The Effect of Early Computer Science Exposure in Choosing Computer Science as a College Major

Joseph T. Isaac, Jr.
School of Computing
Clemson University
Clemson, SC 29634
jisaac@clemson.edu

ABSTRACT
This paper describes how early aged students' exposure to computer science affects their later life decision in choosing their college major to be computer science. In this paper, an experiment is conducted to justify this statement. During a 4 week period, two groups of students without prior computer science experience will be lectured and complete assigned tasks. One group, the treatment group, will receive computer science lecturing and programming assignments. The other group, the control group, will receive freedom to choose their own lecturing and assignments. Preliminary results suggest that having early exposure to computer science will lead to studies, a career, and/or hobby in computer science. With underrepresented groups in computing at a low enrollment, this paper seeks to pave way to solve this problem.

Author Keywords
Computer Science Education; broadening participation; teaching; education; software engineering.

ACM Classification Keywords
Human Factors; Design; Performance.

INTRODUCTION
Since the year 2000, computer science enrollments in college have fluctuated in the United States. It has come to understanding that underrepresented groups in computing are less likely to enroll in computer science (CS). As defined by the National Science Foundation (NSF), underrepresented groups in computing are women, African Americans, Hispanics, native American, and persons with disabilities [2]. The U.S. population is estimated to be about 316,000,000. About 70% of that population is underrepresented groups [6]. Diversity in computer science is necessary not only because of job openings, but producing effective solutions to worldwide problems. This study will explore the affect of computer science exposure prior to college and reasoning of choosing computer science as a college major.

The objective of this study is to prove that exposure to computer science at an early age leads to pursuits of computer science education. This study examines the effect of computer science instruction and programming activities on students with very minimal prior knowledge of computer science. It was hypothesized that students who have a strong understanding and experience with computer science are more likely to choose computer science as a college major.

BACKGROUND
In "Ubiquity Symposium The Science In Computer Science" [6], the article investigates broadening computer science enrollment in the United States. It's an interview with Jan Cuny, a diversity program leader at NSF, about reasons of low diversity enrollment in CS, why diversity in the field is important, and efforts made. The article features various stats that are relevant to the premise of my research topic, computing education.

RELATED WORK
In "Glitch Game Testers", a group of young African American men were introduced to video game testing as a gateway to lead them to the field of computing [3]. This involved training and an additional contextualized computing curriculum for them to gain more exposure to computer science education [4]. It address the cultural differences in how African American males practice in computing, which lead to a unique educational intervention that's different from their majority counterparts. In conclusion, its findings were that there were strong indicators that both testing video games and pairing them with Computer Science workshops have increased interest in computer science their group of young African American males.

In "LilyPad Arduino", it investigates computer science education through a fabric based construction kit that allows novices to design and build their own soft wearables [1]. This work challenges students to think of computing outside of typical desktop applications exclusively. With Arduino, embedded systems are used to landscape computational media and equipping them with adaptive sensors. In this

study, students aged 10-14 met each weekday for one week to build their own electronic fashion.

METHODS

Target Sample & Participants

To conduct this experiment, the experimenter will target 100 random underrepresented 10th grade high school students without computer science experience. The high school of choice will be Greenville Senior High Academy in Greenville, SC. Computer science experience would be detected through access to student course records to see that they have not taken any computer programming courses. Also, a pre-screening survey will filter out students with computer science experience outside of school. These students will be the participants of the experiment. In addition, the experiment will have one seasoned computer science teacher, one general studies teacher, and five computer science lab assistants.

Data Collection Equipment

Two computer lab lecture rooms, a pre-test, and a post-test will be used for data collection. Each computer lab lecture room will have 55 computers, 55 computer desks and chairs, one projector, and one large white board. The pre-test and post-test will be identical. They will have fields that will obtain their name, student ID number, and ranking of their prospective college majors in order of preference. The following majors will be listed for ranking: Computer & Information Sciences, Psychology, Nursing, Biology/Biological Sciences, Education, English Language & Literature, Economics, Communications Studies/Speech Rhetoric, Political Science & Government, and Engineering.

Experimental Design

This experiment will use a between-subjects design. The independent variable will be the subject matter of lecturing. The dependent variable will be the difference in performance on the pre-test and post-test. The treatment group and control group will be split 50 students each. The treatment group will take an introductory computer science course for one hour per session over the course of four weeks. The control group will have a study hall session for one hour per session over the course of four weeks where they will be able to learn or do any academic endeavor.

Experimental Task

For the treatment group, there will be a 30 minute lecture and 30 minute lab assignment in each lecture. These lectures and assignments will teach students basic programming skills and concepts. All participants are expected to give their best efforts to attend all lectures, engage in lectures, and perform to the best of their abilities on lab assignments. For the control group, students will be able to work on any academic endeavors they choose. They are expected to take advantage of the general studies teacher and assistants to help tutor them in any of their academic assignments.

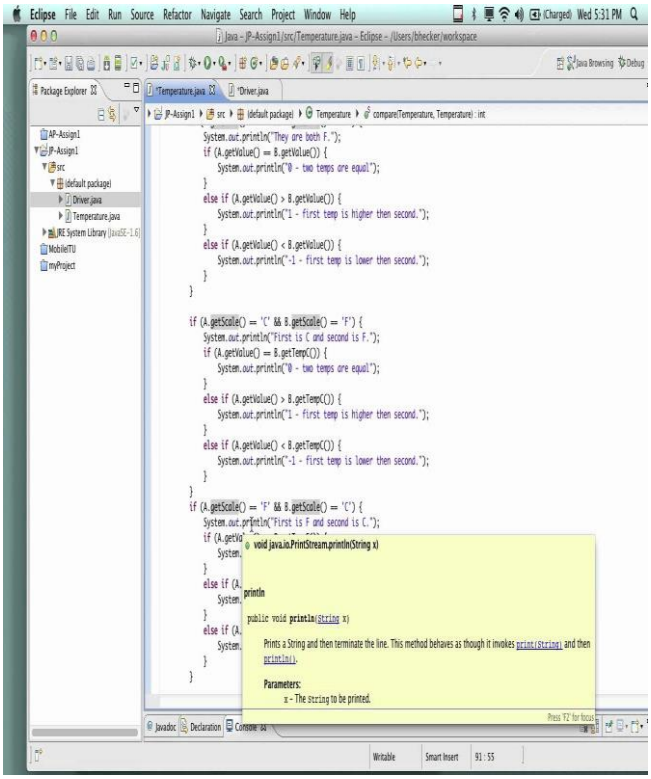

Figure 1. Java Programming Assignment

Procedure

At the beginning of both group's first class session, they will take the prospective college major ranking pre-test. Each week, the treatment group will learn the Java programming language and apply what they learned in their lab assignments. The computer science teacher will instruct the sessions while the five assistants help students. Over the 4 weeks, the control group will have study hall in the same classroom, but different times than treatment group. The control group is allowed to do any academic activity as there will be no formal lecturing. At the conclusion of the four weeks, both groups will take the same test from the pre-test as a post-test.

1. Computer & Information Sciences
2. Psychology
3. Nursing
4.Biology/Biological Sciences
5.Education
6.English Language & Literature
7.Economics
8.Communications Studies/Speech & Rhetoric
9.Political Science & Government
10.Engineering

Figure 2. Potential College Major Ranking

Measures

The differences in rankings between the pre-test and the post-test will be used to measure data. From that data, the experimenter will gather the mean, media, mode, and standard deviation. With such statistics, the experimenter will be able to display it in visual formats and analyze it to bring meaning to my research (See Figure 1).

$$Score = (Pretest\ Rank) - (Posttest\ Rank)$$

Figure 3. Score Formula

RESULTS

Sample Description

The expected total enrollment of students at Greenville Senior High Academy is 1,407. From that total, 59% of them are minority students. In the 10^{th} grade, there are expected to be 300 students. A sample of 100 students will be used for this experiment. Sampling a third of the 10^{th} grade enrollment is a significant sample size. 100% participation is expected by all participants.

Data Treatment

The dependent variable is the difference in ranking between the pre-test and post-test. This calculation is performed by subtracting the pre-test rank number from the post-test rank number. The score range is -9 to 9. The larger the score, the more interested a student is interested in choosing computer science as a major. The smaller the score, the less interested a student would be in choosing computer science as a major. A score of zero means there was no change. Both the control and treatment group will use the same data collection processes. This data will be collected to determine whether being exposed to computer science has an impact on choosing the field as a major.

Presentation of Results

Prospective College Major	Average Ranking Score	Standard Deviation
Computer Science	7	2
Engineering	5	1
Psychology	3	5
Nursing	-2	4
Biology	0	6
Education	3	3
English	2	2
Business	4	6
Communications	3	3
Political Science	-1	3

Table 1. Expected Data from Treatment Group

Prospective College Major	Average Ranking Score	Standard Deviation
Computer Science	2	5
Engineering	3	3
Psychology	4	2
Nursing	2	1
Biology	3	5
Education	2	4
English	4	2
Business	1	6
Communications	2	4
Political Science	4	5

Table 2. Expected Data from Control Group

For the treatment group, computer science is expected to have the largest increase in average score for choosing a computer science major as opposed to other majors. The average score for this group is expected to be seven with a standard deviation of two (See Table 1). For the control group, Political Science & Government was expected to be the largest increase in average score at four. The average computer science major score is expected to be two with a standard deviation of five (See Table 2).

DISCUSSION

Results from the experiment to evaluate the likelihood of choosing computer science as a college major after being exposed to computer science are confident. They significantly advocate that students are more willing to pursue computer science when computer science is demystified with first-hand activities. The combination of Java programming lectures and lab programming assignments engages students to inquire more about computer science [5].

With the treatment group achieving a score of a seven in comparison to a score of a two in the control group, it is evident that there is a misunderstanding of computer science (see Tables 1 and 2). This data shows that students do not understand what is entailed in computer science. The stigma of computer science plays a role in students avoiding the discipline, thinking that it's a nerdy and extremely difficult field.

The scores in the control group can be attributed to its opened ended structure (not necessarily any lecturing or assignments). It's likely that students were influenced by outside factors like other students, instructors, assistants, and family. Scenarios can include students seeking more information about majors on the pre-test, learning about different subject areas that their classmates are in, college discussions, and family discussions that may include occupations.

This study compliments the work in "Glitch Game Testers" by it analyzing a broader young audience and directly measuring computer science interest on a ranking system [3]. Glitch Game Testers focuses a lot on video game testing and linking it back to computer science with its workshop. The study in this paper accesses students only in the classroom environment. Therefore, its measurements are strictly traditional classroom related like college.

This study compliments the work in "LilyPad Arduino" by it supporting the idea of exposing young children to computing and computer programming [1]. The LilyPad Arduino mixes traditional desktop computer and textile, embedded systems fabric to create fun exposure. The study in this paper further showed that with computer science exposure students will not only enjoy such activities, but pursue computing in college.

CONCLUSION

As the National Science Foundation shows, there is a low underrepresented group enrollment in Computer Science [7]. Diversity is necessary for job openings and various problem solving scenarios. This study used a significant high school sample size to show that exposure to computer science has a positive effect on the likelihood of choosing computer science as a college major. In conclusion, demystifying computer science before college will increase the chances of a student pursuing studies in computer science.

RECOMMENDATIONS

Further development and future work of low enrollment of underrepresented groups in computer science will include researching why students are not choosing computer science as a major and how it should be presented be more appealing. Interviewing this group would be a necessary next step to better understand this issue.

ACKNOWLEDGMENTS

I would like acknowledge my course instructor, the writing center, and my fellow classmates for helping with the development of this assignment.

REFERENCES

1. Buechley, L., Eisenberg, M., and Catchen, J. The LilyPad Arduino: Using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. *CHI '08 Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, (2008), 423–432.

2. Cuny, J., Forbes, J., Taylor, H., et al. Understanding NSF Funding. *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, ACM (2014), 83–84.

3. DiSalvo, B. and Guzdail, M. Glitch game testers: African American men breaking open the console. *Digital Games Research Association Conference*, (2009), 1 – 7.

4. Forte, A. and Guzdial, M. Computers for communication, not calculation: media as a motivation and context for learning. *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, (2004), 10 pp.– .

5. Goldberg, D.S., Grunwald, D., Lewis, C., Feld, J.A., and Hug, S. Engaging Computer Science in

Traditional Education: The ECSITE Project. *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education*, ACM (2012), 351–356.

6. Snodgrass, R. Ubiquity symposium: The science in computer science. *Ubiquity 2013*, February (2013), 1–8.

7. Ward, B. Computer Science Enrollments Drop. *Computer 41*, 4 (2008), 87–89.