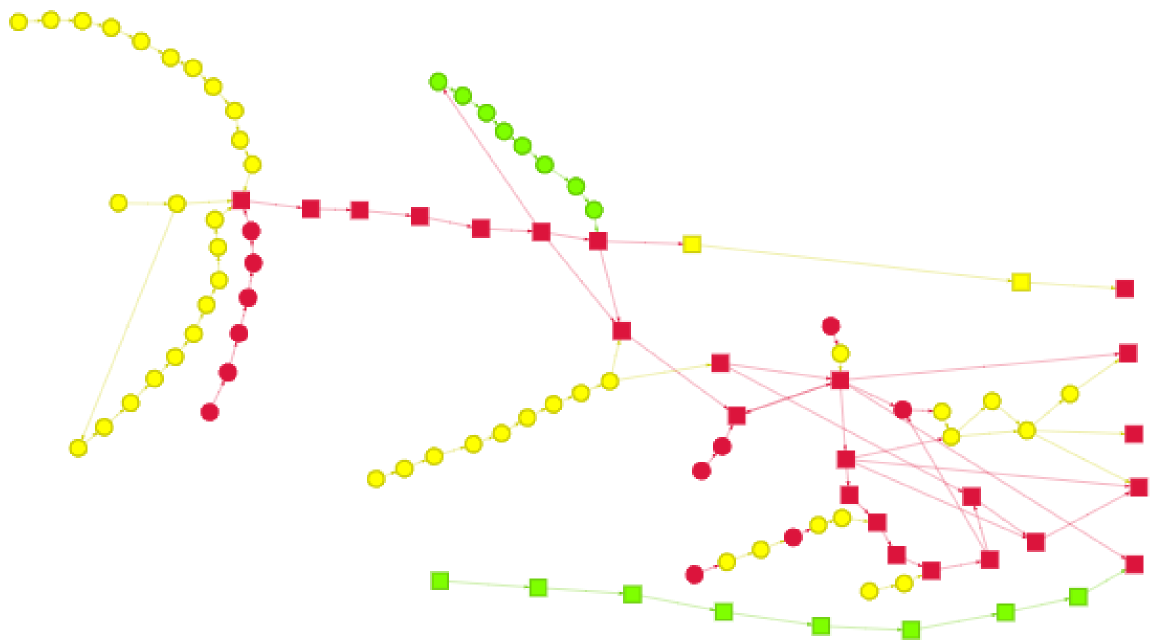

Figure 3.2: ALICE - System Biology Pilot Knowledge Map

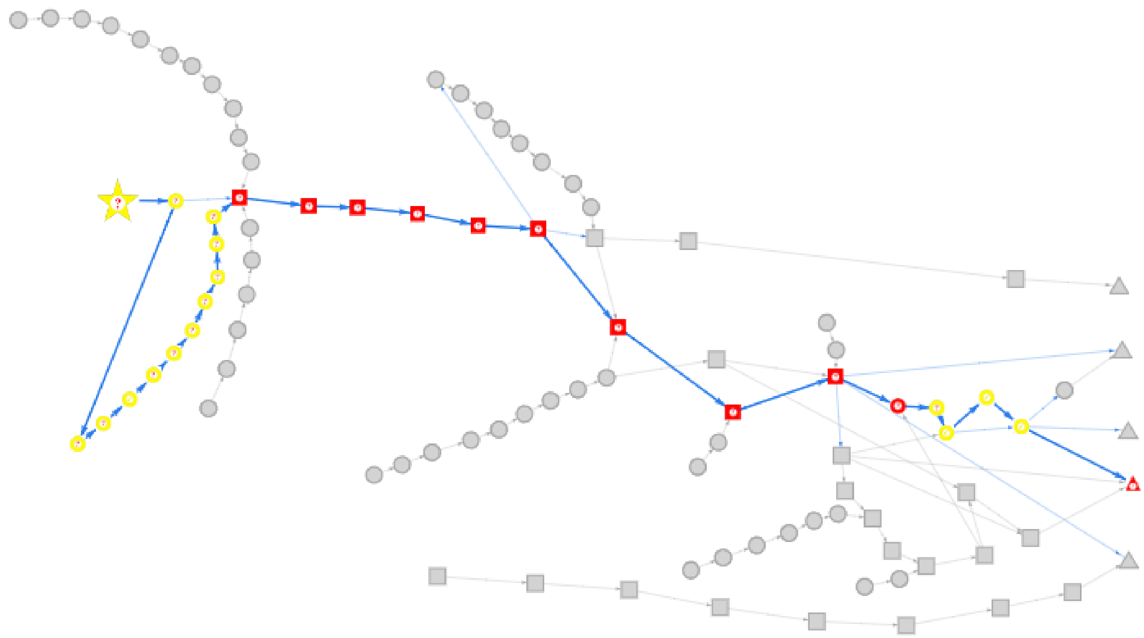

Figure 3.3: ALICE - A Unique Student Path

3.3 Pilot Study and Future

To begin exploring viability and assess ALICE's impact, a pilot study in Systems Biology is being conducted at the University of Georgia. Systems Biology is an ideal interdisciplinary course combining skills from biology, mathematics/statistics, and computer science. Students taking this course come from a variety of disciplines and represent a diverse population of educational levels ranging from advanced undergraduates to post-doctoral fellows. The interdisciplinary nature of this course makes it an excellent candidate to test ALICE. Throughout this study, sections of Systems Biology are being offered at both the senior undergraduate and junior graduate levels, and are being contrasted in various ways. Some section are serving as the "control" group and are being taught in a traditional lecture-style classroom setting exactly as it has been taught in the past using best current practices. Other sections are acting as the "experimental" groups and are using ALICE for mediating learning and teaching. This study will conclude in mid-2018 and results will be reported soon after.

In implementing ALICE, many modifications must be made to traditional instructional design. As we set up the initial ALICE pilot, we reported the considerations of the architecture and guidelines to incorporate an adaptive learning sytem like ALICE in a course [63]. The traditional lecture style was no longer possible as students acquire knowledge personalized to their individual background and goals. With ALICE, the classroom evolves into a discussion based setting to encourage collaborative thinking and problem solving between disciplines. The instructor's role changes from lecturer to facilitator of learning. Normal course elements such as centralized assessment techniques, due dates, course progression, and grading schemes must all adapt with personalized instruction. Aguar et al., 2017 [63] discusses in further detail the influence of ALICE on these instructional design aspects.

Though certain adaptations to instructional design should be anticipated when transitioning to a new teaching method, there is an enormous amount of effort currently required by the instructor to develop adaptive content. Even with the most dedicated educators, keeping up with the demand needed to implement ALICE in the Systems Biology pilot study was a struggle. While many adaptive learning systems have had promising outcomes, the task of creating adaptive content is an obstacle preventing widespread incorporation and success. These observations identified a clear need for the sharing and reuse of adaptive content.

ALICE aims to solve a unique problem in STEM education, by addressing the needs of interdisciplinary STEM courses such as Systems Biology. While the ALICE pilot is still ongoing and results are pending, it has the potential to transform the way interdisciplinary courses are taught by minimizing the huge overhead needed for students from various backgrounds to work together to solve complex, real-world problems. However; this interdisciplinary framework is not one that can be applied to all or even most STEM fields. Many STEM subjects require a more traditional instruction approach, where all students in the class must progress through the same linear ordering of competencies. Even though these subjects may not be "interdisciplinary" in the ways that Systems Biology is, most STEM subjects do have high applicability in other disciplines. That is, many STEM subjects such as Mathematics, Computer Science, or Engineering may be "monodisciplinary" - but many times, these "monodisciplinary" skills can have a wide range of applications across other fields. This realization, along with the bottleneck problem in adaptive content creation inspired the creation of SAIL - A System for Adaptive Interest-based Learning in STEM education.

4 | SAIL

This chapter uses significant portions of textual materials, graphs, tables, and/or figures from Aguar et al. 2017 "Towards Interest-based Adaptive Learning and Community Knowledge Sharing" ©2017 International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'17).

4.1 Motivation

Adaptive learning strategies have been shown to improve student performance, with adaptation usually based on a student's previous knowledge, pace, or learning style [2]. This research takes an alternative, less-explored approach by enabling the adaptation of content, practice problems, and examples based on a student's interests. Incorporating personal interest into learning has been shown to increase intrinsic motivation and provide positive learning outcomes [3]. Most studies of personalized interest in education have been implemented manually, but initial studies incorporating personal interest into adaptive technologies have indicated tremendous potential [4].

We feel many levels of STEM education could benefit from the widespread incorporation of interest-based learning in their curriculum, but have found no tools available to the community that easily facilitate adaptation based on interest in a user-friendly way that is applicable in multiple domains. Evidence shows that re-designing introductory CS courses to highlight the high applicability with other fields can help recruit a larger and more diverse subset of students. We feel an adaptive learning system that stimulates interest while learning introductory programming concepts can accomplish similar goals. In addition to utilizing adaptive, interest-based learning to show the interconnectivity of STEM fields with the world around us, evidence shows that interest-based learning can aid in improving student learning and achievements.

We propose a System for Adaptive Interest-based Learning (SAIL)- a web-based adaptive learning framework that empowers adaptation based on interest, is not limited by domain, designed with the user (both student and instructor) in mind, and available for widespread use. The vision is to provide a framework where educators from multiple domains can contribute to and access adaptive content available to the community to provide students with an improved and individualized learning experience. Such a system can help fill the gaps in educational STEM resources, help motivate students to pursue STEM fields, and help attract and retain a larger and more diverse populations of students.

4.2 SAIL Design

The development of a new adaptive learning tool, SAIL (System for Adaptive Interest-based Learning), could be the solution to deliver customized curriculum based on students' interests to help attract and retain a more diverse population of students while promoting a standard of achievement across STEM subjects.

SAIL provides a **System** that enables an **Adaptive** learning experience, using **Interest-based** examples to tailor content and exercises for each student. Interest-based adaptive learning has the potential to provide a better overall **Learning** experience, customized to each student, to increase intrinsic motivation and enjoyment of STEM fields while ensuring a standard of achievement is met.

4.2.1 Evolution

SAIL is the evolution of ALICE (Adaptive Learning for Interdisciplinary Collaborative Environments), an unprecedented adaptive learning system for interdisciplinary instruction. SAIL harnesses many of the exciting potentials in ALICE, but significantly expands the goals, adaptive framework, and usability in the following ways:

1. **Goals:** ALICE aims to solve a unique but important problem in interdisciplinary STEM education, providing individualized instruction students to traverse different learning paths based on their backgrounds and goals. SAIL was inspired by this interdisciplinary approach to instruction, but expands the vision for a wider variety of STEM issues. Most STEM courses follow a linear ordering of course concepts (ex: addition precedes multiplication), but these topics are often highly applicable in other areas (ex: adding up points in a game of putt-putt or adding up items on your grocery list). SAIL expands the adaptive framework to support STEM subjects with linear course progression and high applicability in other fields to

help address the well-known issues of low enrollment, a lack of diversity, a lack of educational resources, and often inconsistent educational standards.

2. **Adaptive Framework:** The adaptive framework with ALICE supported goal-based adaptivity so students with different backgrounds and different goals in the course could traverse unique paths with varying starting and ending goals at the same time. SAIL preserves this interdisciplinary design, but expands the adaptive framework to support linear course progression of course topics while adapting practice problems and examples based on interest. With the high applicability of many STEM fields to the surrounding world, highlighting this interconnectivity through customized problems based on a student's interest could increase intrinsic motivation of students studying STEM concepts as they learn fundamental skills through avenues that are already exciting to them. This increase in motivation has been shown to lead to better performance outcomes and can hopefully help with attracting more students to continue their STEM studies.

3. **Enhanced Authoring and Usability:** Adaptive learning systems (ALS) are expensive to create - requiring enormous efforts for content creation, usability, and the adaptive engine. All efforts going into creating a successful ALS are then limited to a specific domain, as successful, easy, authoring of these powerful systems has not been feasible. In the design of SAIL, we emphasize usability for both student and instructor, to make a system that can be accepted for use by the community and lowers the skills required to author the adaptive design. In doing so, we lay a groundwork that supports future of community knowledge sharing of educational content to help alleviate the expense required to create adaptive materials.

Initial studies of the incorporation of students' interests into adaptive learning systems have demonstrated many positive learning outcomes, and we believe it has enormous potential to aid in issues throughout STEM education. Currently, no widespread solution exists to easily facilitate IBL in a user-friendly design spanning multiple disciplines. A System for Adaptive Interest-based Learning in STEM education is a novel solution to many of the issues in widespread STEM education. To begin testing the viability and impact of SAIL in STEM classrooms, an initial system was built and tested for introductory level computer science at the University of Georgia. The following sections describe how SAIL was designed to facilitate the missions listed above.

Interest-based Adaptation

In the evolution from ALICE to SAIL, there are some significant differences in both ideology and adaptivity. ALICE's adaptivity was implemented through a network connecting multiple disciplines, while SAIL evolves to include *intradisciplinary* adaptivity - aiming to highlight the interdisciplinary nature of many STEM fields (Computer Science, Mathematics, Statistics, Engineering) through interest-based learning to show the impact and interconnectivity with other surrounding fields.

With ALICE, students learn different content at different paces with multiple starting points and end goals. While this interdisciplinary design is still supported in SAIL, the adaptive framework is expanded to support linear traversal through lexias/lessons in a pre-determined order while adapting practice problems and examples to a student's individual interests. The design of this adaptive framework can be seen in Figure 4.1. This expansion allows students to move through a linear ordering of lexias: (A, B, C,

...), while practice problems and examples branch off based on the student's interests (D1, D2, D3, .. Dn). Students may follow the adaptive path (solid arrows) or adapt their own path (indicated by dotted arrows) through the examples.

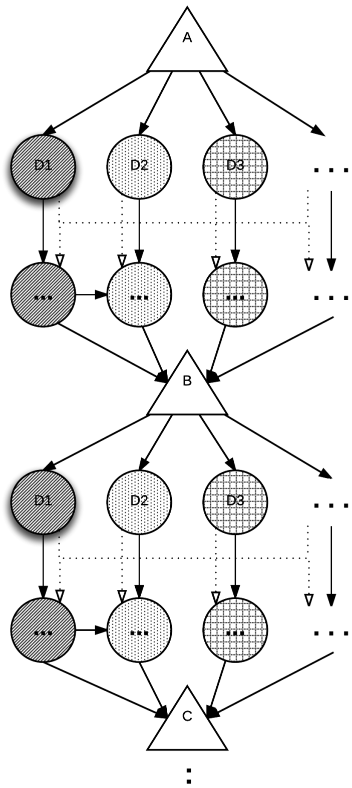

Figure 4.1: SAIL's Interest-based Adaptive Framework

For example, students in an introductory programming course could all have the same lesson on loops but will then follow a path of examples and practice problems based on their indicated interests. A student interested in Sports may take one path where they practice loops with Sports-related examples and problems, while other students may encounter problems in other domains more aligned with their interests. Adapting content based on interest can increase a student's intrinsic motivation as stimulating interest has been shown to significantly influence learning in regards to an individual's attention, goals, and level of cognition [3]. However, developing adaptive content is time consuming and too large a task for a single instructor. SAIL aims to alleviate this hurdle by laying the foundation for a framework to support future community knowledge sharing of adaptive content.

Community Knowledge Sharing

SAIL could transform the approach to adaptive learning by providing a medium for future community knowledge sharing of educational content. SAIL is being built as a scalable system - to grow over time as the community provides more interest-based examples and course content. In making such a system scalable, the usability and ease for instructors to add content and facilitate adaptive branching in their courses was important in SAIL's design. The interactive knowledge map - used previously for students to progress through course topics in ALICE, was enhanced and expanded to include an easy-to-use administrative interface to help instructors set up their courses.

The Interactive Knowledge Map

While the focus of SAIL at this point is not on Human-Computer Interaction (HCI), the usability of such a system is an important consideration for widespread success. With ALICE, students were presented with their unique path through an interactive knowledge map. In early observations of the pilot study, the interactive knowledge map seemed to resonate with students and instructors alike as an exciting way to see one's personalized syllabus throughout a course. With SAIL, we kept the interactive knowledge map for students, enhanced its usability, and expanded this concept to help facilitate easy authoring for instructors to customize their adaptive courses.

Instructor Use

An interactive drag and drop interface is included in SAIL, allowing instructors to more easily curate the adaptive interactive knowledge map (IKM) for their course. This interface allows instructors to easily add content (via uploading or reuse of community resources) to their course and seamlessly create adaptive branches in course instruction.

The interactive knowledge map (IKM) is a graphical syllabus of course content - lessons, exercises, etc. It can be thought of as a graphical view of the lesson plan during the course - where each "node" in the map represents either a lesson or homework assignment. Figure 4.2 shows a sample lesson plan for an elementary math class.

Sample Lesson Plan - Addition
1. Basic Addition Lesson (2+2=4) 2. Basic Addition Homework 3. Multiple Digit Addition (12+12 = 24) 4. Multiple Digit Addition Homework

Figure 4.2: Sample Lesson Plan - Addition

In this sample lesson plan, items 1 and 3 are lessons, while items 2 and 4 are homework assignments. Transforming this typical lesson-plan into SAIL's IKM is as simple as creating four different "nodes" - one node per item in this list - using the "add node" button, and creating an ordering of lessons by dragging and dropping directed edges between nodes. The "Basic Addition" node would have an outgoing edge to the "Basic

Addition Homework" node, implying the order in which these nodes or "modules" will be visited during course progression. Figure 4.3 shows how easy it is to create a directed edge between two nodes using the "add edge" tool.

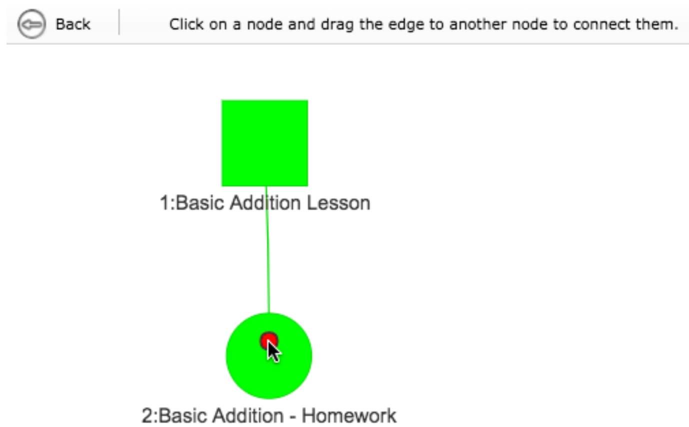

Figure 4.3: SAIL - Creating Directed Edges between Lessons

There are three types of nodes designated in SAIL:

- **Root Node** - A node that all students must progress through, regardless of interest - indicated by a square shape. An example would be the "Basic Addition" lesson - all students must progress through this lesson in this sample course.
- **Example Node** - These nodes are the practice problems, examples, or homeworks that will be given to students. Indicated by a circular shape, these are the nodes that will hold the interest-based exercises. To facilitate the adaptive branching based on interest, each example node must be assigned an interest category from a drop-down menu in the form (seen in Figure 4.4.

- **Pre-Req Node** - A node that holds pre-requisite content. Indicated by a triangle (not pictured), these nodes help preserve the interdisciplinary functionality of ALICE, where instructors can incorporate pre-requisite knowledge in a student's path, while not confusing these pre-requisite nodes with the root nodes for that course. These nodes are carried over from the ALICE design for interdisciplinary instruction, and while tying pre-requisite concepts to root lessons can be implemented in SAIL's design, we do not include discussion of these in our demonstrations.

Once a node is created via the "Add Node" button, customizing the node is accomplished by simply clicking on the node in the interactive knowledge map. The right side of the interface populates a form (seen in Figure 4.4) with that node's information, allowing the instructor to specify the node type, the node category, and add content all on the same page. Currently supported instructional materials include: embedded videos (via youtube links), PDF documents, and PowerPoint filetypes.

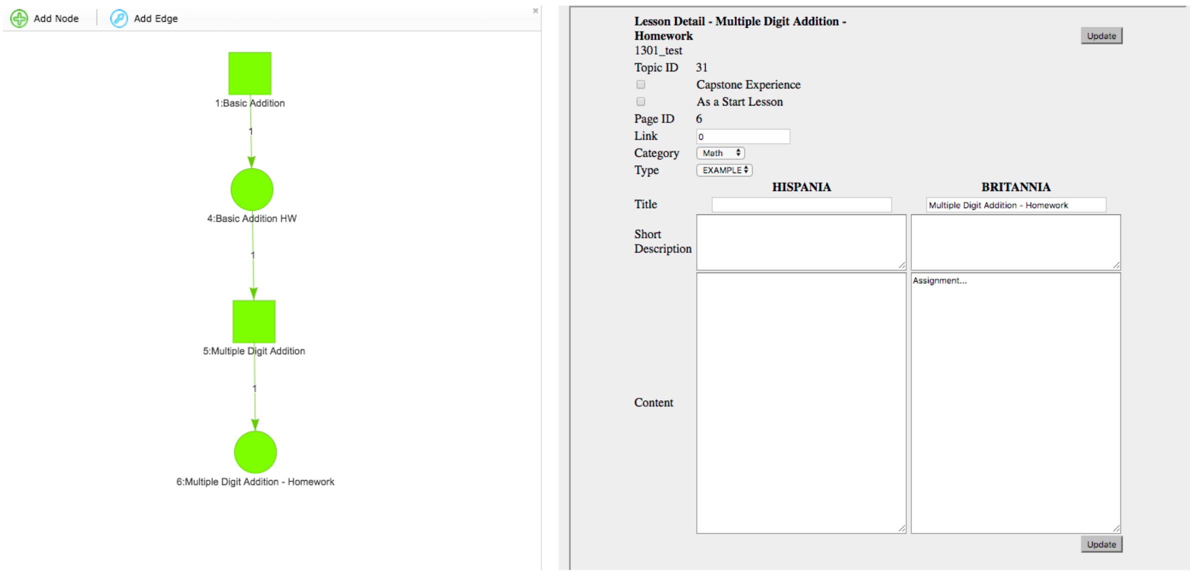

Figure 4.4: SAIL - Addition Lesson Plan

Translating the Addition lesson plan into SAIL took only a few minutes using the tools described above. SAIL's representation of the Addition lesson plan can be seen in Figure 4.4, where "lesson" modules were created as root nodes and "assignment" modules were created as example nodes. In this example, all nodes are green in color, indicating that all modules belong to the same "category" - Math. As this knowledge map is a direct translation of the sample lesson plan seen above, it implies that all students will go through the same lessons and homework assignments throughout course progression. The true value of SAIL is that these tools to add nodes and edges can help instructors easily facilitate adaptive branches in their course. For example, instead of all students receiving the same homework assignment on "Basic Addition", the instructor could easily load two or more versions of this assignment into adaptive branches in the course - see Figure 4.5. With adaptive exercises, one student could practice basic addition through sports-themed problems such as adding up points in a soccer game, while another student who may be less interested in sports could receive practice exercises more fitted to their individual interests. The adaptive lesson plan uses colors to distinguish the different interest categories. In this example, there is a "Sports" category, shown in red, an "Animals" category, shown in blue, and a "Music" category, shown in yellow.

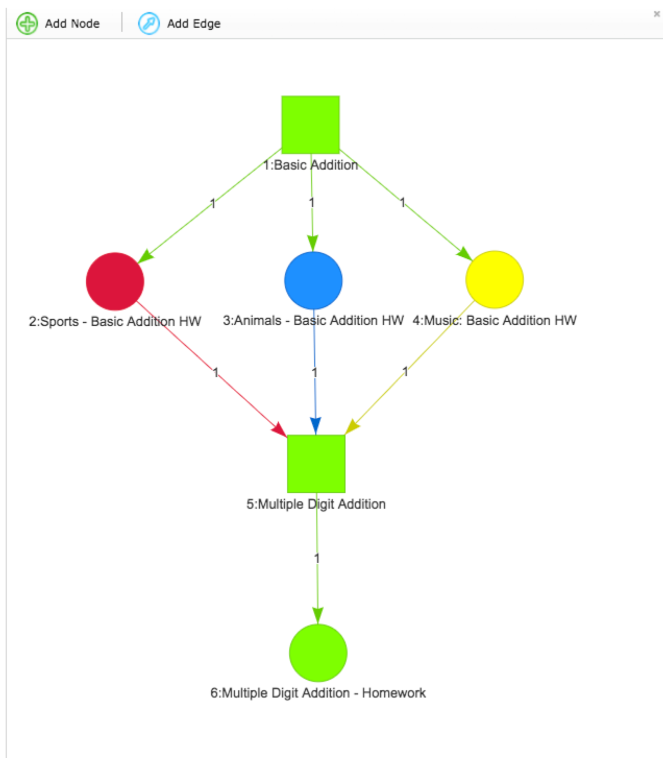

Figure 4.5: SAIL - Adaptive Addition Lesson Plan

These categories for interest-based problems can also be easily created by the instructor. In a separate page, accessible only by administrative users, teachers can easily add interest based categories to their course - see Figure 4.6

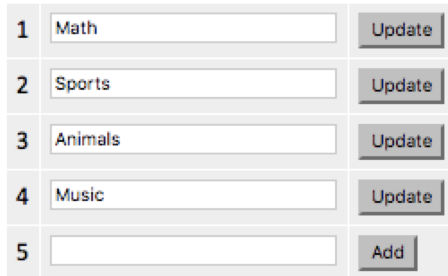

Figure 4.6: SAIL - Adding Interest-based Categories to a Course

Student Use

Once the knowledge map for the course is created by the instructor, it becomes an interactive tool for students to customize their independent learning path throughout the course. The interactive knowledge map presents students with their unique adaptive path highlighted throughout the course and helps students track their progress through course modules. Figure 4.7 shows a sample screenshot of the SAIL's interactive knowledge map from the student's view where the student's unique path through the adapted exercises is highlighted. Students must simply double-click a node to be taken to that lesson or activity. Colors and shapes for category and node type are the same as in the administrative interface. In this example, the student has selected "Sports" as their interest, so they will receive sports-themed exercises to practice basic addition. Students select their interest using the dropdown box in the top left. Switching interests can be done at any point without losing progress and will update the student's path in real-time.

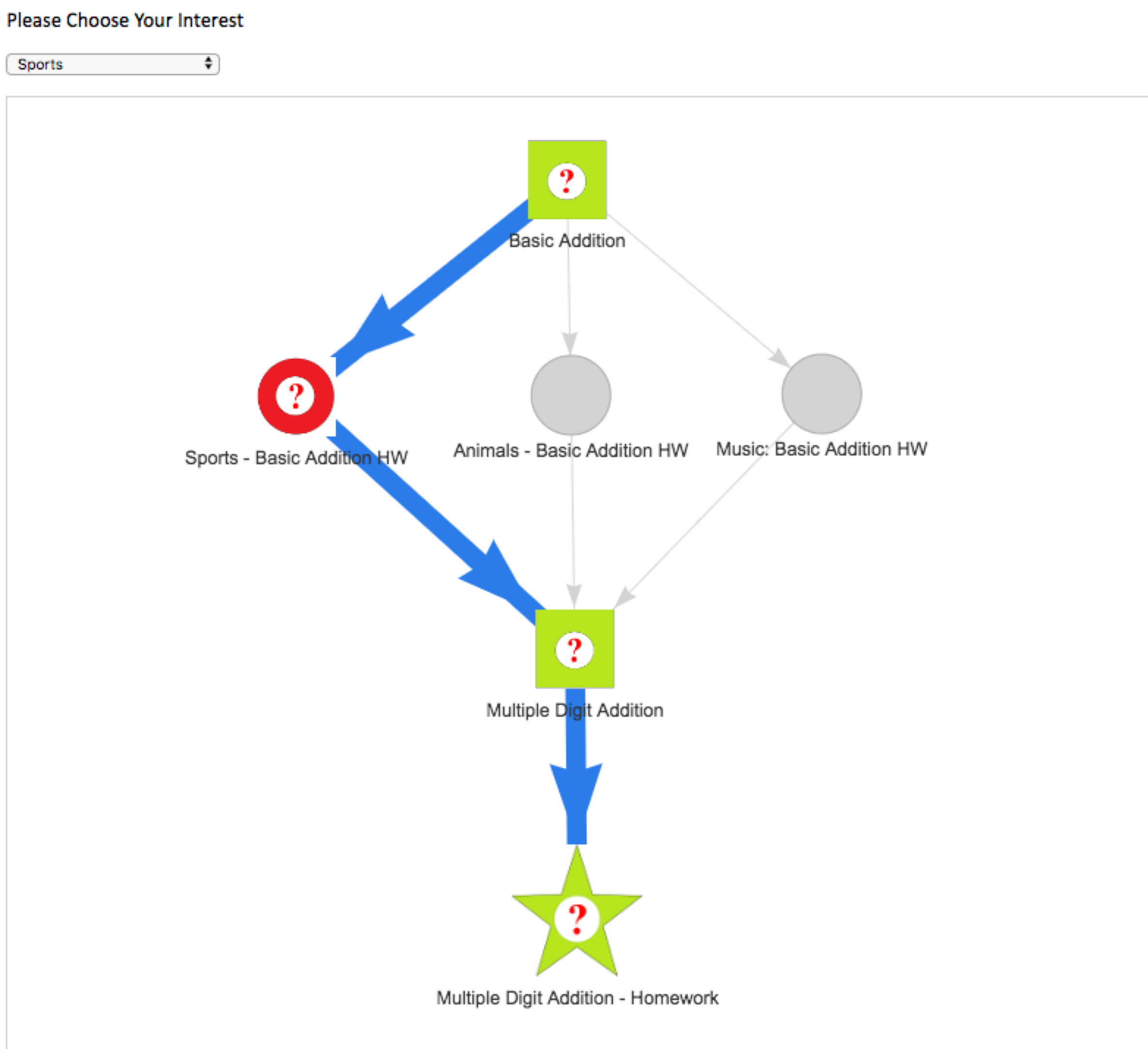

Figure 4.7: SAIL - Interactive Knowledge Map Student View

We use a series of symbols within the icons to demonstrate whether an activity is unknown, known, or graded for each student. The red question-marks inside the nodes in Figure 4.7 indicate that the student has not yet completed that lesson. A yellow check-mark indicates a complete activity, and a green check-mark indicates it has been graded. The final node in the course is indicated by a star-shaped icon to symbolize the finish-line or goal to be attained.

SAIL provides students with both an adapted and adaptable learning experience - the system adaptively suggests the path it thinks relates best to a student's interests, but allows the student to change paths and visit different exercises at any point without losing progress. This allows students to become an active participant in their own learning experience. Other studies have shown that interest-based learning is most impactful when students additionally have a choice in how to personalize their education experience [30].

4.3 Exploring SAIL's Potential

STEM Motivation

In a traditional instructional settings, students rely on the instructor or textbook to gain information and practice problems about a subject. The content and level of difficulty will vary from instructor to instructor. An adaptive curriculum that emphasizes the interdisciplinary nature of STEM subjects could attract, motivate, and retain students while ensuring an overall quality of education. Accentuating this interdisciplinary nature of STEM has the potential to attract a larger and more diverse population of students based on many studies of women and minority attraction and attitudes towards STEM fields such as Computer Science. This is especially impactful in K-12 settings, where a lack of trained STEM instructors can mean a lack of prepared students coming from the already limited number of high school courses. There is a need for more STEM instructors as well as a more adequate and engaging STEM curriculum. To begin assessing the viability and impact of SAIL, a pilot study was conducted in an introductory programming class at the University of Georgia. Our next chapter explores the reason for choosing computer science as the context for our initial study and discusses the areas, specific to computer science education where SAIL could have positive impacts.

5 | SAIL for Computer Science

This chapter uses significant portions of textual materials, graphs, tables, and/or figures from Aguar et al. 2017 (in-press) "Reviving Computer Science Education through Adaptive, Interest-based Learning" ©2017 IEEE.

5.1 Motivation

Technology is deeply integrated into our modern society. With its broad impact in nearly every industry, there is a desperate need for workers skilled in technology and programming - i.e.: Computer Science. Even with high salaries and exciting companies, the growing demand for technology professionals far exceeds the number of qualified graduates. Many challenges prevent the widespread education of Computer Science including

- A dominating misperception surrounding Computer Science/coding
- A lack of resources such as qualified instructors
- A competent and compelling curriculum

In a traditional instructional settings, students rely on the instructor or textbook to gain information and practice problems about a subject. The content and level of difficulty will vary from instructor to instructor. This is especially impactful at the K-12 level where a lack of trained CS instructors can mean a lack of prepared students coming from the already limited number of CS high school courses. More CS instructors need to be trained in CS concepts and need help developing an adequate and engaging curriculum.

An adaptive curriculum that emphasizes the interdisciplinary nature of CS could attract, motivate, and retain students while ensuring an overall quality of education. Accentuating this interdisciplinary nature of CS has the potential to attract a larger and more diverse population of students to the field based on many studies of women and minority attraction and attitudes towards computer science.

Based on the underlying problems existing in CS education, I propose the use of SAIL, a System for Adaptive Interest-based Learning - for Computer Science (SAIL-CS) to deliver introductory programming curricula to both students and educators to help attract, retain, and diversify CS.

5.1.1 The Need for K-12 CS Education

With technology impacting nearly every part of our lives, culture, and society, the demand for computer scientists has become a necessity in the job industry. A recent study reported that non-technical fields such as healthcare, banking, and manufacturing account for over half of the unfilled technology job demand [64]. Even though nearly everyone interacts with technology on a daily basis, the number of people educated in computer science (CS) falls far behind the need for technology employees.

The projected growth of CS related jobs significantly exceeds the projected growth rates of other occupations, including other STEM fields [5, 65]. CS enrollment is on an upward trend [66], but even if this trend continues, predictions estimate over one million unfilled computer science jobs by 2020 [65].

A lack of exposure to CS hinders major selection in college. Freshmen entering college are often not aware of what Computer Science entails because they have never had a course [2]. Even though today's infrastructure revolves around technology, Computer Science is not a required course in the U.S. K-12 system. Statistics show that 90% of parents think CS should be a part of their child's K-12 curriculum while only 40% of schools offer it [1]. Most schools who do include CS in their curriculum offer it as a high school elective not counting towards graduation [2, 20].

CollegeBoard (collegeboard.org) reported in 2011 that students' major selection in college was heavily influenced by any AP exams they took in high school [21]. The number of students in the U.S. taking the AP CS exam has been well below the national numbers for comparable subjects due to the low enrollment in the already low number of courses offered at U.S. high schools. A 2007 report from CollegeBoard showed that females students who took the AP Computer Science exam were **ten times** more likely to select CS as a major [67] in college. This makes the argument that exposure to CS should occur before college.

In Fall 2016, a new "AP computer Science Principles" (APCS-Principles) course launched nationwide aiming to broaden participation in CS by introducing high school students to computational thinking and how computers affect the world [22, 23]. This focus on computational thinking - defined as "encompassing a set of concepts and thought processes from CS that aid in formulating problems and their solutions in different fields" [20] - differs from standard introductory programming courses that teach syntax and debugging skills. The Principles AP course is not meant to replace, but to complement the existing AP Computer Science course - which has now been renamed to AP Computer Science A (APCS-A) [23]. Figure 5.1 below compares the topical outlines for both courses.

AP CS- Principles	**AP CS - A**
Creativity	Object-Oriented Program Design
Abstraction	Program Implementation
Data and Information	Program Analysis
Algorithms	Standard Data Structures
Programming	Standard Operations and Algorithms
The Internet	Computing in Context
Global Impact	

Figure 5.1: Comparing the APCS-P and APCS-A Curriculum Outlines

APCS-Principles has the potential to peak student interest in technology and perhaps positively influence the perspective of CS, but does not replace learning fundamental CS Programming skills [22]. The two AP Computer Sciences courses (Principles and A) can be taken independently, in either order [23]. However, AP Computer Science A naturally succeeds AP Computer Science Principles for students wanting to continue their education in CS as it teaches essential syntax, debugging, and logic in a Java based environment. Unfortunately, the AP Computer Science A course has far fewer resources and many challenges preventing success [24].

5.1.2 The Need for More Resources

With the desperate need to train more computer scientists, it follows that there are a lack of trained CS educators at the K-12 level. To better educate teachers about CS and how to best incorporate it in their classrooms, numerous free teacher training opportunities have been made available [68–71]. More resources are needed as even willing teachers are often overworked, underpaid, and lack the time needed to develop a sufficient CS curriculum when they are just learning the concepts themselves. As more teachers are educated in CS and more classes are offered at the high school level, ensuring that the content taught adequately prepares the students for the next university level course(s) in CS is paramount.

More often than not, students entering the University of Georgia (UGA) with high grades in their high school Computer Science course should be allowed to bypass the CSCI 1301 entry-level programming (Java 1) course, but are unprepared. It has been our experience here at UGA that exempting students from our 1301 course and beginning with our CSCI 1302 (Java 2) class has set the student up for failure. It should be noted that the incoming students at the University of Georgia are of high caliber, with the incoming freshman of 2016 possessing a 3.98 GPA on average [72]. Though they have credit for the entry-level programming course on paper, they are ultimately unprepared to continue their CS education. Setbacks like these can discourage students from pursuing the major. More resources are needed at the K-12 level to ensure an adequate foundation in CS concepts is established. Educating more K-12 teachers in CS concepts and increasing the number of high-school classes is insufficient if the content being taught is not inadequate.

5.1.3 Community College Impact

The same argument for ensuring a more standardized CS curriculum can also be applied to community colleges. It is becoming a more common practice for students to complete their "core work" - 30-59 hours at a community college before transferring into a higher ranked university [73]. While the caliber of these colleges differ from universities, ensuring that students gain an adequate foundation in their CS curriculum is paramount for the retention of transfer students as majors, and in producing graduates of the expected merit.

5.2 Diversity Awareness and Impact

In addition to the underpopulation of computer science, there is also a lack of diversity within the field. According to data from the National Science Foundation, in 2014 only 18% of all Computer Science Bachelor's degrees were awarded to females. Students identifying as either black / African American or Hispanic combined received less than 20% of all awarded bachelor's degrees in CS [74].

Attracting a more diverse group of computer scientists can help grow the field, but is also important in the cultivation of today's technological society to avoid biased products, better cultivate innovation, and better overall production [6]. In Computer Science in particular, having predominantly men working on a software could produce gender-biased end-results that are disconnected from over half of the population.

Computer Science is applicable in nearly every industry, field, or hobby in today's society - from healthcare to finance to physical education and more - but Computer scientists are often stereotyped to be "nerdy" and solely interested in technology [11,18]. A study where women were shown a computer science classroom filled with the typical stereotypes (i.e.: StarTrek posters, gaming consoles, etc.) reported that women were less inclined to enroll in a CS course than women who were shown a classroom without the stereotypical props [75, 76]. Correcting the image associated with CS to make it inclusive to a wider variety of students can have a huge impact on who pursues CS as a major/career. [17, 77]. Efforts such as breadth-first introductory CS courses and earlier research opportunities to explore the interconnectivity of CS with other fields have shown encouraging results. Multiple studies have reported success in attracting more women to their programs by simply changing the way that introductory programming is advertised [2, 14, 17].

A nationwide effort to incorporate these changes came with the launch of the new AP- Computer Science Principles course (Fall 2016). This course has a high potential to successfully attract a more diverse population of K-12 students to the CS field. If successful in attracting more CS students to the field, measures will need to be taken to ensure that students are retained and remain interested beyond this initial introduction. Described above was the need to improve fundamental CS K-12 courses, such as the AP Computer Science A course offered in high schools nationwide. In addition to acquiring sufficiently trained teachers and a more standardized curriculum is the need to ensure students enjoy their fundamental programming experience and remain excited about CS. Programming should be taught in a way that reiterates the interconnectivity of CS with other fields to be inspirational and interesting to a diverse population of students with differing cultures, races, background knowledge, and interests.

5.3 SAIL-CS

SAIL is an innovative new web-based adaptive learning system that customizes a student's individual development path (IDP) based on their interests. The overall construct of SAIL can be applied to any discipline that has a broad applicability to other domains - including most STEM fields (other CS courses, Mathematics, Engineering, etc). This proposal focuses on a pilot study (SAIL-CS) using adaptive, interest-based learning to provide improved, standardized, adaptive solutions for introductory courses at the K-12 and community college levels.

The development of SAIL-CS (System for Adaptive Interest-based Learning - Computer Science), could be the solution to deliver customized introductory CS curriculum based on students' interests to help achieve better attraction, diversity, retention, and a standard of competency for students at the high-school or community-college level.

With SAIL-CS, each student will indicate subjects or topics of interest to them. Based on these interests, SAIL-CS will create a unique learning path to guide the student throughout the learning process. SAIL-CS aims to help change the way CS is perceived by showcasing the integration of CS with other fields while effectively teaching students much needed technological skills.

5.3.1 SAIL-CS and the Instructor

The proposed system will follow the curriculum standards for APSCS-A and introductory university level CS1 (programming) courses by leading students through a linear ordering of topics that students must complete irrespective of instructor. In traditional instructional settings, students rely on the instructor or textbook to gain information and practice problems about a subject. The content and level of difficulty will vary from instructor to instructor. This is especially impactful at the K-12 level where a lack of trained CS instructors can mean a lack of prepared students coming from the already limited number of CS high school courses. SAIL-CS eliminates this problem by providing the teacher with meaningful and adequate content that is ready-to-use in the classroom.

In addition to providing a more standardized, meaningful educational experience to students, SAIL-CS can also be used both to train educators and to provide a standardized level of content across K-12 and community college CS courses nationwide. With the lack of CS instructors, many training workshops have surfaced aiming to educate teachers (specifically K-12 educators) in introductory programming topics and provide curriculum resources for them to implement in their classrooms [68–71]. Although many report success, more trained instructors and better resources are needed. SAIL-CS can act as a training tool for educators - ensuring they too are provided an adequate and motivating curriculum. By using the same tool to *teach the teachers*, educators will become familiar with the software to later implement in their classrooms while gaining fundamental CS competencies adapted to their interests. This tool could help educate more instructors nationwide and ensure that the information they pass on to their students is at a competent level. The burden is then taken off of the teacher to create instructional material to inspire and adequately educate students as SAIL-CS does the heavy lifting.

5.3.2 Interest-Based Learning

Another common issue in the traditional educational setting is a lack of meaningful connections between what is being taught and how these concepts apply to the real-world. For example, a student may learn how to print out the numbers 1-10 with a for-loop but not understand why learning such a concept could be applicable. Even when instructors provide examples to show the interconnectivity with the real world, it is impossible for a single example to be of interest to all students in the class. An example problem that uses baseball to teach some fundamental skill would intrigue only a subset of the students involved in the course, as some students may be uninterested in sports or may be from different backgrounds and not understand the rules. Problems such as these exist in the traditional instructional setting where all content, practice problems, and assignments depend on the unique instructor and are universal for all students. SAIL-CS addresses this problem by adapting the learning content to a student's individual interests to ensure that each student is taught new ideas in a meaningful way.

With SAIL-CS, A high school student who has enjoyed past English/grammar courses could be introduced to the concept of iteration (loops) by implementing a word-count feature. Alternatively, a student interested in sports or physical education (PE) might find it more exciting to learn iteration by calculating an average heart-rate given a dataset of heart-rate readings over time - examples illustrated in Figure 5.2. Adapting content to personal interest can increase a student's intrinsic motivation as personal interest has been shown to significantly influence learning in regards to an individual's attention, goals, and level of cognition [3].

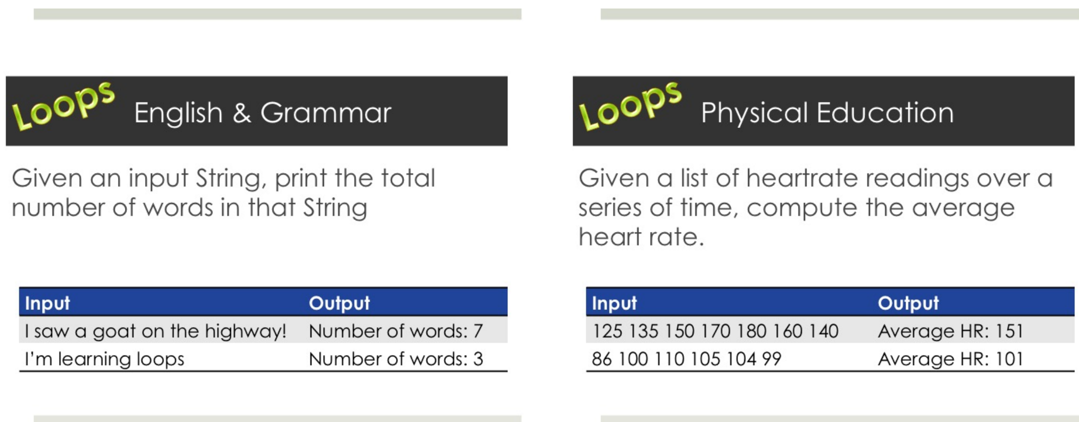

Figure 5.2: Example Interest-based Problems for Looping

5.4 SAIL-CS Design Details

5.4.1 Curriculum Design

SAIL-CS is meant to teach introductory Computer Science skills to adequately prepare a student for a subsequent university-level programming course. While the system itself is not limited to a particular domain, it will be used to teach the most fundamental programming concepts that would coincide with the APCS-A course or a Introduction to Programming course at the college-level. The APCS-A curriculum listed in Figure 5.1 provides a good example of the content and order expected to be taught. SAIL-CS will teach fundamentals via Java - the suggested language for the APCSA course [23] and most widely-used in entry-level college/university programming courses nationwide.

SAIL-CS is built as an evolutionary system, to constantly improve with additional time and resources. A small subset of "interests" for the examples will be implemented in the pilot version, but over time, teachers and the community may discover more diverse interests among students, and can come up with more examples and extend that level of the graph.

5.5 Conclusion

There are many challenges to successful widespread Computer Science education. I propose the use of SAIL (System for Adaptive Interest-based Learning) for Computer Science courses (SAIL-CS) to help address many of these issues. Specifically, SAIL-CS could:

- Address diversity issues by showcasing the interconnectivity of CS with other fields.
- Provide a standard of competency that each student should receive in K-12/community college CS1 courses.
- Attract and retain a larger population of students in CS
- Adequately train more educators in CS fundamentals

These issues in Computer Science education are mirrored throughout many other STEM domains. SAIL was created to help alleviate these issues in STEM education, with SAIL-CS proposed as the initial pilot study for testing and implementation. SAIL-CS aims to address the outstanding needs in both K-12 and community college level Computer Science education. This system can be used to educate both students and instructors in an intriguing context. The following chapter details the SAIL-CS pilot study, where several modules of the Computer Science curriculum were created and implemented in a University-level Introduction to Programming course.

The interest-based nature of SAIL-CS could increase intrinsic motivation of students as they learn fundamental programming skills through avenues that are already exciting to them. The misperception that Computer Science is solely about technology could be corrected by highlighting the interconnectivity of Computer Science with a variety of other subjects. Widespread adoption of this system could help educate more instructors and students to an adequate level of competency and recruit and engage a larger, more diverse representation of students in Computer Science.

6 Pilot

In this chapter we outline how we incorporated SAIL into an Introduction to Programming course at the University of Georgia. We present all relevant information about the experiment including the setup of treatment vs. control groups, pertinent demographic information, which variables we were able to control, the design of SAIL content, how SAIL was implemented, and the data collection measures utilized. This study was conducted with prior approval from the Institutional Review Board (IRB).

6.1 Participants

SAIL was tested in the context of introductory computer science programming (CSCI 1301) at the University of Georgia during the Fall 2017 semester. Four Introduction to Programming (CSCI 1301) courses were taught by three instructors with a total of 307 students participating in the study. One of the four courses (47 students) was used as the treatment group, where they interacted with SAIL for all lecture content and assignments. The remaining three courses were used as control groups for comparison - see Figure 6.1.

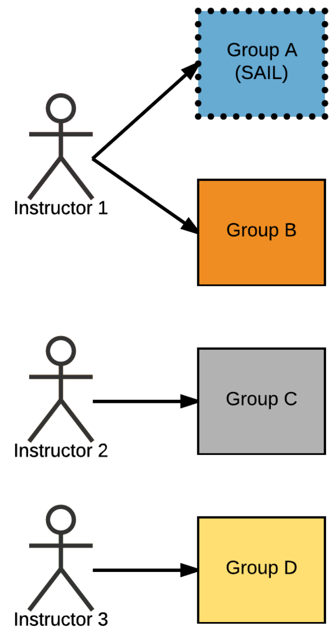

Figure 6.1: Pilot Study Treatment and Control Groups

Group A (the SAIL treatment group) and B were both taught by the same instructor, eliminating any concern for instructor bias. These courses were designed to follow the same flipped-classroom approach, described in Section 6.2. Group C and Group D were taught by separate instructors who have demonstrated excellence is teaching abilities. Both Group C and Group D followed the traditional lecture style classroom setting with in-class lectures and at-home exercises.

Group A was chosen to act as the treatment group for the following reasons:

1. Instructor 1 taught two sections - Group A and B. In order to compare the treatment group to a control group without instructor bias, either Group A or B would be the best choice.

2. Groups A and B were being taught with a hybrid (flipped-classroom) approach to instruction. Though SAIL could be implemented in a traditional lecture-style classroom, the video aspect of SAIL lends itself well to independent learning outside of the classroom - making SAIL well suited for online and hybrid courses. This strengthened that either Group A or Group B would be the best choice as the treatment group.

3. The classroom assigned to Group A was equipped with computers and the classroom assigned to Group B was not. As SAIL was to be used in-class, to ease the expectation that students bring their own computers to class, Group A was chosen as the treatment group.

6.1.1 Demographics

Demographics for students were collected via a survey approximately 8 weeks into the semester, one week prior to the withdrawal deadline. Due to the timing of collecting demographic data, this data may more accurately reflect the demographics of students retained in the courses, and not of students who were originally registered and may have dropped the course.

Demographics across all sections consisted of 70% male and 30% female participants, while students in the treatment group were 62% male and 38% female. Figure 6.2 provides the breakdown of gender per group, as well as a visualization of the enrollment differences between sections. Numbers inside the bars represent the total count of for each gender per group, with the percentage of female students per group included for easy comparison among sections.

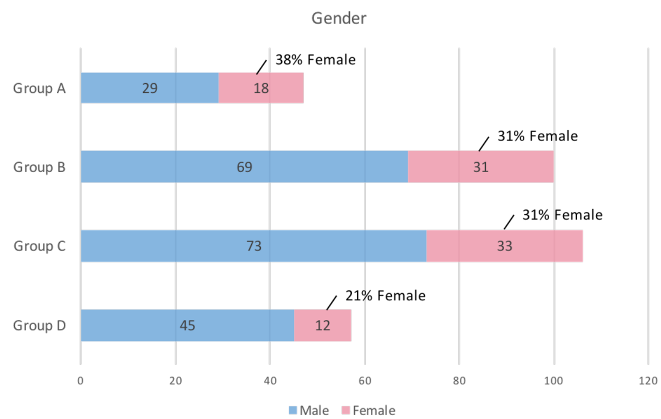

Figure 6.2: Gender - All Groups

The 2016 Taulbee survey indicated that the male to female ratios of enrolled CS students was 81.7% male, 18.3% female, with similar distributions for earned bachelor degrees (82.1% male, 17.9% female) [7]. When compared to the 2016 Taulbee Survey data, our female enrollment appears higher than expected; however, not all students enrolled in this introductory course are majors so our enrollment by gender in this course does not directly translate to CS major enrollment. The total number of CS majors across all sections is 173 (approximately 56% of all participants), with 75% male and 25% female. These distributions are still slightly higher than expected, yet closer to the national data from the Taublee Survey. All courses individually exceed the national averages for female enrollment. With high female enrollment in all sections as both majors and non-majors comes the opportunity to gather meaningful data about how male and female performance in introductory CS may differ.

Ethnicity demographics across all sections can be seen in Table 6.1. Figure 6.3 provides the breakdown of ethnicity by group, as well as a visualization of the enrollment differences between groups. Numbers inside the bars represent the total count of each student identifying with that ethnicity per group.

	Group A		Group B		Group C		Group D		Total	
White	28	59.6%	54	54.5%	56	53.8%	19	33.3%	157	51.1%
Black or African American	2	4.3%	5	5.1%	10	9.6%	5	8.8%	22	7.1%
Hispanic or Latino	3	6.4%	11	11.1%	6	5.8%	7	12.3%	27	8.7%
Asian	13	27.7%	27	27.3%	29	27.9%	24	42.1%	93	30.0%
Other	0	0.0%	2	2.0%	3	2.9%	2	3.5%	7	2.3%
Total Known Ethnicity	**46**		**99**		**104**		**57**			**0.0%**
Ethnicity Unknown	1		0		0		0		1	
Grand Total	**47**		**99**		**104**		**57**		**307**	

Table 6.1: Ethnicity Statistics by Group

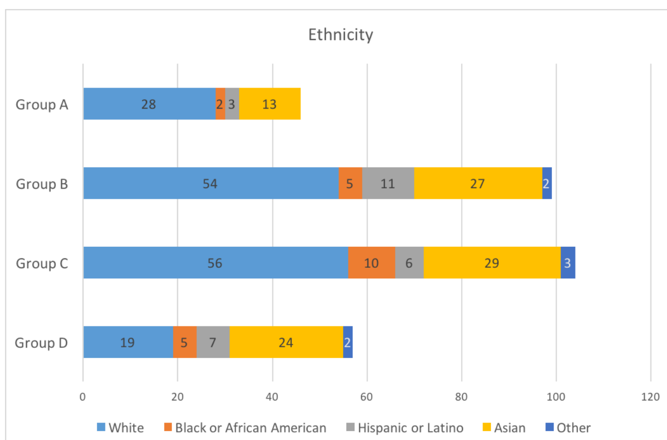

Figure 6.3: Ethnicity - All Groups

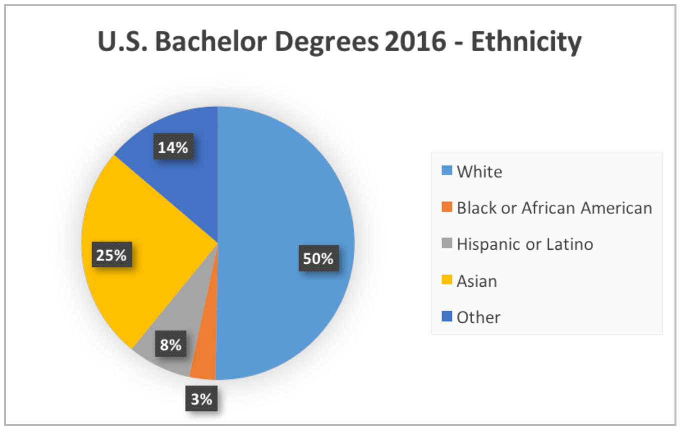

Figure 6.4: Ethnicity - National CS Bachelor Degrees
Data from: 2016 National Taulbee Survey [7]

Where "other" includes Native Hawaiian/Pacific Islander, Multiracial, Non-Hispanic, American Indian or Alaska Native, and Nonresident Alien. [7].

The ethnicity data as a whole for our pilot mirrors the ethnicity data expected for a computer science class based on the data from the 2016 Taulbee Survey presented in Figure 6.4. The enrollment of black or African American students is slightly elevated from the national average (7% compared to 3%); however, some of our sections feature very low enrollments (2 or 3 students). Due to the very small sample sizes of some minority groups (black/african american and hispanic/latino), we may be unable to draw conclusions from qualitative and quantitative data regarding SAIL's impact differing by ethnicity.

6.2 Learning Environment

All groups had 75 minute class times two days a week. Recent studies of college aged individuals' sleeping and performance patterns indicated that early start times for classes may negatively impact learning [78]. Our sample featured varied start times (Group C - 9:30am, Group A - 12:30pm, Group B - 2:00pm, Group D - 3:30pm), with the control group in the middle of this range to help control for any bias caused by class time.

SAIL was used to teach two modules in an introductory level computer science course. Module 1 covered Branching and Module 2 covered Looping - both fundamental topics in introductory programming.

Group A used SAIL both at home and in class as part of a hybrid approach to instruction. Students would watch lecture videos in SAIL at home - short quizzes were given to ensure student's were progressing through content at home. In place of a traditional lecture, students grouped up based on their interests and practiced with the adaptive interest-based exercises. This hybrid approach allowed instructors to facilitate deeper discussion to reinforce the content students learned through the videos in SAIL and gave students the ability to ask questions and gain assistance during class time.

While SAIL could be utilized in many classroom designs such as a distance or online learning environment, implementing SAIL in a flipped classroom allowed us many benefits for our study purposes:

1. Group B, taught by the same instructor as Group A (The SAIL treatment group) is also a flipped course. This seasoned instructor has taught this course many times as a hybrid approach to instruction - utilizing the best known techniques for a flipped

classroom. Both Groups A and B follow a nearly identical course structure, where students in Group B gains their competencies outside of the classroom through at-home textbook reading assignments, take the same reading quizzes as Group A who watched SAIL videos at-home, and work as groups during class time to complete practice exercises. By implementing the same learning framework, we were able to control many variables that may have influenced our results for comparison such as (1) the instructor, and (2) a flipped classroom approach. Figure 6.5 details the similarities and differences of the learning structure for Groups A and B.

2. Students in Group A (the SAIL treatment group) had three weeks of experience with a normal, hybrid approach to instruction before SAIL was incorporated. These students can offer insight into a comparison between experiences using SAIL without the bias of traditional-lecture vs. flipped approaches to learning.

3. The instructor did not have to change their teaching style. They simply assigned SAIL modules for at-home content instead of textbook readings and instead of printing in-class handouts, they merely instructed students to complete their selected in-class activity in SAIL.

4. In addition to comparing SAIL with another hybrid class, we have the opportunity to compare SAIL to two separate sections (Groups C and D), taught by two separate teachers using the traditional lecture-style approach for teaching.

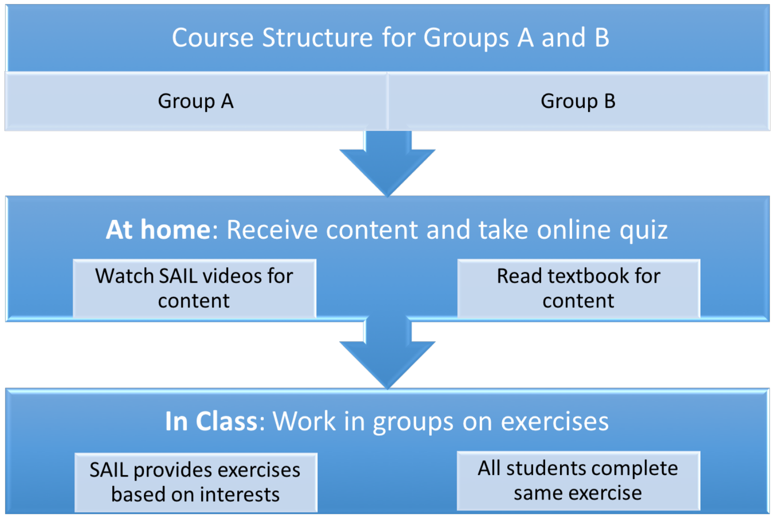

Figure 6.5: Group A and B Course Structures

6.3 SAIL Content and Procedures

The treatment group used SAIL as the instructional medium for six weeks. All topics related to Branching and Looping were covered. As students used SAIL as part of a flipped-classroom approach, we divided content into two sections: (1) Lessons - completed at home and (2) Exercises - completed in class. The sections below describe these contents in detail.

6.3.1 Lessons

Students in Group A had previously been completing reading assignments in the course textbook at-home to gain new knowledge competencies. After completing their textbook reading, students took an online "Reading Quiz" - a simple quiz to ensure they read the content and were gaining basic competencies. After three weeks of instruction, we began using SAIL in the treatment group. While Group B continued their textbook reading assignments at home, Group A transitioned to completing specified lesson nodes in SAIL. For lesson delivery, we created short, professional quality instructional videos covering the same content that would have been in the textbook reading. The videos featured both lecture slides outlining major concepts, and heavily incorporated coding examples. Instructional videos were created by the researchers, also seasoned 1301 instructors, and lasted approximately 20 minutes each - Table 6.2 shows an outline of all videos created for this pilot. Students in Group A and B both completed the same "reading quizzes" after completing their at-home assignment.

Branching Videos	
Video 1	The basic if-else statement Boolean expressions Java comparison operators Compound boolean expressions
Video 2	Terminating a program The conditional operator Practice with Boolean operators When not to use == Comparing Strings
Video 3	Nested branching statements Multi-branch statements
Video 4	Short circuiting Switch statements Enumerated types
Looping Videos	
Video 1	The while statement The do-while statement Infinite loops Input Validation
Video 2	Nested loops
Video 3	The for statement The break and continue statements The for-each statement

Table 6.2: Outline of videos created for SAIL-CS

6.3.2 Exercises

SAIL adapted practice problems for in-class exercises based on student interest. The interest categories available within SAIL-CS were:

- Sports
- Entertainment
- Science

Only three interest categories were available in the pilot - as it has been well documented that creating adaptive content is time consuming. The goal was to come up with three very broad categories at this time so that most students would identify with one of these three as somewhat interesting. As part of SAIL's vision is to facilitate community knowledge sharing of adaptive content, interest categories could become more narrow and perhaps more intriguing to students with more time and resources. Keeping in mind that this initial pilot had a limited selection for interests, any positive results seen from adapting exercises in this broad scope of interest has the potential to improve over time.

When using SAIL, students were able to indicate which of these categories interested them most. Based on the student's selection, SAIL created a customized path for each student to receive practice problems based on their interest. SAIL, therefore, adapted the content and path to the individual student (or group of students), but was also adaptable – in that students selected their interest and could change this selection at any time. By allowing students to choose their interest, we are guaranteeing that students do indeed get problems interesting to them throughout the entire duration of the study. Figure

6.3 shows a sample problem that was adapted from the generic version for Branching practice and Figure 6.4 shows an adapted exercise for Looping practice.

All adapted exercises followed a very precises adaptation to ensure that students in both groups (treatment and control) were receiving the same types of problems at the same difficulty level. This helped ensure that students in both the treatment and control group were practicing the same skills at the same level, so we could best assess the true impact SAIL had on student performance and overall experience with as many variables controlled as possible.

Branching In-Class Exercise Sample Problem	
Interest	Problem Description
Generic	Magic 8-ball: Write a magic 8-ball program. The program will randomly select amongst 10 different outcomes and print one to the user. Have fun with this. You can have the user enter a question or they can just have the question in their heads before running the program. Here are the possible outcomes: Ex: It is decidedly so, Without a doubt, Definitely, ...
Sports	Fantasy Football Name Generator: Fantasy football is a growing hobby where people create imaginary football leagues, scout and draft players, and compete against other fantasy teams for the championship. To create a fantasy league, you first need to name your fantasy league. Come up with 10 possible names for your fantasy football league and write a program that randomly selects your team name from those 10 options. Print the selected name to the user.
Science	Time Travel: You've been tinkering in your garage and invented a machine that allowed you to travel space and through time. The problem is, that it won't let you pick the time period. Write a program that will decide which of 10 time periods you will be sent to. Below is a list of 10 possible time periods – feel free to customize these with any historic event or time period. Ex: Prehistoric, Middle Ages, Renaissance, Victorian Era, ...
Entertainment	Shuffle Play: iPods and most music streaming apps like Spotify or Pandora allow you to "shuffle" the order in which you play songs in a playlist. Write a program that randomly selects from 10 songs and print the name of that song chosen to play to the user. Have fun with this. You can customize the 10 songs on your playlist – Here's an example playlist: Ex: "Bye bye bye", "Georgia on my mind", "Twinkle Twinkle Little Star", ...

Table 6.3: Branching In-Class Exercise Sample Problem

Looping In-Class Exercise Sample Problem	
Interest	Problem Description
Generic	Assume a user inputs a String s that contains at least one character. Print the number of lowercase letters in the string. The program should work for any valid input String s. Examples Enter a String s: BanAnA Output: BanAnA has 3 lowercase letters
Sports	Soccer Win/Loss/Draw: In soccer, each team must report its wins-losses-and draws for the entire season. For each game played, the team reports either W, L, or D symbolizing a win, loss, or draw on that game. Assume a user inputs a sequence of soccer stats s that contains at least one character. Print the number of wins (W) for that season. The program should ignore the case of the letter and work for any input String s. Examples Enter a String s: WLLDWDLWD Output: WLLDWDLWD has 3 wins
Science	DNA Sequence Analysis: A DNA sequence is a succession of letters that indicate the order of nucleotides within a DNA. The possible letters are A, C, G, and T, representing the four nucleotide bases of a DNA strand — adenine, cytosine, guanine, thymine. Researchers analyze these DNA sequences to understand its features, function, structure, or evolution (source: Wikipedia). Assume a user inputs a DNA sequence s that contains at least one character. Print the number of adenine bases (a) found in the DNA sequence. The program should ignore the case of the letter and work for any input String s. Examples Enter a String s: AAACCCTTAG Output: AAACCCTTAG has 4 adenine bases
Entertainment	Twitter Hashtags: You have a new job working for Twitter. They want to know how many hashtags (#) people normally include in their tweets. Assume a user inputs a tweet t that contains at least one character. Print the number of hastags (#) found in the tweet. The program should work for any valid input String t. Examples Enter a tweet: Go dawgs! #UGA #ATD Output: Go dawgs! #UGA #ATD has 2 hashtags

Table 6.4: Looping In-Class Exercise Sample Problem

6.4 Data Collection

Both quantitative and qualitative data were collected from the 307 participants who consented to participate in the study.

6.4.1 Quantitative

Reading Quizzes Short, online quizzes meant to ensure students progressed through course competencies at home. Only Groups A and B (controlled instructor) participated. The treatment group (Group A) watched specified SAIL videos on branching and looping concepts while Group B read the specified sections of the textbook before taking each quiz. Students in each section completed five reading quizzes in total (three for Branching and two for Looping). As this quizzes were given in an uncontrolled environment (at home, any resources allowed to answer questions), they may not be a good measure to compare performance between students in Groups A or B but can offer insight into if students are progressing through content with SAIL in an adequate way.

Quizzes Two quizzes (one for branching, one for looping) were given immediately after students completed these modules in all courses to evaluate performance. Quizzes were designed to have 5-6 multiple-choice questions, all increasing in difficulty. Each question featured a fragment of code and four multiple-choice options for the code's output. Students from all groups took the same quiz to allow data comparison across all courses. The questions in this quiz are "neutral" in context - where the content and subject does not favor a particular interest. Students in the control group will be accustomed to receiving "neural" content in both the branching and looping modules. By evaluating with "neutral"-themed questions, we are able to test if SAIL students can transfer their knowledge gained through interest-based exercises to non-related (or non-interesting) domains.

Ideally, quizzes would have been counted as a graded item across all courses to help ensure accurate student assessment. Unfortunately, this is the ultimate decision of the instructor and we were unable to control two of the external factors that may influence success. In Groups A and B, the instructor chose not to count quizzes as a grade, but rather gave the quizzes unannounced, allowing them to count for "bonus" points on a student's overall grade. Groups C and D did allow these quizzes to count for a grade, giving students advance notice that the quiz would take place for adequate preparation. This introduces many uncontrolled factors (extrinsic motivation, incentive to prepare, and adequate time to prepare). Limitations such as these are the realistic nature of educational studies where one cannot control all factors that may influence success. We do, however, still include these measures in our analysis, as results are still important, though they should be viewed through the lens of the limitations outlined above. Due to the unevenness of these assessments across sections, we chose to look at grades for Exam 1 and 2, described below, to offer us more insight into comparison of assessment.

Exams Due to the limitations of the quiz assessments detailed above, we put more emphasis on student performance on two separate exams:

- **Exam 1:** Exam 1 was the first heavily weighted item (15% of the overall grade), given to students in all sections under the same conditions. It was created by the most seasoned instructor, following similar formats from past semester where questions consisted of: true/false, multiple choice, short-answer, and code-writing problems. All four classes were given the exact same exam on the same date, counting for the same percentage of their overall grade. This is considered a highly controlled assessment for comparison of student performance among all sections.

- **Exam 2:** Exam 2 was the second heavily weighted item (20% of the overall grade), given to students in all sections under the same conditions. Like Exam 1, this exam was also created by the most seasoned instructor and followed a similar format. Again, all four classes were given the same exam, on the same date, counting for the same percentage of their grade. However, as the exam grades overall may offer insight into comparison of performance between sections, there are many uncontrolled factors about how SAIL was used in preparation for this assessment that we must consider. We describe these limitations below.

Exam 1 was given six weeks into the semester, covering all topics taught thus far in the course. In preparation for the exam, students in all sections underwent their traditional instructional process for the first three weeks, without any intervention in the treatment group. During this time, students were introduced to the basics of programming, covering topics such as declaring variables, data types, arithmetic, and System I/O. In the three weeks leading up to the exam, SAIL was introduced as the intervention to the treatment class where they learned all Branching-related concepts for the course. As described above, SAIL was used as the primary instructional tool both in class and at home. In assessing performance on Exam 1, it must be taken into consideration that students utilized SAIL in the three weeks immediately prior to the exam, learning what were the most difficult concepts in the course at that point. In an analysis of Exam 1, 67% of questions directly correlated to Branching concepts taught with SAIL, while the other 33% tested on foundational skills. We feel that because students must understand basics of programming such as data types and basic syntax before successfully learning Branching skills, all questions on this test are relevant to assess.

The conditions under which Exam 2 was adminsitered are the opposite of the setup described for Exam 1. Exam 2 was given ten weeks into the semseter, testing over all content (approximately 4 weeks of material) since Exam 1. In this four week span, students in the treatment group used SAIL as the intervention for the first two weeks - to learn all Looping related topics. In the two weeks leading up the exam, the intervention was removed and students assumed their traditional instruction. Unlike Exam 1, where the content taught with SAIL was the pinnacle of the exam, Exam 2 tested on Looping concepts and as well as other, more advanced topics (arrays and the basics of classes/objects). Due to this, we did not feel Exam 2 grades as a whole would tell us much about about how SAIL might have impacted performance in the treatment group. In analyzing Exam 2, we determined that 10 out of 33 questions related only to Looping topics (taught with SAIL in the treatment group) with no inclusion of future topics. We extracted these 10 questions from the exam and compared performance across all sections to best assess how SAIL may have truly impacted student knowledge of Looping topics. Unfortunately, there are still many uncertainties about how reliable and controlled this measure is, as students in all sections had more time to practice looping in advanced domains (looping through arrays, writing an equals method that involves looping, etc). Unlike Exam 1, Exam 2 does not test students immediately after utilizing the intervention for knowledge acquisition. Also, with only 10 questions on the exam, we feel students may not have been as thoroughly tested on Looping knowledge (in isolation from other topics) as we were able to measure with Branching in Exam 1. Keeping these differences and possible limitations in mind can help us best interpret the data, and we believe comparing performances even with these uncontrolled factors can still provide insight into SAIL's impact on performance measures.

6.4.2 Qualitative

Like with most educational studies, we believe that quantitative data only tells part of the story. As we have seen, quantitative data often leads to many questions as to what may or may not cause variances among the data. To fill in some of the gaps, we chose to conduct a survey as an additional measurement of SAIL's impact in many areas.

This survey doubled as the consent process for the study, which means that we only quantitatively assess the grades of students who responded to the questionnaire. This provides an even ground for comparing quantitative and qualitative results, as we are looking at the same subset of students in both parts of the data. The survey was optional, with incentives provided to increase participation, and completely anonymous to anyone other than the researchers to avoid any bias caused by students feeling their responses may affect their grade in the course. The 307 participants are the students who opted into the study, as almost all students in each class chose to participate. With any data, knowing the background, demographics, and other surrounding variables of students can help us understand the responses and any trends found. For this reason, many questions were asked to help best inform our results. Student demographic information was collected to help determine if SAIL's impact on performance, motivation, or perception differed based on these trends. We also collected information about whether students had prior exposure to computer science before taking this course and whether or not they were a CS major.

Students were asked to rate their agreement with a series of questions on a 1-7 likert scale where 1 = strongly disagree...7 = strongly agree. Students from all sections were asked the same questions to assess their perception and enthusiasm towards the course, computer science, and programming. Additionally, students in the treatment group were asked a series of questions about their overall experience using SAIL. Table 6.5 shows the questions asked across all four sections to make inferences about perception, enthusiasm and interest in CS. Table 6.6 shows additional questions we asked the treatment group to gauge their overall experience with SAIL as well as how the different aspects of SAIL (interest-based exercises and lecture videos) were perceived.

Please indicate how strongly you agree or disagree with the following statements:
I am interested in Computer Science
I feel programming has the potential to positively impact my other interests (ex: sports, art, ...)
I enjoy programming more than I expected
I plan to continue my Computer Science education after this course
I referenced the SAIL videos while studying or working on programming assignments.
I feel I understand the material in this course
I enjoy this course

Table 6.5: CS Perception/Enthusiasm Questionnaire

Please indicate how strongly you agree or disagree with the following statements about your overall experience using SAIL:
SAIL provided an overall positive learning experience
I felt the SAIL website was easy to use.
I would choose to use SAIL for learning future topics
Please indicate how strongly you agree or disagree with the following statements about your overall experience using SAIL:
I referenced the SAIL videos while studying or working on programming assignments.
I prefer watching the videos in SAIL to the traditional learning approach.
I felt the videos in SAIL helped me learn.
Please indicate how strongly you agree or disagree with the following statements about the video lectures in SAIL:
The exercises tailored to my interests increased my enjoyment of programming.
I found the adapted exercises relevant to my interests
I prefer the interest-based exercises to the generic exercises given to the entire class
I liked the option to adapt the class exercise based on my interest

Table 6.6: SAIL Perception Questionnaire

6.5 Research Questions

Below we describe which quantitative and qualitative measures will be utilized to help answer our intended research questions.

1. **RQ 1 - How does the use of SAIL impact performance measures?** We attempt to answer this question by comparing the quantitative measures for all Branching and Looping assessments, with the highest emphasis placed on Exams.

Assessments to be examined include:

- Reading quizzes
- Branching and Looping Quizzes
- Exam 1 and Exam 2

2. **RQ 2 - How does the use of SAIL impact perceived learning and confidence in the course material?** As confidence and perceived learning are subjective feelings, we rely on qualitative measures for our information. We ask students in all sections rank their agreement with the following statement on a 1-7 likert scale: "I feel I understand the material in this course".

3. **How does the use of SAIL influence students' attitudes and perception towards computer science (or STEM fields)?** We intend to study this through the combination of several qualitative measures. For each of the following statements, students will rank their agreement on the 1-7 likert scale:

   - "I enjoy this course" - A simple question that allows us to compare overall enjoyment of the course between sections.
   - "I enjoy programming more than I expected" - An alternative version of the previous question that eludes to a higher than expected enjoyment of the course. This is included as the literature review suggested a stereotype associated with CS. This question may give us some insight into expected vs. actual enjoyment of CS.

- "I am interested in Computer Science" - We want to gauge student interest in CS across sections
- "I feel programming has the potential to postiviely impact my other interests (ex: sports, art, ...)" - Do students in CS realize how impactful CS can be to other fields, specifically to other domains of interest to them. If so, this can say something about perception and motivation.
- "I plan to continue my education in CS" - Can elude to future retention, can also speak to perception and enjoyment of CS.

4. **RQ 4 - How did students perceive the overall experience of SAIL when compared to the traditional mode of instruction?** As the treatment group have been exposed to both learning with SAIL and the traditional learning approach, we asked students to indicate their agreement with the following questions regarding their overall experience using SAIL:

   - "SAIL provided an overall positive learning experience"
   - "I would choose to use SAIL for learning future topics"

   In addition to overall experience, we surveyed them specifically about the interest-based exercises and video lectures in SAIL to gauge the impact of these tools independently.

5. **RQ 5 - Does the impact of SAIL on (1)performance, (2)confidence, and (3)perception differ based on diversity factors such as gender, ethnicity, or prior exposure to CS?** To answer this, we separate all quantitative and qualitative measures by gender and prior exposure to CS. Unfortunately, with the low enrollment of students from minority ethnicities, we are unable to make meaningful comparisons regarding SAIL's impact on ethnicity.

6. **RQ 6 - Does SAIL demonstrate a potential to impact attraction and recruitment to STEM disciplines?** To answer this, we will look to the answers of our preceding research questions - student perception, motivation, performance, confidence, and diversity - to interpret how the interrelation of these results might affect attraction and recruitment in the future.

7 | Results

SAIL was used to teach two modules in an introductory level computer science course. Module 1 covered Branching and Module 2 covered Looping - both fundamental topics in introductory programming. Students in Group A used SAIL to progress through course topics in Branching and Looping as well as completing in-class exercises adapted to their interests. This section details the results collected for both quantitative and qualitative analysis to assess SAIL's overall impact. Results are divided into three parts, one quantitative analysis section for each module and a qualitative analysis at the end:

- **Quantitative Analysis - Branching** - All *branching*-related assessments for treatment and control groups
- **Quantitative Analysis - Looping** - All *looping*-related assessments for treatment and control groups
- **Qualitative Analysis** - (1) Treatment vs. control comparisons on perception and enthusiasm and (2) treatment group's assessment of their experience using SAIL.

In each section we report and begin to interpret the results from our data with a detailed discussion to follow. In addition to observing the results overall, we divide many of our measures into subgroups to determine if the data varies based on (1) gender or (2) prior exposure to CS. Unfortunately, with the low enrollment of some ethnicity groups in our pilot study, we are unable to offer insight into comparisons between minority groups. We look at the comparison of data between sections, the overall trends in the data sets, and statistical significance determined by z-tests.

7.1 Quantitative - Branching

In Module 1, students in Group A learned Branching concepts through SAIL. Three quantitative measures were used for comparing student performance between the treatment and control groups:

- Reading Quizzes
- Branching Quiz
- Exam 1

Of these three assessments, Exam 1 provided the most controlled setting for obtaining a meaningful comparison between Group A and the control groups (B, C, and D). We explore SAIL's impact in Exam 1 scores overall as well as the difference in impact between various groups. Reading Quizzes and the Branching Quiz scores are presented first, though as these were not considered strong measures to assess performance, we present only the overall comparisons.

7.1.1 Quizzes

In this section, we explore data from two types of quizzes:

- **Reading Quizzes (RQ)**:

  Short, online quizzes meant to assess student learning immediately after gaining new competencies at home. Only Groups A and B (controlled instructor) participated. The treatment group (Group A) watched specified SAIL videos on branching concepts while Group B read the specified sections of the textbook on branching concepts before taking each quiz. Students in each section completed three separate SAIL modules/readings on Branching for three separate Reading Quizzes: RQ 5, RQ 6, and RQ 7.

- **Branching Quiz**: A six-question quiz given in-class to all four sections immediately after learning all Branching concepts.

Overall

When looking at all quiz data, Groups A and B had similar performance on the reading quizzes, and all four sections had similar performance on the branching quiz. A visualization of average grades for all three reading quizzes (RQ 5 -7) taken by Groups A and B combined with the Branching quiz scores from all four sections is shown in Figure 7.1. Statistical analysis through z-tests concluded that there were no significant differences between the data sets.

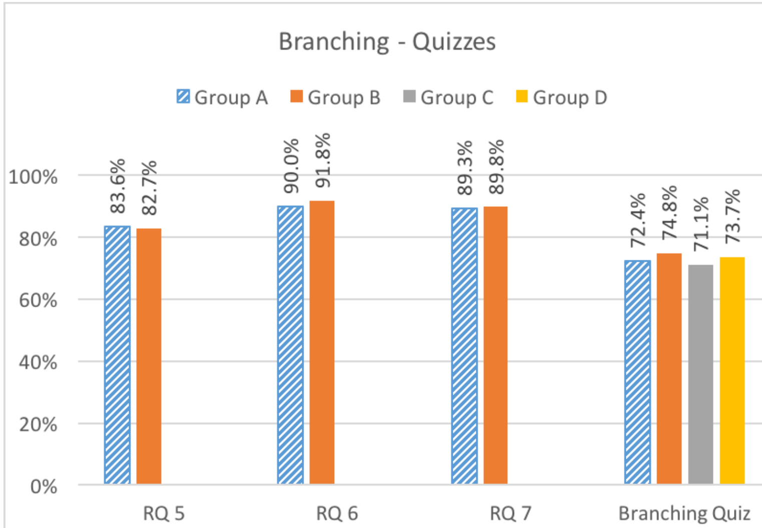

Figure 7.1: Branching - Quiz Performances

Reading quizzes were taken online, allowing for students to use textbooks, videos, and any other resources to complete the quizzes. While it was emphasized that the quizzes should be taken individually, there were no measures to guarantee individual assessment. We thus infer that reading quizzes may not accurately asses student learning of a concept, but rather that students were able to solve problems with all resources at their disposal. The similarity of the reading quiz scores for Groups A and B suggests that students were able to use SAIL as a positive resource to gain information and complete these assessments. These grades do not, however, allow us to accurately compare student performance between the groups.

The Branching quiz was given in-class, controlling the use of outside resources and assuring individual assessment. While this data may provide some insight into student performance, there were varying extrinsic factors among the courses that may have impacted results. As mentioned previously, the quiz was unannounced for Groups A and

B and was counted as "bonus", with no opportunity to harm their grade. Conversely, Groups C and D announced the quizzes ahead of time, giving students adequate time (about one week) to prepare and counting it as a quiz grade in their overall course. Due to the unevenness in preparation and extrinsic motivation for Groups A and B to perform well on the Branching and Looping quizzes, we do not feel these are strong metrics for accurate comparison of student performance among sections. It is interesting that even without prior notice or the possibility of harming their grade, students in Groups A and B performed roughly the same, on average, as Groups C and D. We hypothesize, that students in sections A and B may have demonstrated higher performance if given the same advance and extrinsic incentive to study and prepare. Based on this hypothesis, we place higher emphasis on examining student grades in Exam 1 - the first highly weighted metric, given in all four courses, that all students had the same advance and incentive (15% of their grade) to do well on.

7.1.2 Exam 1

The first exam, Exam 1, was given six-weeks into the semester. The exam was worth 15% of the student's overall grade, being the highest weighted grade earned at this point in the semester. The exam was created by Instructor A - the most seasoned instructor, and all four sections (Groups A-D) were given the same exam on the same date. The exam consisted of multiple choice, true-false, short-answer, and code-writing problems. Exam 1 follows a routine format, with only minimal deviations (changing context of questions, etc) from other exams given in past years.

The exam covered all topics taught in the first six weeks of the semester. The treatment group used SAIL for the three weeks of instruction prior to the exam to learn "Branching" concepts. In an analysis of the exam, 67 points directly correlated to Branching concepts taught with SAIL, while the other 33 points tested on foundational skills - declaring variables, using the Scanner class, etc., that are all considered prerequisites to the Branching concepts.

The Exam 1 metric is the first, heavily-weighted item that students across all four sections had the chance to study for. As students in all sections had the same extrinsic motivation and advance notice to prepare, this exam is considered the most reliable metric for comparison.

Overall

Figure 7.1 shows the average Exam 1 scores for all four sections where Group A demonstrated the highest performance (nearly 5% above the second highest class - Group B). The differences in treatment and control group scores were determined to be statistically significant when comparing Group A to each control group individually, and aggregated - resulting p-values and levels of significance are displayed in Table 7.1.

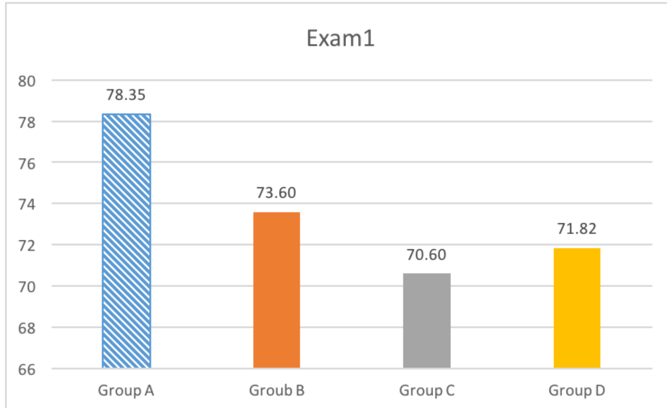

Figure 7.2: Exam 1 Performance - Overall

Group A vs. Controls

Exam 1	p-values	Siginif.
Group B	0.03027	*
Group C	0.00123	**
Group D	0.00460	**
Aggregate	0.00590	**

Significance: *p < .05 **p < .01 ***p < .001

Table 7.1: Exam 1 - Overall z-tests

The above table displays the p-values for comparing Group A with each Group (B, C, and D) independently and aggregated. Asterisks indicate the level of statistical significance determined.

As Group A significantly outperformed all control groups, including Group B with the same instructor and the same flipped classroom approach, we reason that the positive shift in exam grades is likely related to SAIL. To further explore the difference in performance between Group A and the control groups on Exam 1, we analyze Exam 1 performance by various factors below.

Gender

Figure 7.3 shows the average Exam 1 score for all four groups by gender. Both the male and female students in Group A outperformed males and females in all three control groups. The overall trend across the sections is that female students performed lower than male students in each group, with the smallest disparity seen in Group A at only 1.25%, and the largest disparity seen in Group D at nearly 13%. Among the control groups, male performance is fairly consistent (within 3 points), while female performance in the control groups varied heavily depending on section (within 11 points).

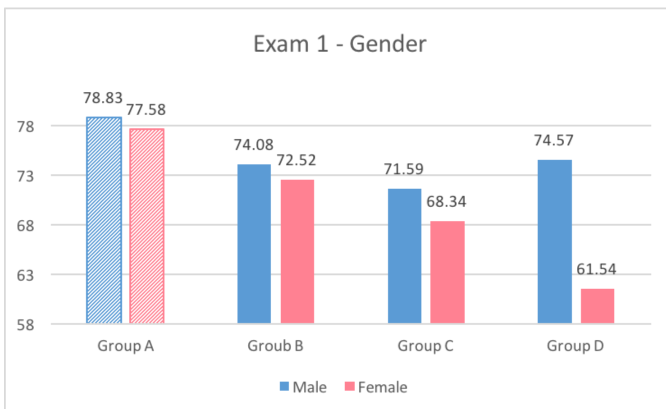

Figure 7.3: Exam 1 Performance - Gender

Z-tests were performed comparing Group A males and females to all males and females in the three control groups, both individually and aggregated. Table 7.2 shows the results of these tests. The first four rows show the results of Group A males vs. all control groups' males and females, and last four rows show Group A females vs all control groups' males and females.

Exam 1 - Gender

Group A - Males vs. Controls

	Male	Siginif.	Female	Signif.
Group B	0.11718		0.00389	**
Group C	0.00920	**	0.00374	**
Group D	0.07114		0.00002	***
Aggregate	0.04421	*	0.00216	**

Group A - Females vs. Controls

	Male	Siginif.	Female	Signif.
Group B	0.36242		0.06788	
Group C	0.08929		0.04412	*
Group D	0.31415		0.00182	**
Aggregate	0.21784		0.03463	*

Significance: *p $<$ .05 **p $<$ .01 ***p $<$ .001

Table 7.2: Exam 1 - Gender - z-tests

p-values for (1) Group A males vs. all control groups (M & F) individually and aggregate and (2) Group A females vs. all control groups (M & F) individually and aggregate.

	A - Female		B - Female		C - Female		D - Female	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Signif.
A - Male	0.66154		0.00415	**	0.00000	***	0.00000	***
B - Male	0.36242		0.59950		0.05005		0.00776	**
C - Male	0.08929		0.73485		0.22607		0.01997	*
D - Male	0.31415		0.37720		0.00639	**	0.00039	***

Significance: *p $<$ .05 **p $<$ .01 ***p $<$ .001

Table 7.3: Exam 1 - Male vs. Female (all sections) z-tests

Comparison of male (rows) vs. females (columns) performance across all sections.

Males from the treatment group performed significantly better than all females in control groups ($p < .01$ for all, $p < .001$ for Group D). The only female group not significantly outperformed by the male treatment group was the female treatment group with a p-value of .66 (shown in Table 7.3). Males from the treatment group performed significantly better than males from Groups C and D ($p < 0.01$ and $p < .05$, respectively) as well as the aggregate of all male control groups ($p < .05$).

Females from the treatment group performed significantly better than females in control groups C and D ($p < .05$ and $p < .01$, respectively), as well as the aggregate for all female control groups ($p < .05$). Females did not perform significantly better than females from Group B with the controlled instructor and learning environment, though the averages were higher by $> 5\%$. No significant differences were seen between treatment females and any male control groups, though Group A females had higher averages than all male and female control groups. Though the higher grades for Group A's females vs. all control males were not determined to be significant, Group A was the only group where female students did not perform statistically lower than at least one male control group. The comparison of all female vs. all male groups is shown in Table 7.3.

The treatment group achieved the highest similarity between male and female performance, elevating both male and female scores significantly higher than many of the controls. Group A females were the only group that were statistically comparable to all other male groups, suggesting that SAIL helped level the playing field between male and female students. In addition to helping remedy gender disparities in performance, the treatment group raised the overall bar, as both Group A male and females outperformed all control groups, regardless of gender. This suggests that SAIL helped to not only raise bar for females to perform comparable to males, but helped elevate performance overall.

Prior Exposure to CS

Exam 1 grades were divided to assess how students with prior exposure to computer science (CS) before taking the course compared to students with no prior CS experience. Analyzing the differences between these two groups can help us understand if SAIL helps impact the learning curve of students who are brand new to this subject in a different way than students who many have been introduced to the fundamentals of these topics before.

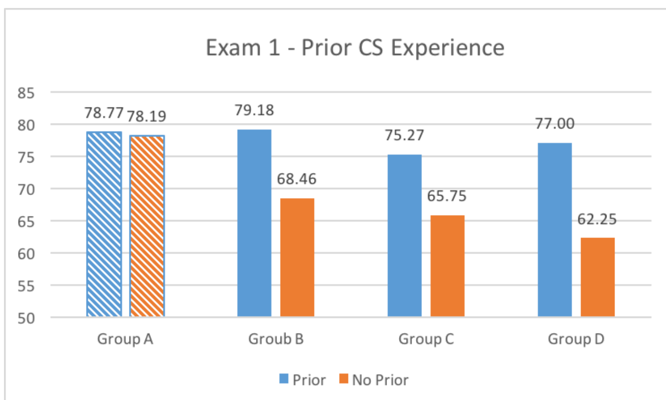

Figure 7.4: Exam 1 Performance - Prior vs. No Prior Experience

Students who had prior exposure to CS before taking this course all performed similarly on Exam 1 (75.27% - 79.18%). The treatment group falls on the high end of this range at 78.77%, though with such similar averages, no significance difference was found between students with prior experience in Group A vs. students with prior experience

Group A - Prior Experience vs. Controls

Exam 1	Prior	Siginif.	No Prior	Signif.
Group B	0.91077		0.01232	*
Group C	0.43255		0.00248	**
Group D	0.60314		0.00059	***
Aggregate	0.66921		0.00077	***

Group A - No Prior Experience vs. Controls

Exam 1	Prior	Signif.	No Prior	Signif.
Group B	0.66303		0.00013	***
Group C	0.28931		0.00000	***
Group D	0.57135		0.00000	***
Aggregate	0.65069		0.00000	***

Significance: *p < .05 **p < .01 ***p < .001

Table 7.4: Exam 1 - Prior Experience z-tests

	A - No Prior		B - No Prior		C - No Prior		D - No Prior	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Signif.
A - Prior	0.79892		0.00000	***	0.00000	***	0.00000	***
B - Prior	0.66303		0.00000	***	0.00000	***	0.00000	***
C - Prior	0.28931		0.00250	**	0.00003	***	0.00029	***
D - Prior	0.57135		0.00000	***	0.00000	***	0.00000	***

Significance: *p < .05 **p < .01 ***p < .001

Table 7.5: Exam 1 - Prior vs. No Prior Experience z-tests (all sections)

Comparison of prior experience (rows) vs. no prior experience (columns) performance across all sections.

in the controls (individual or aggregate). Students without prior exposure to CS had a much more drastic variation in Exam 1 grades (62.25% - 78.19%). This nearly 16% variation is largely due to the high performance of the treatment group (78.19%) as it differed substantially from all control groups' students with no prior experience.

Not surprisingly, a z-test indicated statistical significance between both subsets of the treatment group (with and without prior experience) vs. students without prior experience in all control groups. Table 7.5 showed the comparison of prior experience in all groups (rows) with no prior experience in all groups (columns). This data clearly shows that Group A is the only group who performed as well as students with prior experience in any section. These results suggest a decreased learning curve for students in the treatment class who had no prior experience to CS while maintaining the expected high scores of those students who did come with some CS exposure.

7.1.3 Quantitative - Branching Summary

This section presented several metrics for analyzing SAIL's impact on student performance in Module 1 - Branching. Of these assessments, Exam 1 is likely the most accurate measure for comparison due to a controlled high extrinsic motivation (15% of overall course grade), and a uniform preparation opportunity across all sections. Less controlled assessment measures including Reading Quizzes and the Branching Quiz showed statistically similar performance across all groups. These results suggest that students in both Group A and Group B had adequate resources to effectively solve problems at home on reading quizzes, and that students in all sections performed statistically similar on the Branching quiz, even with lower extrinsic motivation and no opportunity for Groups

A and B to prepare. In looking at Exam 1 scores, students utilizing SAIL performed significantly better than all control groups in individual and aggregate comparisons. In addition to increased overall performance, gender disparities between male and female students were minimized in the treatment group, as females from Group A were the only female subset who performed statistically similar to *all* male groups when no other female group performed similar to *any* male groups. It is also suggested that SAIL decreased the learning curve for students with no prior CS exposure, as the treatment group without prior experience was statistically similar to all groups with prior exposure while students without prior experience in control groups all performed significantly lower than all groups with prior exposure. These results suggest an impactful teaching approach and a decreased learning curve for the students utilizing SAIL.

7.2 Quantitative - Looping

In Module 2, students in Group A learned Looping concepts through SAIL. Three quantitative measures were used for comparing student performance between the treatment and control groups:

- Reading Quizzes
- Looping Quiz
- Exam 2

Following the structure of the former section, We first present the overall comparisons for the Reading Quizzes and the Looping Quiz scores. We then explore SAIL's impact in Exam 2 scores overall as well as the difference in impact between subgroups.

7.2.1 Quizzes

In this section, we explore data from two types of quizzes:

- **Reading Quizzes (RQ)**:

  Short, online quizzes meant to assess student learning immediately after gaining new competencies at home. Only Groups A and B (controlled instructor) participated. The treatment group (Group A) watched specified SAIL videos on looping concepts while Group B read the specified sections of the textbook on looping concepts before taking each quiz. Students in each section completed two separate SAIL modules/readings on looping for three separate Reading Quizzes: RQ 8 and RQ 9.

- **Looping Quiz**: A five-question quiz given in-class to all four sections immediately after learning all Looping concepts.

Overall

When looking at all quiz data, Groups A and B had similar performance on the reading quizzes, and all four sections had similar performance on the branching quiz. A visualization of average grades for all three reading quizzes (RQ 8 & 9) taken by Groups A and B combined with the Looping quiz scores from all four sections is shown in Figure 7.5.

Statistical analysis through z-tests concluded that there were no significant differences between the data sets.

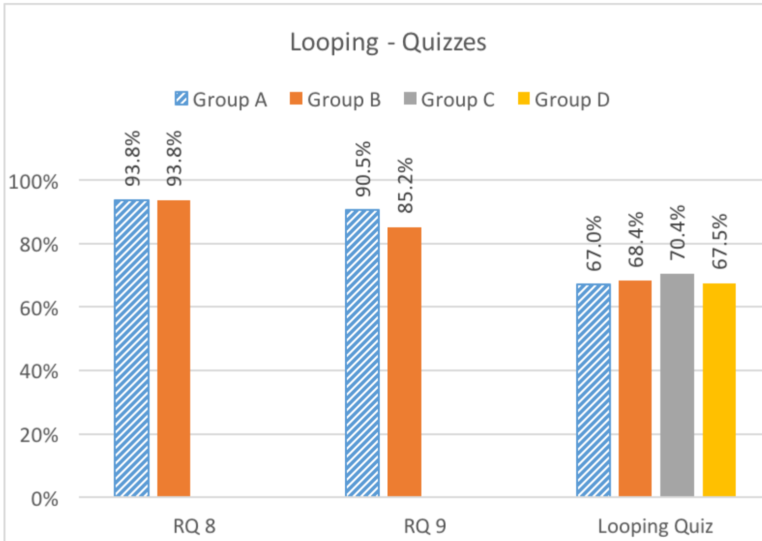

Figure 7.5: Looping - Quiz Performances

As detailed in Module 1, reading quizzes were short online quizzes that students took in an uncontrolled environment, utilizing any resource to find the correct answer to the problems. These quizzes may not be a good comparison for how SAIL impacted student performance, but allow us to conclude that students in both Groups A and B were able to attain similar levels of achievement with the resources they were given. This suggest that SAIL may be at least an equal resource in attaining desired information, but does not tell us much about impact on learning. Similar to the Branching Quiz, described in Module 1, the Looping Quiz was given to all sections, though only Groups C and D had prior notice to prepare and extrinsic motivation that their answers would count for a grade. Students in all sections indicated similar performance in the Looping Quiz, including Groups A and B who lacked the same extrinsic motivation and advance to study. While we present this data as pieces of information, we lean more heavily on Exam 2 grades to indicate meaningful results about performance between sections.

7.2.2 Exam 2

The second exam, Exam 2, was given ten weeks into the semester. The exam was worth 20% of the student's overall grade, being the highest weighted grade earned at this point in the semester. The exam was created by Instructor A - the most seasoned instructor, and all four sections (Groups A-D) were given the same exam on the same date. The exam consisted of multiple choice and true-false problems. Exam 2 follows a routine format, with only minimal deviations (changing context of questions, etc) from other exams given in past years.

The exam covered all topics taught since Exam 1 - about four weeks of material. The treatment group used SAIL for the first two weeks of instruction prior to taking Exam 2 to learn "Looping" concepts. In the subsequent two weeks leading up to the exam, the treatment class resumed their non-SAIL flipped classroom approach - reading the textbook at home and receiving the generic in-class exercises, the same as Group B. For this reason, Exam 2 grades as a whole would not be considered a good measure for how students performed with SAIL vs. the traditional approach. For best comparison, only items on the exam directly associated with Looping (without incorporating future topics) were considered in the analysis below. Out of the 33 multiple choice questions on Exam 2, 10 out of 33 questions featured only Looping concepts and were analyzed for comparison among all groups.

The Exam 2 metric is important to include in analysis as it is the second heavily-weighted item that students across all four sections had the chance to study for. Unlike Exam 1, Exam 2 provides fewer opportunities to accurately assess student performance (10 questions as opposed to 37) as well as a different timeline for when students were tested on these Looping concepts. With Exam 1, students were tested on Branching immediately after learning these concepts (while the treatment group utilized SAIL). With Exam 2, students in the treatment group halted the use of SAIL after learning Looping topics, and then went on to learn two more sections (Arrays and Classes/Objects) before taking the Exam. We further discuss the impacts that this deviation from immediate assessment may have on Exam 2 grades in the next chapter.

Overall

Class averages for Exam 2 did not differ significantly, with the lowest average at 81.37% and the highest (the SAIL group) at 83.48%. Figure 7.6 compares the overall averages for each group and Table 7.6 shows the resulting z-test output for comparing Group A exam scores against control group scores (Groups B, C, and D).

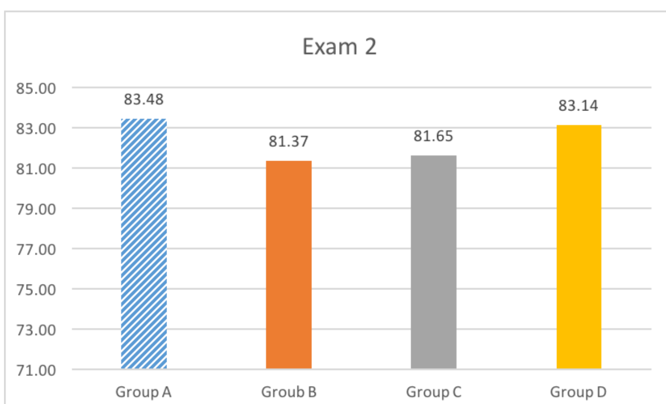

Figure 7.6: Exam 2 Overall Performances

Group A vs. Controls

		p-values	Siginif.
Exam 2	Group B	0.35093	
	Group C	0.42431	
	Group D	0.84609	
	Aggregate	0.45380	

Significance: *p $<$ .05 **p $<$ .01 ***p $<$ .001

Table 7.6: Exam 2 - Overall z-tests

The above table displays the p-values for comparing Group A with each Group (B, C, and D) independently and aggregated. Asterisks indicate the level of statistical significance determined.

We find the similarity of these grades unexpected based on the evidence presented in Exam 1. One must keep in mind that Exam 2 grades looked at far fewer questions and that this assessment was not given immediately after learning looping concepts. We discuss these variations more in the following chapter. At this point, it is important to note that students taught with SAIL were observed on Exam 2 to uphold the performance measures expected in a University level course.

Gender

Figure 7.7 shows the average Exam 2 score for all four groups, by gender. Although Exam 2 scores overall were very comparable, male vs. female performance for Exam 2 showed different trends. Male students in all sections performed similarly, whereas female performance was variable across all sections. For all three control groups, male students performed higher than the class average (+0.94% to +2.34%), while female students performed lower than the class average with a higher variation (-2.04% to -10.92%). In the treatment group, the opposite is true where the male students performed slightly under the class average (-0.62%) while the female group performed slightly above the class average (+0.96%). This is the first instance (from both Exam 1 and Exam 2 data) where a female group has outperformed *any* male group on the same assessment. As with Exam 1, Group A shows the smallest disparity (1.58%) between male and female performance, while Group D shows the largest disparity (13.26%) for Exam 2.

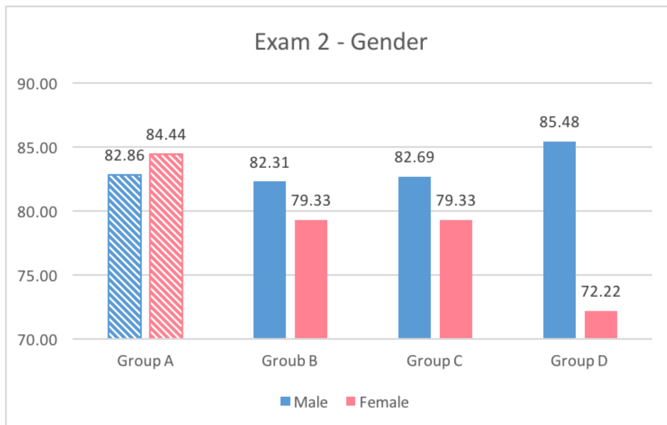

Figure 7.7: Exam 2 Performances by Gender.

Group A - Males vs. Controls

Exam 2 - Gender	Male	Siginif.	Female	Signif.
Group B	0.84500		0.25515	
Group C	0.95147		0.27287	
Group D	0.17297		0.00005	***
Aggregate	0.89025		0.14909	

Group A - Females vs. Controls

Exam 2 - Gender	Male	Siginif.	Female	Signif.
Group B	0.54213		0.18572	
Group C	0.61505		0.20226	
Group D	0.66689		0.00020	***
Aggregate	0.70729		0.11659	

Significance: *$p < .05$ **$p < .01$ ***$p < .001$

Table 7.7: Exam 2 - Gender z-tests

p-values for (1) Group A males vs. all control groups (M & F) individually and aggregate and (2) Group A females vs. all control groups (M & F) individually and aggregate.

Z-tests were performed comparing both males and females in the treatment group to all males and females in the three control groups. Table 7.7 shows the results of these tests.

Only the females from Group D (average of 72.22%) showed significant difference ($p < .001$) from either gender in Group A. When comparing female to male treatment in all sections, Group A is the only group where females did not perform worse than male students with some statistical significance. Table 7.8 shows the comparison of all male and female groups where females in Group D are significantly lower than all male groups, while females in Group B and C differ significantly from males in Group D.

This data continues the trend from Exam 1 where females consistently perform lower than males across the board, with the exception of the treatment group females in Exam 2. Just as with Exam 1, female and male performance is comparable in three of the four sections (treatment, Group B, and Group C), with the most similar scores between males and females seen in the treatment group. This strengthens the argument that SAIL has perhaps helped the playing field to be gender neutral. Group D has the largest disparity again between male and female performance - a 13.03% difference for females in Exam 1 and a -13.26% difference for females in Exam 2. As mentioned in the Exam 1 analysis, Group D did have the lowest female to male ratio at time of data collection (8 weeks into the semester), but the reason for this large disparity is unknown at this point.

	A - Female		**B - Female**		**C - Female**		**D - Female**	
	p-val	**Signif.**	**p-val**	**Signif.**	**p-val**	**Signif.**	**p-val**	**Signif.**
A - Male	0.62654		0.16314		0.16314		0.02115	*
B - Male	0.54213		0.27330		0.27330		0.04190	*
C - Male	0.61505		0.21556		0.21556		0.03428	*
D - Male	0.66689		0.00094	***	0.00094	***	0.00009	***

Significance: *p < .05 **p < .01 ***p < .001

Table 7.8: Exam 2 - Male vs. Female z-tests (all sections)

Comparison of male (rows) vs. females (columns) performance across all sections.

Prior Exposure to CS

Exam 2 grades were divided to assess how students reporting prior exposure to computer science (CS) before taking the course compared to students who reported no prior CS experience. Analyzing the differences between these two groups can help us understand if SAIL helps impact the learning curve of students who are brand new to this subject in a different way than students who many have been introduced to the fundamentals of these topics before.

Figure 7.8 shows a comparison of averages in all sections divided by whether or not the student had prior exposure to CS when taking the course. In Exam 1, we saw a significance in the way students without prior experience performed in the treatment group. The Exam 2 data presented here does not follow this trend, as all of the averages are statistically comparable. Outputs of the z-test determining no significant differences between any of the groups can be seen in Table 7.9.

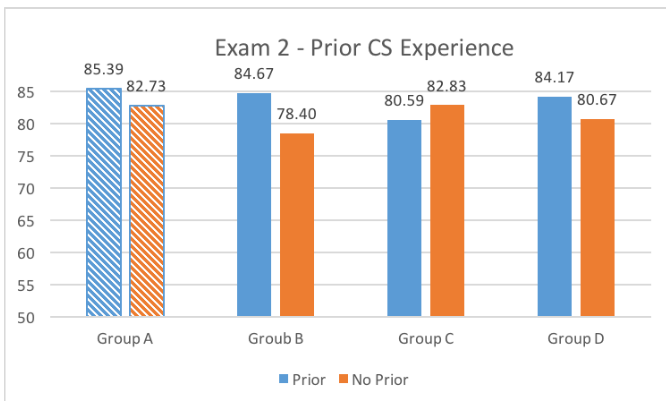

Figure 7.8: Exam 2 Performance - Prior vs No Prior Experience

Again, it is important to consider the circumstances under which Exam 2 was taken as well as observe that SAIL has had no negative effects on performance at this level.

7.2.3 Quantitative - Looping Summary

Across the board, Exam 2 scores were very comparable for overall performance as well as prior exposure vs. no prior exposure to programming. The trend of minimizing

Exam 2

Group A - Prior Experience vs. Controls

	Prior	Siginif.	No Prior	Signif.
Group B	0.86031		0.10274	
Group C	0.28867		0.53048	
Group D	0.68402		0.23670	
Aggregate	0.54424		0.32477	

Group A - No Prior Experience vs. Controls

	Prior	Signif.	No Prior	Signif.
Group B	0.4488		0.10726	
Group C	0.4509		0.96921	
Group D	0.44349		0.41027	
Aggregate	0.928		0.62310	

Significance: *p < .05 **p < .01 ***p < .001

Table 7.9: Exam 2 - Prior Experience z-tests

	A - No Prior		B - No Prior		C - No Prior		D - No Prior	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Signif.
A - Prior	0.25111		0.00020		0.19203		0.16952	
B - Prior	0.44880		0.00259		0.39606		0.29225	
C - Prior	0.45090		0.34244		0.35173		0.98513	
D - Prior	0.44349		0.00016		0.39943		0.20902	

Significance: *p < .05 **p < .01 ***p < .001

Table 7.10: Exam 2 - Prior vs No Prior Experience z-tests (all sections)

gender disparities was continued from Exam 1, with female performance in the treatment group surpassing any group of male performance for the first time. Though the gender comparisons for Exam 2 were not as statistically significant, the consistency of trends between both exams is important. While the difference between Exam 1 and Exam 2 data outcomes raise many questions, it is important to note that Exam 2 did not show any harmful effects to the treatment group.

7.3 Qualitative

As with most educational studies, we believe that quantitative data only tells part of the story. As we have seen, quantitative data often leads to many questions as to what may or may not have caused variations in the data. To fill in some of these gaps, we assess qualitatively how students perceived SAIL's impact in many areas. This analysis is divided into two parts:

- **Perception / Enthusiasm about CS** - This is where we analyze responses from all section regarding their overall perceptions, interests and feelings about the course.
- **SAIL Perception** - This is where we analyze the questionnaire given to students who interacted with SAIL to analyze their perception of SAIL's impact as a whole and looking at individual aspects.

7.3.1 Perception / Enthusiasm about CS

Students in all four sections were asked to rate their agreement on a 1-7 likert scale (1=Strongly Disagree to 7=Strongly Agree) with several statements regarding their enjoyment and perception of their Computer Science course and Computer Science / programming as a whole.

Overall

Figure 7.9 shows the average responses for each question by section. Statistical significance between the treatment group (Group A) and all three control groups are indicated by the appropriate asterisks (*p < .05, **p < .01, ***p < .001) for each question - displayed in Table 7.11.

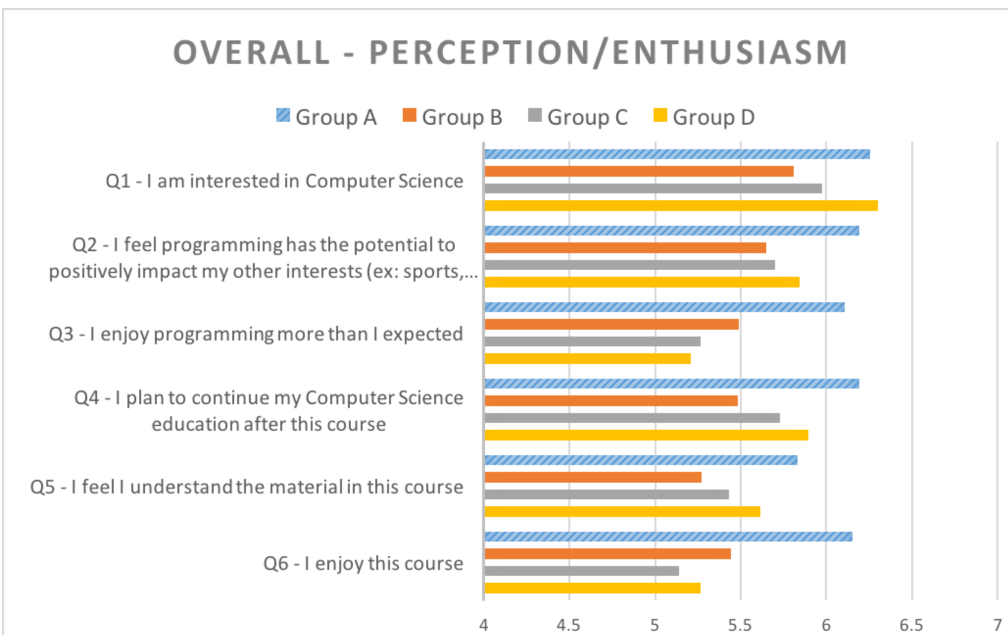

Figure 7.9: Perception / Enthusiasm - Overall

OVERALL

	Group B		Group C		Group D		Aggregate	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Signif.
Q1	0.02817	*	0.14412		0.7922800		0.1548000	
Q2	0.00534	**	0.01252	*	0.07820		0.01441	*
Q3	0.00572	**	0.00019	***	0.00002	***	0.00054	***
Q4	0.01152	*	0.05787		0.20356		0.04263	*
Q5	0.02291	*	0.10572		0.31926		0.08040	
Q6	0.00140	**	0.00003	***	0.00016	***	0.00020	***

*p < .05, **p < .01, ***p < .001

Table 7.11: Perception/Enthusiasm - Overall z-tests

The above table shows the output for overall perception comparison of Group A against the three control groups (Groups B, C, and D) individually and aggreagated. First, Group A is compared to the controls overall, second, Group A females are compared to all control females, and last Group A males are compared to all control males.

When compared to the control groups individually, Group A responses are significantly higher than at least one control group for each question. Group A is significantly higher than all three control groups for Questions 3 - "I enjoy programming more than I expected" and Question 6 -"I enjoy this course". Group A also reported significantly higher responses than 2 out of 3 control groups for Question 2 - "I feel programming has the potential to positively impact my other interests (ex: sports, art, ...). When isolating the comparison between Groups A and B, controlling for instructor and classroom design - Group A responses are significantly higher for every question.

In looking at aggregated results where Group A was tested against the control groups' population as a whole, Group A had significantly higher responses for Questions 2 (feeling CS is applicable to their other interests), 3 (enjoying programming more than expected), 4 (intending to continue CS education), and 6 (overall enjoyment of the course).

For items not significantly different from the aggregate of the control groups, (Q1 and Q5), Group A still demonstrated significantly higher agreement when compared to at least one other control group (Group B) and indicated a higher than average response when compared to all groups.

These results indicate that SAIL had an overall positive impact on enjoyment, understanding how CS can impact other fields, and continued pursuit of CS education. We divide these results into our different subgroups below.

Gender

The Perception/Enthusiasm analysis is divided into male and female responses to assess any differences related to gender. Perception/Enthusiasm ratings were high for both male and female students in the treatment group when compared to other groups. Figures 7.10 and 7.11 below show the comparisons.

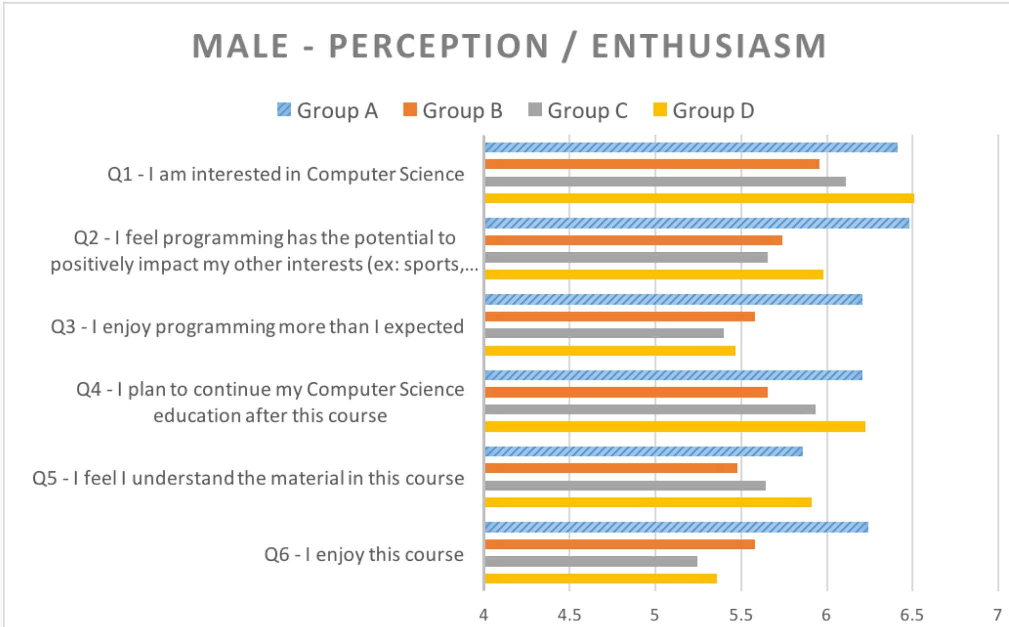

Figure 7.10: Perception / Enthusiasm - Male

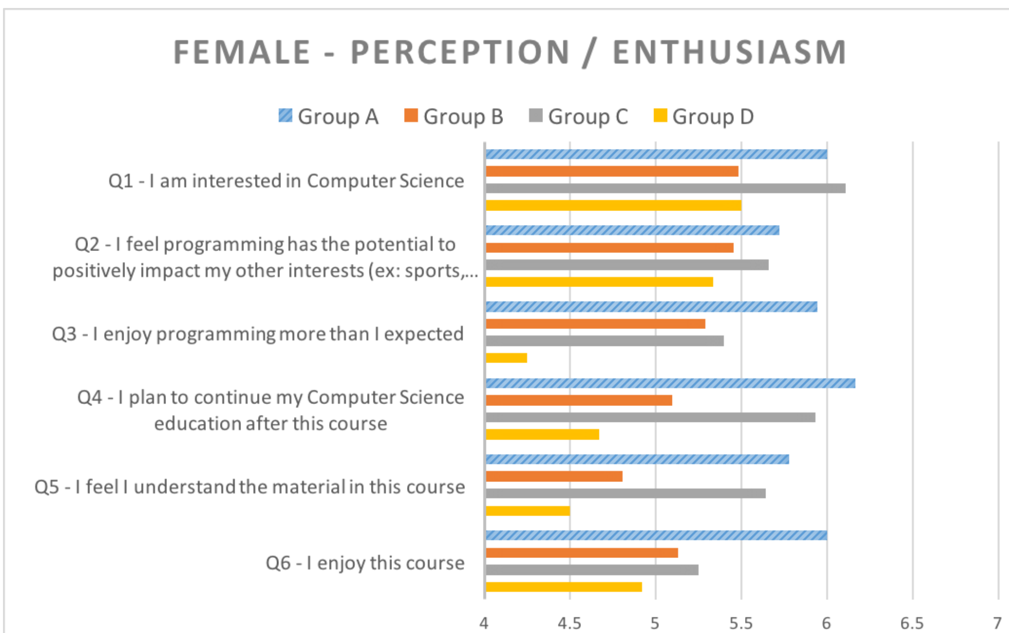

Figure 7.11: Perception / Enthusiasm - Female

FEMALES

	Group B		Group C		Group D		Aggregate	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Siginif.
Q1	0.07617		0.20801		0.27186		0.15687	
Q2	0.38170		0.78232		0.35268		0.61134	
Q3	0.05046		0.01844	*	0.000003	***	0.01177	*
Q4	0.01919	*	0.01425	*	0.00974	**	0.01494	*
Q5	0.01543	*	0.03838	*	0.00873	**	0.01958	*
Q6	0.01108	*	0.00194	**	0.02929	*	0.00684	**

*p < .05, **p < .01, ***p < .001

MALES

	Group B		Group C		Group D		Aggregate	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Siginif.
Q1	0.08870		0.23861		0.42810		0.27386	
Q2	0.00318	**	0.00271	**	0.02524	*	0.00490	**
Q3	0.03263	*	0.00192	**	0.00229	**	0.00738	**
Q4	0.11971		0.44934		0.94049		0.34049	
Q5	0.22462		0.54506		0.81752		0.49532	
Q6	0.02098	*	0.00192	**	0.00112	**	0.00436	**

*p < .05, **p < .01, ***p < .001

Table 7.12: Perception/Enthusiasm - Gender z-tests

The above table shows the output for male and female perception comparison of Group A against the three control groups (Groups B, C, and D) individually and aggreagated. First, Group A females are compared to all control females, and last Group A males are compared to all control males.

While the male group indicated average or above average responses for most questions, the most significant differences in enthusiasm/perception can be seen for female students. Similar to overall perception, both male and female students using SAIL indicated a high statistical significance from almost all control groups. Their comparisons, however differ drastically when compared to other controls for Q2 and Q4. Male students differed significantly from all other male controls in response to Q2 - Feeling that CS has the potential to impact their other interests, while female comparison to this question was not significant. These results indicate that SAIL may be positively impacting the perception of male students to demonstrate the high applicabiliyy of CS in many areas.

Female students in the treatment group reported significantly higher responses to Q4 - planning to continue their CS education after this course. This indicates that SAIL may be helping motivate female students to continue with their CS education. This is important to create a diverse culture of computer scientists. Male student responses in the treatment group were also higher than many controls, though not by any significance. In comparison, male and female students both reported similar ratings for continuing their CS education, demonstrating the potential of a larger, more diverse population recruited via SAIL.

Prior Exposure

While all responses for students with prior experience between all sections are fairly similar, with little statistical significance, students reporting no prior CS experience had significantly different responses for every question across almost all control groups and significance against the aggregates for all questions. Figures 7.12 and 7.13 show the comparisons of average agreement across all sections, while Table 7.13 displays their significance.

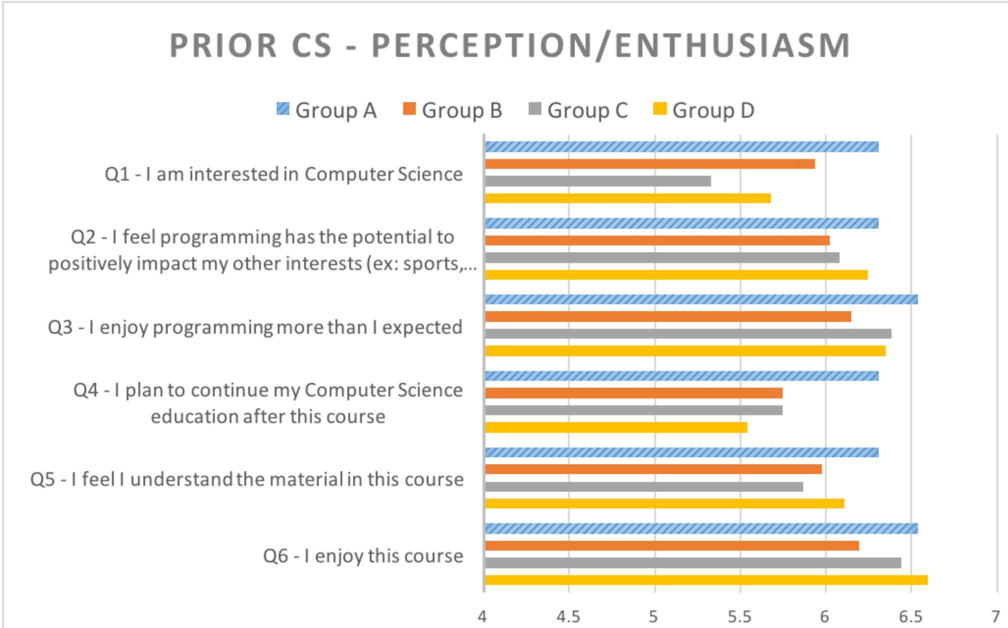

Figure 7.12: Perception / Enthusiasm - Prior Experience

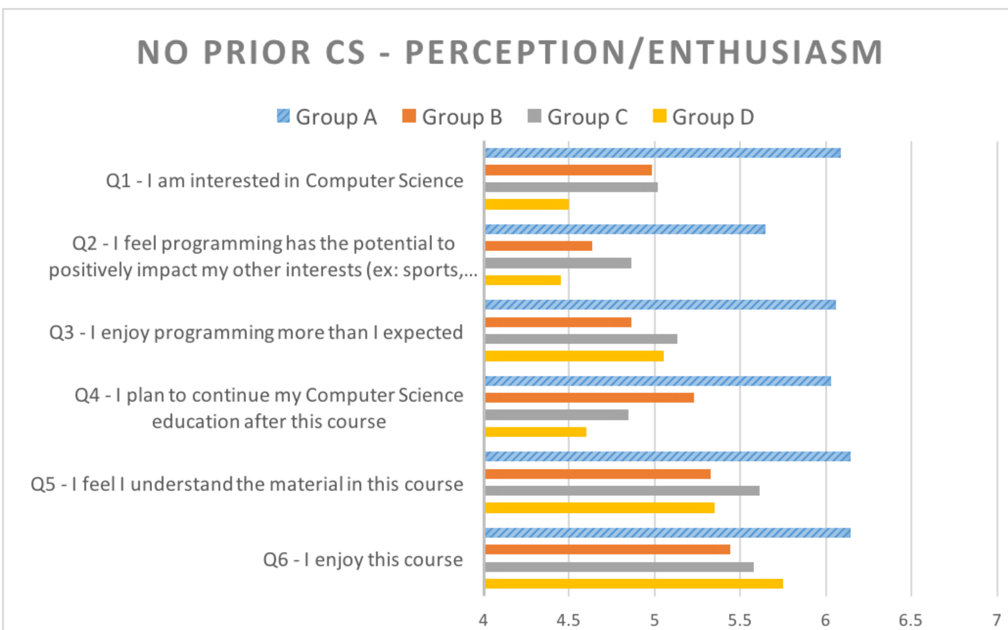

Figure 7.13: Perception / Enthusiasm - No Prior Experience

PRIOR EXPERIENCE

	Group B		Group C		Group D		Aggregate	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Siginif.
Q1	0.36919		0.72310		0.73547		0.63373	
Q2	0.36535		0.24027		0.48116		0.33195	
Q3	0.14431		0.09002		0.03099	*	0.08235	
Q4	0.38446		0.63662		0.56869		0.51074	
Q5	0.37747		0.54772		0.80769		0.53859	
Q6	0.29990		0.04016	*	0.11223		0.10900	

*p < .05, **p < .01, ***p < .001

NO PRIOR EXPERIENCE

	Group B		Group C		Group D		Aggregate	
	p-val	Signif.	p-val	Signif.	p-val	Signif.	p-val	Siginif.
Q1	0.00286	**	0.00418	**	0.14417		0.00824	**
Q2	0.00041	***	0.00504	**	0.00811	**	0.00225	**
Q3	0.00579	**	0.00002	***	0.00000	***	0.00016	***
Q4	0.00061	***	0.00115	**	0.00233	**	0.00095	***
Q5	0.00049	***	0.00127	**	0.00001	***	0.00042	***
Q6	0.00006	***	0.00002	***	0.00000	***	0.00001	***

*p < .05, **p < .01, ***p < .001

Table 7.13: Perception/Enthusiasm - Prior Experience z-tests

A quick glance at Table 7.13 shows you that there was a drastic difference between responses of students with and without prior experience. The students without prior experience who utilized SAIL are the only group of students without prior experience whose perception, in all aspects, is comparable to students who came with prior exposure. In comparison with the "overall" perception discussed earlier, students with no formal experience had significantly hiher responses for Q1,Q2,Q4, and Q5, demonsotrating an increased interest in CS, feeling that CS can impact their other interests, plans to continue their CS education, and feeling they understand the course material. Additionally, students with no prior experience using SAIL reported a higher average for plans to continue their CS education than students in control groups *with* experience. Though these numbers did not differ significantly, this trend shows us that SAIL has effectively increased intent to pursue CS across the board when compared to to control groups. Students with no prior experience using SAIL reported feelings of understanding the course material similar to students in all sections who had prior exposure, unlike students in the control groups without experience. This suggests a decreased learning curve through using SAIL.

And it is no surprise that students with prior experience in CS may report high levels of interest, enjoyment, and feeling comfortable with the course material - but students without prior experience using SAIL are now able to attain similar levels of enjoyment interest, and feeling they understand the course material as students who have some background.

7.3.2 SAIL Perception

To help understand how different aspects of SAIL may play a role in the impacts we have observed, we asked students who used SAIL a series of questions to gain some insight into their perceptions of the learning experience, Students were asked a series of questions to gauge their overall experience using SAIL and which parts of SAIL they perceived as useful. Response to SAIL was overwhelming positive with details outlined below.

As this section was only completed by the treatment group, we compare feelings within the treatment group by grouping responses into three categories to aid in our discussion:

- **Negative** - responses 1-strongly disagree, 2-disagree, and 3-somewhat disagree. Warmer colors (red, orange yellow) indicate negative responses in our figures.
- **Neutral** - response 4 - neither agree nor disagree. Gray was chosen to indicate neutrality in our figures.
- **Positive** - responses 5-somewhat agree, 6-agree, and 7-strongly agree. Blue shades were chosen to indicate positive responses in our figures.

Overall Experience

To capture student's perception of their overall experience with SAIL, a series of statements were provided where students were asked to rate how strongly they agreed or disagreed with each statement on a likert scale where 1-Strongly Disagree and 7-Strongly agree. Figure 7.14 provides a breakdown of the percentage of students (out of 47 total) who selected each response.

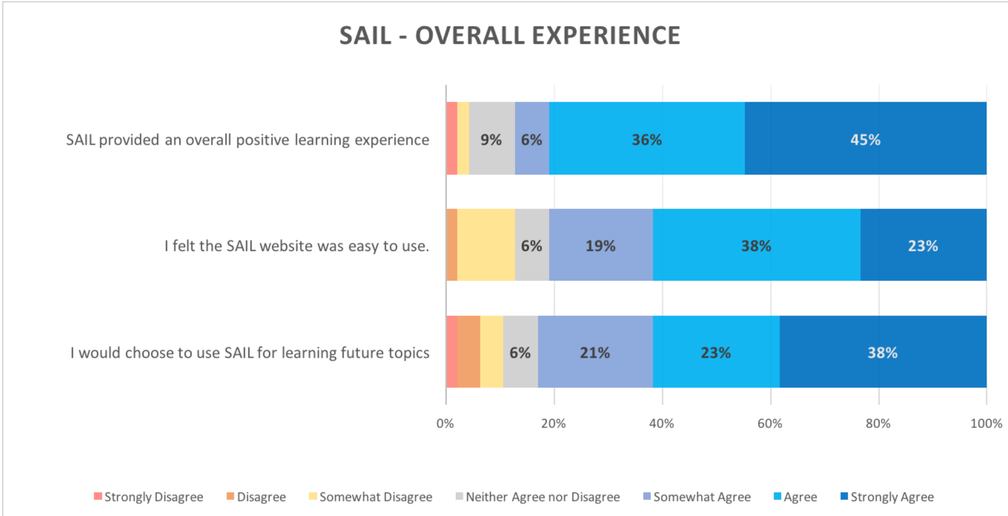

Figure 7.14: SAIL Perception - Overall

Student responses about their overall experience using SAIL were overwhelmingly positive, with 87% of students positively responding to "SAIL provided an overall positive learning experience", and 82% positive that "[They] would choose to use SAIL for learning future topics". These numbers lead us to conclude that SAIL was perceived as a beneficial experience and that students seemed receptive to the intervention.

In analyzing other responses about the use of SAIL, 80% of responses were positive that "[They] felt the SAIL website was easy to use". This indicates that students were receptive to the acceptance of SAIL and that its usability was acceptable for the purposes of this study.

Interest-based Exercises

We asked students several questions specifically about the interest-based exercises to gain an understanding of how students perceived their impact.

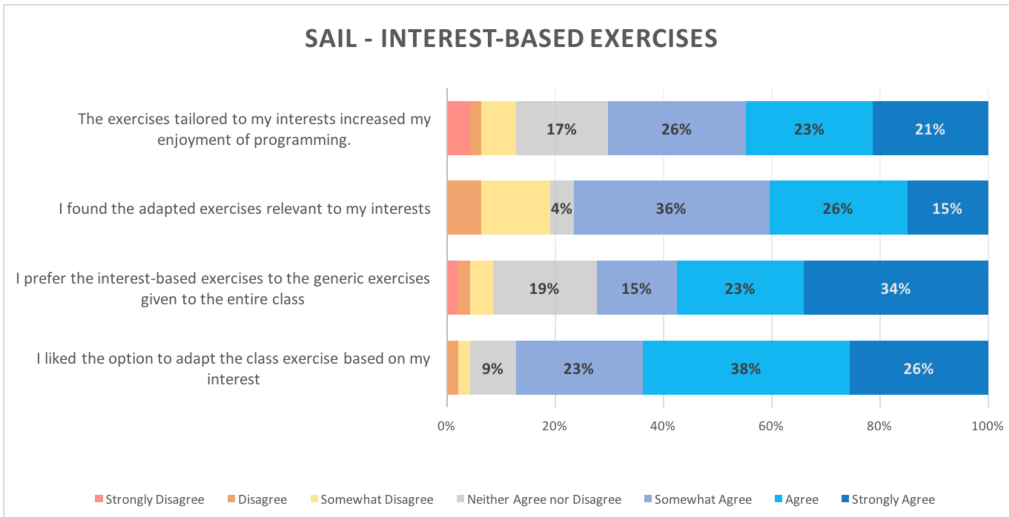

Figure 7.15: SAIL Perception - Interest-based Exercises

70% reported a positive indication that the interest based exercises helped increase their enjoyment of programming. As we saw the trend in 7.3.1 indicating SAIL's influence on an increased level of course/programming enjoyment, this data offers a clue as to what role the interest-based learning exercises played. With a 70% positive response, 17% neutral, and only a 13% negative response, it seems most students felt these interest-based exercises aided in increasing enjoyment.

77% of students positively assessed that the exercises were relevant to their interests. A similar amount, 72%, identified positive feelings of preferring these interest-based exercises to the generic exercises given before. Additionally, 87% of students reported that they liked the option to adapt the class exercise based on their interests.

SAIL Lecture Videos

Students were asked questions about the lecture videos in particular to assess what role they may have played in SAIL's overall impact, Figure 7.16 shows their responses.

81% of students reported referencing the SAIL videos when working on programming assignments, and 75% of students indicated they preferred watching the lecture videos in SAIL to the traditional learning approach. Additionally, 89% of students felt the videos in SAIL helped them learn. These numbers indicate that students perceived the videos as a positive resource in their learning experience.

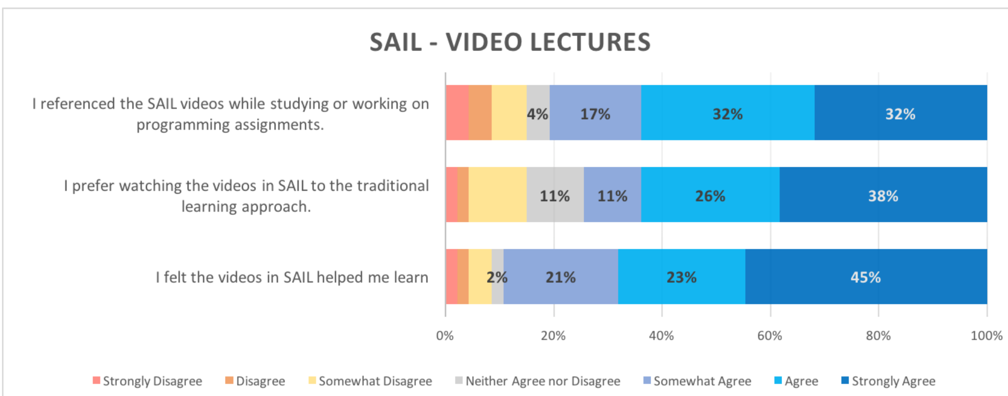

Figure 7.16: SAIL Perception - Lecture Videos

7.3.3 Qualitative Summary

The quantitative data showed that SAIL had overall positive implications on student perception of motivation, interest, enjoyment, confidence, and overall learning experience. Most significantly overall was the difference between the treatment and control groups in enjoyment of the course. Large disparities between male and female responses were seen for all questions, females in the treatment group indicating responses similar to their male peers. Isolating the female responses from all sections, SAIL demonstrated an increase in female students intending to continue their CS education. Even larger disparities were seen between students who had prior CS experience vs. students reporting no prior CS experience. While responses for students with prior experience were roughly comparable, though, the treatment group lead these averages in many areas, difference between students with no prior experience were seen to be significant for all questions. These results demonstrate that for students with no prior CS exposure, SAIL may help with decreasing the learning curve, increasing confidence and lead to higher enjoyment and understanding of SAIL's impact on other fields.

Additionally, students in the treatment group had an overwhelmingly positive responses regarding their experience using SAIL. Among these results, the most impactful are the positive feelings that SAIL provided an overall positive learning experience and that students would choose to use SAIL for learning future topics. These results demonstrate that students perceived that the benefits to their learning was directly correlated with SAIL and that they preferred its use over the traditional instruction approach. In analyzing student perception of the different parts of SAIL,the most significantly positive responses were that the majority of students reported that the lecture videos helped them learn, and liking the option to adapt the class exercises based on their interest.

8 Discussion

8.1 Quantitative Results Discussion

The results from this experiment illuminated many meaningful insights into SAIL's potential impact on performance measures. Overall, performance measures were extremely high for the first exam, possibly our strongest controlled measurement, while overall performance in all other assessments (Exam 2, all quizzes) indicated no preference between the treatment or control groups.

A limitation of our study was the lack of controlled testing across all sections. Ideally, the branching quiz and looping quiz would have been given to students in all four sections with the same advance notice and same extrinsic motivation to prepare. Unfortunately, with educational studies, especially with large studies as this one, we cannot control all variables relating to the instructor and their decisions in the course. Even with uneven preparation and uncontrolled extrinsic factors, we found it interesting that students from Group A and B performed just as well as their peers with time to prepare in Groups C and D. While we cannot draw any firm conclusions that Group A and B would have performed better if given the same opportunities, this theory is in line with elevated scores from Group A observed for Exam 1.

Additionally, we saw clear trends that SAIL helped minimize the gender disparities by elevating female performance to significantly higher levels than their female peers, and to comparable levels of their male peers. Though this effect was more evident for female students, male students in Exam 1 also demonstrated a significant elevation in performance than their male peers. Students using SAIL with no prior CS experience also saw significantly higher grades for Exam 1, suggesting a decreased learning curve for these students as their grades directly correlated with experienced students. We suspect that these increases in performance outcomes can be explained by the arousal of situational interest during learning. Past studies have claimed that arousing situational interest during learning lead to increased motivation and therefore higher learning outcomes. One goal of SAIL was to use the interest-based exercises to stimulate situational interest to increase interest in CS as well as increase performance measures. We believe these results confirm this effect.

While Exam 2 measures were extremely similar for overall comparisons and when looking at the prior/no prior experience subgroups, we noted many differences in the way the exam was administered that may have impacted an accurate assessment. While we feel Exam 1 is our most accurate assessment of SAIL's impact, all other measures still provide us with insightful information about how SAIL may have impacted various aspects of performance. It is important to note, that in no scenario was the SAIL observed to negatively impact performance measures. In fact, the treatment group was at least equal to, if not superior than all controls on every measure assessed. This demonstration of performance utilizing SAIL translates to a high possibility for improved performance if utilized in K-12 STEM educational settings in which the same excellence in education is not often upheld.

As we discussed heavily in Chapter 2 and Chapter 5, there is much need for a competent and compelling curriculum in CS and other STEM fields that can uphold a standard of education while helping to attract and interest more students in STEM careers. As our study showed, at a minimum, that SAIL can help students achieve similar if not better performance measures as a University level introductory course with proven instructors, such a system could be incredibly impactful when applied to a K-12 or community college setting. With future development of lecture videos and adaptive content embedded within the system, SAIL could be used to instruct more educators and students alike in an interesting, impactful, and standardized way. As many other studies have shown that an interest-based learning approach or an adaptive learning system have increased student performance, the careful design to target STEM issues and embed usability and a community knowledge sharing potential sets SAIL apart as a system that could not only be impactful, but could also be accepted and adopted by the community.

8.2 Qualitative Results Discussion

8.2.1 Perception / Enthusiasm about CS

In this section, we look at each question regarding perception and enthusiasm about CS. We first state what our initial hypothesis was for the outcome, and then discuss how this compared to our results. This section can thus serve to highlight the important results seen for each question, what we can learn from this, and what new questions arise from our data. We revisit the literature, when needed, to attempt explanations for results that were different from our expectations. In addition to discussing the results for each question in isolation, we discuss how the results relate between these questions and look for any correlations among the data - for example, did higher confidence lead to higher interest?.

To best present the flow of our discussion, the questions are not presented in the order they were asked or analyzed - we display the question numbers for easy reflection on earlier graphs and charts.

Q2 - *"I feel programming has the potential to positively impact my other interests (ex: sports, art, ...)"*

Based on the existing literature, we understand that CS is often misunderstood to be solely about technology and that this misperception has influenced who pursues the field. Therefore, with SAIL, we attempted to harness the interconnectivity of CS with the world around us to help students understand and become motivated to learn CS. Our hypothesis was that utilizing the interest-based examples in SAIL would lead to a positive shift in student feelings that CS could impact their other areas of interests. We expected this positive shift to occur overall, but be particularly evident for students 1) with no CS exposure and 2) for female students.

Our overall results closely imitated our exceptions. Students using SAIL indicated significantly higher agreement than two out of three control groups and the aggregate of all controls. While not determined statistically significant from one control group, the p-value was still very low (p = .078).

In comparing these responses by gender, we found that male students using SAIL were statistically significant from all three control groups and their aggregate while female students using SAIL showed no significance from their female peers. We expected females to be most impacted by seeing the applicability of CS with other fields via the interest-based examples in SAIL, but this was not the results that were shown. These unexpected results cause us to reflect on what these interest-based exercises were and why they had an apparent difference for male feelings of impacting their interests and not female.

Recall that we presented students with three interest categories - Science, Entertainment, and Sports. With the enormous amount of instructor effort required to create adaptive content, we were limited in how many categories we could implement in this pilot. We chose these categories to be broad and somewhat distinct in hopes that almost everyone would have some amount of interest in one of the three subjects. Looking at the apparent differences between male and female feelings, one possible explanation is that these categories did not appeal to female interest as much as they did to males. To add to this argument, we also included the word "sports" in the question text - "*... impact my other interests (sports, art, ...)"*. Could male students have been more interested in sports, and thus, having a sports-themed category for interest problems and having the word "sports" in the text of the question triggered higher responses from those males? Perhaps in a future work, one could survey a group of students about what they are interested in, and then create the adaptive problems accordingly. These questions are subject to further investigation.

Though we cannot explain this completely, we can look ahead to another data source - the SAIL Perception survey where one of the questions posed was "I found the adapted exercises relevant to my interests". For this question, male and female responses showed no statistical difference ($p = 0.49$) in response. This data does not support our earlier explanations, though with subjective measures such as surveys, there may be other measures at play. One further possible explanation comes from the literature in understanding that there are two distinct kinds of interest "situational" vs. "personal".

As we described in our literature review - situational interest is described as a fleeting interest that can be aroused from educational stimuli while personal interest (or individual interest) is described as a permanent interest that sustains over time. With situational interest, students may think that the problem is interesting in the moment, but they do not continue thinking about it or become curious about the subject once the problem has concluded. [35]. In contrast, personal interest is a lasting interest such as someone who loves playing music or sports and their interest sustains over time. In our survey, we asked students about their "interests" in different ways, without specifying which kind of interest we are discussing. Though expecting students to indicate if they are personally or situationally interested does not seem plausible, perhaps we can use this knowledge distinguishing personal vs. situational interests to help explain our data. As both male and female students made similar claims that the adapted problems were relevant to their interests, but differed drastically in feeling that CS has the potential to positively impact their other interests (sports, art, ...), one possible explanation is that the asking if the adapted problems were related to their interests gauged situational interest, while asking if they felt CS was applicable to their other interests indicated personal interest. Fortunately, even if this is true, situational interest aroused during instruction has been shown to lead to higher learning achievements and is almost always the target of educational interventions - as opposed to the more difficult task of targeting a unique personal interest for each student [79].

It should be noted, however, that even though female responses (for feeling programming can impact their other interests) were not significantly impacted, students using SAIL still had the highest response average among all female groups. The question then becomes - even though females did not respond to this question with any significant difference, can we determine if their perception and enthusiasm about CS was impacted in other ways? We keep this in mind and bring it back up, when relevant, in later questions.

Isolating students who claimed to have no prior CS experience showed a significant difference between students who used SAIL and students in all three control groups ($p<.01$ for all and $p<.001$ for Group B), while there was little difference for students with prior experience. This data aligns with our expectations, as SAIL helped students with no prior CS experience understand the applicability of CS in their other areas of interest. Interestingly, even though students with no prior CS experience using SAIL had significantly increased feelings that CS could impact their other interests, these responses were still much lower than the responses of all groups with prior CS experience, with only the treatment group reaching comparable levels.

To summarize, in understanding if SAIL helped students understand how applicable CS is to their other interests, our expectations were upheld for 1) overall agreement and 2) students with no prior experience in CS. Unexpected findings were that female students using SAIL did not show significant differences in these feelings while male students did. However, when looking at the SAIL Perception survey - both male and females indicated similar levels of agreement that the exercises they were given were relevant to their interests. This leads us to believe that the interest stimulated through our adapted exercises is probably situational (or fleeting) interest. As interventions that arouse situational interest have been shown to increase learning performances, we move to assess if we can corroborate these results in other areas. Additional suggestions would be to extend the interest categories based on a questionnaire of student interest in the future.

Q1 - *"I am interested in Computer Science"*

We hypothesized that SAIL may positively impact interest in Computer Science for often underrepresented groups - in our study these include females and students without prior CS experience. This theory is again guided by the literature showing that misperceptions and stereotypes can keep students from pursuing the field and the studies that have successfully recruited and retained more female students (often students without prior experience) to the field through breadth-first approaches. We believed that SAIL could cause similar occurrences as the breadth-first course restructuring by showing the integration of CS with other fields (Science, Entertainment, and Sports). In making this argument, it makes sense that we feel those students who were positively impacted for Q2 - feeling CS has the potential to impact their other interests, would also demonstrate a positive impact in interest.

Our data showed an overall increase in interest when compared to Group B ($p < .05$). Group B is the course with the controlled instructor and the same flipped-classroom design. This data is somewhat similar to the results for Q2, where SAIL students differed statistically from Groups B and C overall. We are unsure what caused this difference between only Groups A and B but do not want to hastily draw conclusions from this data alone.

Our expectations were upheld for students with no prior CS experience, indicating significantly higher feelings of interest than two out of three control groups (Group B and C) as well as the control aggregate ($p<.01$ for all). Students with prior CS exposure showed no statistical differences in reported interest, though on average, their interest level was reported higher than all controls. When comparing students with and without experience, students with no prior experience in control groups reported much lower interest, on average, than any of the students with experience. Students without experience who used SAIL indicated a higher average feeling for interest than all control groups with experience.

Our expectations were not upheld, however, as there were no statistical indications that SAIL helped increase reported interest for females. Additionally, no statistical differences were found for male students or students with prior CS background - though students with prior CS experience who used SAIL indicated the highest feelings of interest, on average. It should be noted that in none of these groups was SAIL observed to negatively impact interest. Why did interest seem affected for students without prior CS experience but not females? Could these results indicate a correlation with the data from Q2 where students without prior experience indicated understanding how CS could impact their other interests while females did not? Though there do seem to be some trends, it is uncertain how far these correlations may go. It is possible that, being a

University level class, with all sections taught be a well-liked instructor, that not much difference in interest was detected because all instructors did a good job of peaking student interest. While the literature is clear that stereotypes and misperceptions hinder students from pursuing CS, it is less clear about how students feel once they take an initial course. To return to our overall objective, we reflect on why we surveyed students about their interest in the first place. The overarching goal was to determine if interest-based learning (ie: increased interest while learning) could impact a student's motivation to pursue CS as a field of study, therefore helping with attraction and retention. To further explore the bigger picture, we discuss student's intent to continue their CS education below.

Q4 - *"I plan to continue my Computer Science education after this course"*

We hypothesized that using SAIL would increase students' intentions to pursue further CS education - particularly for female students and students without prior CS experience. This hypothesis was based on previous and later discussed questions where we felt students would enjoy programming more, become more interested in the topic, and want to continue learning. With literature showing that breadth-first applications have helped attract and retain female students (often without CS experience) we hoped SAIL would mimic these results and we could infer something about SAIL's ability to help with attraction and recruitment.

The results corroborated our hypotheses with much significance. Female students who used SAIL had significantly increased intention to continue studying CS than females in all control groups ($p<.05$ for all). When comparing the male and female responses, we see that SAIL helped bring female intent to continue their CS education to a similar level as their male peers. This increase in female intent to continue CS education seems disconnected from the female responses for Q2 and Q1 – SAIL females did not indicate a significant feeling that CS could impact their other interests and only showed a slightly higher interest in CS. However, even without indicating a higher interest in CS or feeling that CS can impact their other interests, there is a clear increase in students intending to continue their CS education. The expected correlation between these questions would be: Q2 $\rightarrow$ Q1 $\rightarrow$ Q4 (feeling programming could impacting interests $\rightarrow$ interest in programming $\rightarrow$ intent to continue learning). This evidence does not support this process, indicating there must be other factors influencing female students' increased intent to continue CS. In contrast, there were no significant differences in males groups' intent to continue their CS education, although there was significant increase in SAIL males feeling CS can impact their other interests. In this case, males only showed a significant increase in Q2, not Q1 or Q4, strengthening the argument that understanding or a lack of understanding of how programming can impact their other interests, at least for males, does not lead to increased interest or intent to continue CS education.

When isolating students based on whether or not they had prior CS experience, we also saw the expected results in our data: SAIL showed a significant impact in students without prior CS experience's intention to continue their CS education, while students with prior CS experience showed no statistical significance between treatment and control groups. The trend for students with no prior experience does correlate to our previous questions, as this group showed significantly higher understanding that CS could impact their other interests, significantly higher interest in CS, and significantly higher intent to continue their CS education. While we cannot confidently say, based on our studies divided by gender, that one of these feelings directly impacts the other, there is a pretty strong sense of consistency throughout the data that students without prior CS experience had an increase in interest and perception of what CS is, influencing their decision to pursue more CS studies.

Looking at the data for this question as a whole, we see a significant increase in females and students without prior CS experience's intent to continue their CS education, though it is less clear what has caused this. Increasing intentions for female students and students with no prior CS experience speaks monumentally to the potential for attracting more students, from more diverse backgrounds. It is possible that SAIL helped correct misperceptions about what CS entails – as students without prior experience seemed to indicate they understood how much it could impact their interests. And for females, though it is unclear at this point why, it is evident that SAIL increased their intention to pursue future CS studies. As discussed in our literature review, CS is in need of a larger and more diverse set of students, and the data seems clear that SAIL can assist in providing this.

Q5 - *"I feel I understand the material in this course"*

The literature well-documents that confidence levels and the image about who is capable of pursuing CS largely impacts students' entry into the field. One of our research goals was to assess how SAIL impacted confidence levels in how students felt they understood materials. Our hypothesis was that SAIL would aid in increasing female confidence. Additionally, as we expect SAIL to lead to higher learning achievements, we anticipate a higher level of confidence overall and in other subgroups.

Our data suggested that Group A significantly differed from only Group B in an overall comparison of “I feel I understand the material in this course”. We found it interesting that Group A would differ significantly only from the control with the same instructor and flipped learning environment. We are unsure what this means as an overall comparison – it could say something about traditional classroom instruction vs flipped classroom approaches for CS, it could say something about the instructors themselves, or it could say something about the students enrolled in each of the classes. While there are many variables that could affect this and many unknowns as to why there are only differences from one group, we place more emphasis on the clearer results when divided by gender and prior experience.

When comparing the perceived understanding of students by gender, we saw a significant increase in female students using SAIL feeling confident about understanding course material against all controls ($p<.05$ all, $p<.01$ Group D). Similarly, we saw a drastic increase for students without prior CS experience using SAIL feeling they understood course material ($p<.001$ - all except Group C, $p<.01$.) There were no significant differences between the treatment and control groups for male students or students with prior CS experience. SAIL increased female confidence to comparable levels with their male peers from all sections and elevated confidence for students without prior experience comparable to those with former experience. In looking at all of the sections, the most highest confidence reports we see are for those students who with prior CS experience – which would be expected, as they may experience a shorter learning curve and be more comfortable with some form of programming concepts or syntax. However, with SAIL we see student confidence without prior experience brought to a similar level overall.

Additionally, we see a direct correlation between all groups for statistically different confidence levels and intent to pursue further CS education across all subgroups. (Q5 $\rightarrow$ Q4. This makes the argument that SAIL helped increase confidence in studying CS, which in-turn encouraged more students to continue their studies. This finding strengthens arguments from our literature that confidence levels, especially for females or new students, significantly impacts the pursuit of CS, and we demonstrate that SAIL could help alleviate these concerns.

Q6 - *"I enjoy this course"*

Our hypothesis was that SAIL would help increase enjoyment for all students, disregarding subgroup. This was based on the literature review of interest-based learning and it's effect on enjoyment and learning [30].

Our expectations were largely met. Students using SAIL, overall, reported significantly higher feelings of motivation than all control groups (B, $p < .01$, all others $p < .001$). When divided by subgroups, both males and females showed significantly increased enjoyment levels than the control groups. Additionally, students without prior CS experience were significantly higher than all control groups, while students with prior experience were significantly higher than one of the three controls (Group C, $p < .05$). This leads us to the conclusion that SAIL helped increase enjoyment overall, but most for students without prior CS exposure (disregarding gender).

We see a very large disparity between student enjoyment when comparing students with and without prior CS experience - as students with prior CS experience reported much higher enjoyment of the course. All students without prior CS experience reported lower enjoyment than those students with prior CS experience, though the disparity is minimized for the students without experience who used SAIL. Contrastingly, we do see a slightly higher overall inclination of enjoyment for male students than female students, but the treatment group females reported a higher level of enjoyment than the all of the control group's males.

These findings indicate that SAIL had a positive experience overall on student enjoyment of programming. As studies have shown that increased enjoyment leads to better learning outcomes [30], this leads us to believe that the results seen in Exam 1 were good indications as to SAIL's positive impact on performance. To dig deeper into student enjoyment and whether it was truly caused by SAIL, we discuss our final perception question below.

Q3 - *"I enjoy programming more than I expected"*

In addition to asking students to indicate enjoyment of the course, students were asked to rate their feelings of "I enjoy programming more than I expected". With overall enjoyment scores high, it is interesting to compare these enjoyment levels to their feelings that they enjoyed programming more than they expected. We hypothesize that all students who show increased enjoyment indicate that they enjoy programming more than expected. Comparing these two questions speak to how students perceived CS when they entered the course (drawing from the literature on misperceptions still prevailing about CS), and if/how SAIL impacted these misperceptions.

Our data shows a direct correlation that treatment and control group students who enjoyed the course, also enjoyed the course more than expected. Similar to the results from Q3, overall, males, and students with no prior CS experience all also claimed they enjoyed programming more than they expected. The only difference in the data is that female students were statistically significant from all groups for Q6, while only statistically significant from 2/3 control groups for Q3, though the third control group (Group B) had a p-value of exactly .05, indicating to us that the data matches pretty closely. Also, students with prior experience also showed significance from only one control group, though those control groups differed (Group C, Q6 and Group D, Q3). Our takeaway from this is that student's who enjoyed the course, often enjoyed it more than they expected. This upholds the literature about the common misperception of who will enjoy computer science, and adds the argument than, in the correct environment, computer science can be enjoyed by all.

8.2.2 SAIL - Perception

The students who utilized SAIL as their primary learning source for the six week pilot study were surveyed about their experience with the system. We asked students about their overall experience, the video lectures, and the interest-based exercises.

Overall Experience

Most significantly, we found that students overwhelmingly felt that SAIL provided a positive overall learning experience (87%) and that they would choose to use SAIL for learning future topics (82%). Such high values for these two responses suggest that students perceived SAIL as a beneficial learning experience and, when compared to the flipped-classroom approach they had previously utilized in the course, SAIL may be the preferred option.

Many systems may have incredible potential to affect learning, but without usability, these systems may not be accepted by the community and therefore their impact will be hindered. The well-known technology acceptance model (TAM) outlines a list of factors that determine if a technology will be accepted by the community. As usability is one of they key factors in determining technology acceptance [60], kept usabilty and user experience (student and instructor) as a priority throughout the design process. These high responses of an overall positive experience and preferring SAIL to the traditional instructional approach are encouraging measures that we are on the right track with our design. As with any initial framework, there is still much room for improvement by incorporating more HCI techniques with additional time and resources. While the pilot featured only an initial version of these designs, we wanted to test students on the initial usability of the system to determine if our design thus far has promoted or hindered technology acceptance. 80% of student responses were positive that "[They] felt the SAIL website was easy to use". This high percentage of agreement about SAIL's ease of use indicates that the majority of students were able to utilize the resources in SAIL without undergoing a major learning curve to use the system. This high percentage also suggests that according to the TAM, SAIL is on the right path for widespread technology use and acceptance.

In this study, the researchers created all adaptive content - including video lectures and adapted exercises. We wanted to control the quality of the video lectures and ensure that all adapted exercises, regardless of interest selection, covered the same topics and the same criteria. These decisions were made to control for any factors about level of difficulty or varying instructor videos that may impact student perception and performance. Due to these controls, a future study would be needed to assess how instructors perceive the usability of the system. However, as the interfaces for both student and instructor are similar and both utilize the interactive knowledge map as their main tool, we believe that high positive response from students would carry over into the instructor usability as well.

Interest-based Exercises

We asked students several questions specifically about the interest-based exercises to gain an understanding of how students perceived their impact.

As discussed in Section 8.2.1, we saw a significant increase in course enjoyment for students using SAIL when compared to their control groups. In this section, we asked students if they felt the interest-based exercises helped increase course enjoyment to infer what role the interest-based learning exercises played in this increase. With a 70% positive response, 17% neutral, and only a 13% negative response, it seems most students felt that the interest-based exercises in SAIL did play a role in increasing enjoyment.

In creating the pilot study for CS, we were only able to create three interest categories for students to choose from due to the overwhelming amount of effort it takes to create adaptive content. Results indicated that even with the limited categories available in the pilot, 77% of students positively assessed that the exercises were relevant to their interests, with no significant differences in these numbers between subgroups. A similar amount, 72%, identified positive feelings of preferring these interest-based exercises to the generic exercises given before. Even with the limited number of interest categories currently implemented in SAIL, students still preferred these exercises to the generic exercises. We believe these numbers help corroborate that situational interest was aroused, in all subgroups of students. As situational interest has been shown to help increase motivation and learning outcomes, we believe this is a positive sign that SAIL is helping to increase student motivation and learning in introductory CS.

Additionally, 87% of students reported that they liked the option to adapt the class exercise based on their interests. Even though only 72% of students preferred these exercises to the generic exercises and only 77% of students felt these exercises were relevant to their interests, 87% liked that they played an active role in choosing their exercises. One design of SAIL was to transform the student from a passive to an active participant in their learning experience. We believe these numbers show that students like the ability to customize their learning experience, and that through the continued creation of more adaptive exercises that may be relevant to more students' interests, the impact of SAIL on student learning can only improve.

Video Lectures

Students were asked questions about the lecture videos in particular to assess what role they may have played in SAIL's overall impact.

81% of students indicated positive responses for referencing the SAIL videos when working on programming assignments. We believe this high number, combined with the high performance for many students who utilized SAIL shows that the lecture videos created for SAIL have aided in helping students learn and have provided a valuable resource.

75% of students indicated they preferred watching the lecture videos in SAIL to the traditional learning approach 89% of students felt the videos in SAIL helped them learn. This, coupled with the students increase confidence and performance lead us to believe that the lecture videos - short, 20 minute videos designed to mix lecture content with working examples - were beneficial in helping students learn.

One key design of SAIL is that it is not limited by domain and can be scalable to fit many instructional designs and class sizes. The high feelings about the lecture videos in SAIL are encouraging, as movements towards online and distance learning programs are growing. We hypothesize that these effects on enjoyment, learning, confidence, and perception of CS can be scalable to larger class size, distance learning, and self-learning courses - helping aid in training more students as well as fill the shortcomings of too few educators. We believe that this data shows that the short, well-designed lecture videos can help students learn while the interest-based exercises can increase enjoyment, motivation, and help them learn better. The design of SAIL couples these two aspects of learning together, and has shown that it can provide an overall positive and preferred learning experience for students - with nothing preventing it from being implemented in an eLearning tool with any class size in the future.

8.3 Answering Research Questions

8.3.1 RQ 1 - Performance

How does the use of SAIL impact performance measures?

Our results show that SAIL provides students with a similar or superior overall educational experience to a best-practice University level class. Additionally, SAIL was observed to minimize performance disparities caused by gender and students new to studying CS.

This data suggests that SAIL could be the solution to deliver customized and compelling curriculum to at multiple levels (including K-12, community colleges, and Universities), for both educators and students alike, while ensuring a standard of competency is met.

8.3.2 RQ 2 - Confidence

How does the use of SAIL impact perceived learning and confidence in course material?

Our results show that students using SAIL reported the highest feeling of understanding course material, overall. This effect was magnified for students that were female or had no prior CS exposure, demonstrating significantly increased perceived learning and confidence against all control groups and reporting confidence level similar to their peers that were male / had prior CS experience.

This overall increase in confidence, especially for female students and those without prior CS exposure is important as the literature showed that confidence was a major hindrance in students selecting CS as their major. We can infer, from our data, that the same students showing an increase in confidence, also show an increase in intent to continue their CS education. We thus conclude that SAIL helped increase confidence in CS, leading to an increased intent to continue CS studies.

8.3.3 RQ 3 - Perception and Motivation

How does the use of SAIL influence students' attitudes and perception towards Computer Science (or STEM fields)?

Overall, the trend is that students who used SAIL reported an equal or higher perception of CS than their peers in the control groups. This effect was greatest in significantly increasing overall enjoyment in the course and programming. These overall effects are also seen with significance in all subgroups (males, females, new students), but less-so for students coming with some prior CS background. Results of increased enjoyment overall speak largely to the motivation of students. It is probable that the increase in enjoyment led to an increase in intrinsic feelings and intrinsic motivation to the course, though further investigation would need to be done to corroborate this theory.

We also saw a significant shift in female students and new students' intent to continue their CS studies. Our data suggests that these feelings may be partially linked to enjoyment and confidence levels inspired by SAIL. Male students and students with no prior experience showed significant indication that SAIL helped them perceive how useful CS can be to their other interests. Though not significantly impacted for females or students with prior CS experience, their feelings were not negatively impacted. It is important to note that at no point in this study, did we find student perception of CS negatively impacted by SAIL.

Our findings demonstrate that SAIL has helped increase enjoyment of the course, that students are enjoying the course more than they expected, that they feel confident about their abilities in the course, that increased interest in computer science is inspried for some, and that many under-represented groups are being encouraged to continue their CS studies.

8.3.4 RQ 4 - SAIL Overall Experience

How did students perceive the overall experience of SAIL when compared to the traditional mode of instruction?

Students indicated that SAIL provided an overall positive (87%) and preferred (82%) learning experience to traditional instruction. With extremely positive ratings for:

- The ability to adapt the exercises based on their interests (87%)
- Feeling the videos in SAIL helped them learn (89%)

As students who used SAIL had exposure to instruction with and without the intervention, their feelings of preferring SAIL to the traditional instructional approach indicate many positive outlooks for SAIL's future. Additionally, 80% of responses were positive that SAIL was easy to use - an important feature for widespread acceptance and usability.

We believe these results are scalable to other class sizes and distance learning programs.

8.3.5 RQ 5 - Diversity

Does the impact of SAIL on (1)performance, (2)confidence, and (3)perception differ based on diversity factors such as gender, ethnicity, or prior exposure to CS?

The overall trend in performance is that male students performed better than their female peers across the board. SAIL helped remedy this divide by elevating female performance on all assessments to be higher than their female peers from control groups and comparable to their male peers. This trend was similar when comparing students with and without prior CS experience, and SAIL also seemed to elevate the performance of students without prior exposure, suggesting a decreased learning curve when compared to control groups.

SAIL was seen to significantly elevate confidence levels for both females and students with no prior experience to CS. Additionally, both females and new students indicated significantly higher intentions to continue their CS education than the control groups.

Our data overall shows that SAIL helped level the playing field for males and females and for new students vs. students with prior CS experience. The measures affected include: performance, confidence levels, interest, enjoyment, and intent to continue CS education.

These results provide an exciting outlook for how SAIL can help alleviate many of the diversity issues within STEM education.

8.3.6 RQ 6 - Recruitment

Does SAIL demonstrate a potential to impact attraction and recruitment to STEM disciplines?

Our data shows that SAIL helped correct many of the misperceptions and hindrances that keep people, particularly minorities, from pursuing a CS field. There was a significant increase in overall enjoyment and a significant increase in females and students without prior CS exposure's intent to pursue more CS courses. These measures, coupled with the increase in performance, confidence, and a reported overall positive learning experience indicate that SAIL could have a significant impact on recruiting a larger and more diverse population of students over time.

9 | Conclusion

9.1 Overview of Work

The aim of this research is to alleviate many challenges faced in STEM education through the creation of a scalable, adaptive learning framework that supports interest-based learning (IBL) in multiple domains. This work presents the design and pilot of SAIL, a System for Adaptive Interest-based Learning, to easily facilitate IBL in an adaptive and scalable platform. SAIL is not limited by domain, but was designed with STEM subjects in mind due to their high applicability in other fields. SAIL was designed to help alleviate many of the concerns in STEM education by providing a competent and compelling curriculum that delivers individualized instruction to increase motivation, performance and fill the gaps in STEM education. Additionally, the interest-based nature of SAIL was meant to showcase the interconnectivity of STEM subjects with other fields, helping to to combat misperceptions, increase motivation, and attract a larger and more diverse population of students. A large pilot study (N=307) in the context of introductory programming (Java) was conducted at the University of Georgia, comparing a class using SAIL to three other classes with varying control conditions.

This study resulted in new quantitative and qualitative knowledge about how SAIL can impact introductory Computer Science (CS) as well as assessing viability for other STEM fields, including K-12 STEM education. Specifically, this work contributes to the following issues in the CS and STEM education communities:

- Remedying gender disparities in confidence and achievement
- Elevating overall performance
- Increasing enjoyment and motivation
- Increasing attraction and retention
- Raising the universal standard of STEM education

Additionally, the design of SAIL demonstrates many novel contributions to the facilitation of interest-based learning in adaptive learning systems, including:

- A framework for IBL applicable in multiple domains
- Addressing the bottle-neck of adaptive content creation through community knowledge sharing
- A scalable design not limited by class size or instructional design

9.2 Future Directions

As with any research, the acquisition of new knowledge inspires new questions. While this works contributed significantly to the design implications of an adaptive learning system supporting IBL in multiple STEM subjects and the knowledge of SAIL's impact and viability in STEM education, these contributions act as a springboard for future research directions.

While the initial concept for SAIL has been designed and tested, many improvements could help SAIL evolve in the future. Now that SAIL has been evaluated showing success, we can look towards the future of growing the system.

As with any new design, we feel the interface and usability of SAIL could benefit from continued improvement. As mentioned throughout this paper, the groundwork has been laid to support a future of community knowledge sharing through an easy-to-use interactive map. At this time, any content uploaded to SAIL can be reused by other courses, but an enhanced organization for materials would be useful for easy sorting amongst a large amount of materials. A future study could also involve a pilot where the teachers were surveyed about their experience using SAIL in a classroom.

As the problems in STEM education are complex, more angles could be observed for SAIL's impact on the STEM challenges. A long-term study looking at major selection and retention could help us understand if SAIL genuinely helped recruit and retain more majors from entry-level to graduation. Additionally, more studies could be conducted to help strengthen the evidence of SAIL's impact on performance and perception of STEM fields.

Finally, our pilot for SAIL was limited to a six-week study in a university setting. As many of the STEM education issues reach back into earlier education, testing SAIL for a full year in a K-12 classroom would be an important step. For this to be possible, material would need to be created to support an entire course. This curation of content would take the work of several, yet allow us to observe the full effect of an entire course taught with SAIL at the K-12 level.

Bibliography

[1] "What's wrong with this picture?" [Online]. Available: https://code.org/promote

[2] K. Aguar, H. R. Arabnia, J. B. Gutierrez, W. D. Potter, and T. R. Taha, "Making cs inclusive: An overview of efforts to expand and diversify cs education," in *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 2016, pp. 321–326.

[3] S. Hidi and K. A. Renninger, "The four-phase model of interest development," *Educational psychologist*, vol. 41, no. 2, pp. 111–127, 2006.

[4] C. A. Walkington, "Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes," *Journal Of Educational Psychology*, vol. 105, no. 4, pp. 932–945, 2013.

[5] C. B. Lockard and M. Wolf, "Occupational employment projections to 2020," *Monthly Lab. Rev.*, vol. 135, p. 84, 2012.

[6] C. Corbett, C. Hill *et al.*, "Solving the equation: the variables for women's success in engineering and computing," 2015.

[7] S. Zweben and B. Bizot, "2016 taulbee survey," *Computing Research Association*, vol. 29, no. 5, 2016.

[8] ——, "2014 taulbee survey," *Computing Research Association*, vol. 27, no. 5, 2015.

[9] S. Beyer, K. Rynes, J. Perrault, K. Hay, and S. Haller, "Gender differences in computer science students," in *ACM SIGCSE Bulletin*, vol. 35, no. 1. ACM, 2003, pp. 49–53.

[10] J. Margolis and A. Fisher, *Unlocking the clubhouse: Women in computing.* MIT press, 2003.

[11] E. Mercier, B. Barron, and K. O'Connor, "Images of self and others as computer users: The role of gender and experience," *Journal of computer assisted learning*, vol. 22, no. 5, pp. 335–348, 2006.

[12] S. Cheryan, V. C. Plaut, C. Handron, and L. Hudson, "The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women," *Sex roles*, vol. 69, no. 1-2, pp. 58–71, 2013.

[13] A. Master, S. Cheryan, and A. N. Meltzoff, "Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science." *Journal of Educational Psychology*, vol. 108, no. 3, p. 424, 2016.

[14] C. Alvarado and Z. Dodds, "Women in cs: an evaluation of three promising practices," in *Proceedings of the 41st ACM technical symposium on Computer science education.* ACM, 2010, pp. 57–61.

[15] T. Newhall, L. Meeden, A. Danner, A. Soni, F. Ruiz, and R. Wicentowski, "A support program for introductory cs courses that improves student performance and retains students from underrepresented groups," in *Proceedings of the 45th ACM technical symposium on Computer science education*. ACM, 2014, pp. 433–438.

[16] H. M. College, "Class of 2016 a milestone for women in physics | harvey mudd college news," https://www.hmc.edu/about-hmc/2016/05/17/class-2016-milestone-women-physics/, Jun 2016.

[17] A. Fisher, J. Margolis, and F. Miller, "Undergraduate women in computer science: experience, motivation and culture," in *ACM SIGCSE Bulletin*, vol. 29, no. 1. ACM, 1997, pp. 106–110.

[18] S. J. Gibbons, L. S. Hirsch, H. Kimmel, R. Rockland, and J. Bloom, "Middle school students' attitudes to and knowledge about engineering," in *International conference on engineering education, Gainesville, FL*, 2004.

[19] Z. Henkel, P. Doerschuk, and J. Mann, "Exploring computer science through autonomous robotics," in *Frontiers in Education Conference, 2009. FIE'09. 39th IEEE*. IEEE, 2009, pp. 1–6.

[20] L. Mannila, V. Dagiene, B. Demo, N. Grgurina, C. Mirolo, L. Rolandsson, and A. Settle, "Computational thinking in k-9 education," in *Proceedings of the Working Group Reports of the 2014 on Innovation & Technology in Computer Science Education Conference*. ACM, 2014, pp. 1–29.

[21] "Cs education statistics," http://www.exploringcs.org/resources/cs-statistics, 2015.

[22] "Ap computer science principles - a new ap course - advances in ap® - the college board | advances in ap," http://collegeboard.org/APCSP., 2015.

[23] "Ap computer science a," http://media.collegeboard.com/digitalServices/pdf/ap/ap-course-overviews/ap-computer-science-a-course-overview.pdf.

[24] "K–12 computer science framework (2016)," k12cs.org.

[25] A. M. Sinatra, "–a review of self-reference and context personalizationin different computer-based educational domains," *Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling*, vol. 4, p. 107, 2016.

[26] C. S. Symons and B. T. Johnson, "The self-reference effect in memory: a meta-analysis." *Psychological bulletin*, vol. 121, no. 3, p. 371, 1997.

[27] D. J. Turk, K. Gillespie-Smith, O. E. Krigolson, C. Havard, M. A. Conway, and S. J. Cunningham, "Selfish learning: The impact of self-referential encoding on children's literacy attainment," *Learning and instruction*, vol. 40, pp. 54–60, 2015.

[28] R. E. Mayer, S. Fennell, L. Farmer, and J. Campbell, "A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style." *Journal of Educational Psychology*, vol. 96, no. 2, p. 389, 2004.

[29] R. Moreno and R. E. Mayer, "Engaging students in active learning: The case for personalized multimedia messages." *Journal of Educational Psychology*, vol. 92, no. 4, p. 724, 2000.

[30] D. I. Cordova and M. R. Lepper, "Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice." *Journal of educational psychology*, vol. 88, no. 4, p. 715, 1996.

[31] C. S. Hulleman, A. M. Durik, S. B. Schweigert, and J. M. Harackiewicz, "Task values, achievement goals, and interest: An integrative analysis." *Journal of educational psychology*, vol. 100, no. 2, p. 398, 2008.

[32] O. Köller, J. Baumert, and K. Schnabel, "Does interest matter? the relationship between academic interest and achievement in mathematics," *Journal for Research in Mathematics Education*, pp. 448–470, 2001.

[33] S. Hidi, "Interest, reading, and learning: Theoretical and practical considerations," *Educational Psychology Review*, vol. 13, no. 3, pp. 191–209, 2001.

[34] M. M. Clifford, "Students need challenge, not easy success," *Educational Leadership*, vol. 48, no. 1, pp. 22–26, 1990.

[35] J. I. Rotgans and H. G. Schmidt, "The role of interest in learning: Knowledge acquisition at the intersection of situational and individual interest," in *The Science of Interest*. Springer, 2017, pp. 69–93.

[36] Hypermedia and multimedia. [Online]. Available: http://www.encyclopedia.com/doc/1G2-3401200220.html

[37] C. Bailey, S. R. El-Beltagy, and W. Hall, "Link augmentation: A context-based approach to support adaptive hypermedia," in *Hypermedia: Openness, Structural Awareness, and Adaptivity*. Springer, 2002, pp. 239–251.

[38] P. Brusilovsky, *Methods and techniques of adaptive hypermedia.* User Modeling and User-Adapted Interaction, 6(2-3), 87-129, 1996.

[39] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in *The adaptive web.* Springer-Verlag, 2007, pp. 3–53.

[40] Y. Akbulut and C. S. Cardak, "Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011," *Computers & Education*, vol. 58, pp. 835–842, 2012.

[41] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering Education*, vol. 78, no. 7, 1988.

[42] J. W. Keefe, *Learning Style Theory and Practice.* ERIC, 1987.

[43] G. Pask, "Styles and strategies of learning," *British Journal of Educational Psychology*, vol. 46, 1976.

[44] H. E. Gardner, *Intelligence Reframed: Multiple Int.* Perseus Books Group, 2000.

[45] J. Van Seters, M. Ossevoort, J. Tramper, and M. J. Goedhart, "The influence of student characteristics on the use of adaptive e-learning material," *Computers & Education*, vol. 58, no. 3, pp. 942–952, 2012.

[46] C. Chen, "An adaptive scaffolding e-learning system for middle school students' physics learning," *Australasian Journal Of Educational Technology*, vol. 30, no. 3, pp. 342–355, 2014.

[47] S. Sparrow, "Discover adaptive elearning - smart sparrow - adaptive learning platform," https://www.smartsparrow.com/adaptive-elearning/, 2014, (Visited on 10/25/2015).

[48] T. Zimmer, "The adaptive advantage: How e-learning will change higher ed - forbes," http://www.forbes.com/sites/ccap/2014/12/01/the-adaptive-advantage-how-e-learning-will-change-higher-ed/, Dec 2014.

[49] B. G. Prusty and C. Russell, "Engaging students in learning threshold concepts in engineering mechanics: adaptive elearning tutorials," in *International Conference on Engineering Education (ICEE2011), Belfast, Australia*, 2011.

[50] R. Riddell, "Adaptive learning: The best approaches we've seen so far," Oct 2013. [Online]. Available: http://www.educationdive.com/news/adaptive-learning-the-best-approaches-weve-seen-so-far/187875/

[51] R. Sottilare, K. Brawner, A. Sinatra, and J. Johnston, "An updated concept for a generalized intelligent framework for tutoring (gift). orlando, fl," *US Army Research Laboratory*, 2017.

[52] A. N. Kumar, "Model-based reasoning for domain modeling, explanation generation and animation in an its to help students learn c+," *Model Based Systems and Qualitative Reasoning for Intelligent Tutoring*, p. 45, 2002.

[53] E. R. Sykes and F. Franek, "An intelligent tutoring system prototype for learning to program java tm." IEEE Xplore. Proceedings of the The 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03), 2003.

[54] S. Abu-Naser, A. Ahmed, N. Al-Masri, A. Deeb, E. Moshtaha, and M. AbuLamdy, "An intelligent tutoring system for learning java objects," *International Journal of Artificial Intelligence and Applications (IJAIA)*, vol. 2, no. 2, 2011.

[55] E. Carter and G. D. Blank, "An intelligent tutoring system to teach debugging," in *Artificial Intelligence in Education.* Springer, 2013, pp. 872–875.

[56] H. C. Lane and K. VanLehn, "Coached program planning: dialogue-based support for novice program design," in *ACM SIGCSE Bulletin*, vol. 35, no. 1. ACM, 2003, pp. 148–152.

[57] R. Z. Cabada, M. L. B. Estrada, F. G. Hernández, R. O. Bustillos, and C. A. Reyes-García, "An affective and web 3.0-based learning environment for a programming language," *Telematics and Informatics*, 2017.

[58] S. Devasani, *Intelligent tutoring system authoring tools for non-programmers*. Iowa State University, 2011.

[59] R. A. Sottilare and H. K. Holden, "Motivations for a generalized intelligent framework for tutoring (gift) for authoring, instruction and analysis," in *AIED 2013 Workshops Proceedings*, vol. 7, 2013, p. 1.

[60] F. D. Davis, "A technology acceptance model for empirically testing new end-user information systems: Theory and results," Ph.D. dissertation, Massachusetts Institute of Technology, 1985.

[61] J. Gutierrez, "Literatronic: The use of Hamiltonian cycles to produce adaptivity in literary hypertext," in *The Bridges Conference*, Aug. 2006, pp. 215–224.

[62] J. B. Gutierrez and M. C. Marino, *Literatronica: Adaptive digital narrative.* Paper presented at the Proceedings of the hypertext 2008 workshop on Creating out of the machine: hypertext, hypermedia, and web artists explore the craft, 2008.

[63] K. Aguar, C. C. Sanchez, D. B. Beltran, S. Safaei, M. Asefi, J. Arnold, P. Portes, H. R. Arabnia, and J. B. Gutierrez, "Considerations on interdisciplinary instruction and design influenced by adaptive learning. sail-sb: A case study involving biology, computer science, mathematics, and statistics," *arXiv preprint arXiv:1703.06010*, 2017.

[64] Make computer science in k-12 count. [Online]. Available: https://code.org/files/convince_your_school_or_state.pdf

[65] "Code.org cs stats," https://code.org/images/cs-stats/more-jobs-than-students.png.

[66] A. Nager and R. D. Atkinson, "The case for improving u.s. computer science education," May 2016. [Online]. Available: https://itif.org/publications/2016/05/31/case-improving-us-computer-science-education

[67] R. Morgan and J. Klaric, "Ap students in college: An analysis of five-year academic careers," 2007.

[68] C. Spradling, D. Linville, M. P. Rogers, and J. Clark, "Are moocs an appropriate pedagogy for training k-12 teachers computer science concepts?" *Journal of Computing Sciences in Colleges*, vol. 30, no. 5, pp. 115–125, 2015.

[69] J. Liu, J. Wilson, D. Hemmenway, Y. Xu, and C.-H. Lin, "Oh snap! a one-week summer computing workshop for k-12 teachers," in *Computer Science & Education (ICCSE), 2015 10th International Conference on*. IEEE, 2015, pp. 663–668.

[70] M. Guzdial and B. Ericson, "Georgia computes!: An alliance to broaden participation across the state of georgia," *ACM Inroads*, vol. 3, no. 4, pp. 86–89, Dec. 2012. [Online]. Available: http://doi.acm.org/10.1145/2381083.2381104

[71] J. Clark, M. P. Rogers, C. Spradling, and J. Pais, "What, no canoes? lessons learned while hosting a scratch summer camp," *Journal of Computing Sciences in Colleges*, vol. 28, no. 5, pp. 204–210, 2013.

[72] "Uga 2016 first-year class profile," https://www.admissions.uga.edu/prospective-students/first-year/fy-profile.

[73] "Which top universities take the most transfers?" WP Company. [Online]. Available: https://www.washingtonpost.com/apps/g/page/local/161which-top-universities-take-the-most-transfers/2036/?tid=a_inl%2C

[74] N. C. f. S. National Science Foundation and E. Statistics, "Women, minorities, and persons with disabilities in science and engineering: 2017. special report nsf 17-310," www.nsf.gov/statistics/wmpd/, 2017, (Accessed on 02/06/2017).

[75] S. Cheryan, V. C. Plaut, P. G. Davies, and C. M. Steele, "Ambient belonging: how stereotypical cues impact gender participation in computer science." *Journal of personality and social psychology*, vol. 97, no. 6, p. 1045, 2009.

[76] S. Cheryan, A. N. Meltzoff, and S. Kim, "Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes," *Computers & Education*, vol. 57, no. 2, pp. 1825–1835, 2011.

[77] V. Chan, K. Stafford, M. Klawe, and G. Chen, "Gender differences in vancouver secondary students," in *Women, work and computerization.* Springer, 2000, pp. 58–69.

[78] M. Evans, P. Kelley, and J. Kelley, "Identifying the best times for cognitive functioning using new methods: Matching university times to undergraduate chronotypes," *Frontiers in human neuroscience*, vol. 11, 2017.

[79] D. Wood, J. S. Bruner, and G. Ross, "The role of tutoring in problem solving," *Journal of child psychology and psychiatry*, vol. 17, no. 2, pp. 89–100, 1976.