2006-141: THE ALICE CURRICULUM: IMPACT ON WOMEN IN PROGRAMMING COURSES

Ashlyn Hutchinson, Colorado School of Mines

Ashlyn Hutchinson (ashutchi@mines.edu) received her B.A. in Mathematics from the University of Colorado at Boulder, and her M.S. in Applied Mathematics from Colorado School of Mines. She is currently working on her Ph.D. in Statistics at Colorado School of Mines. Ashlyn is a Research Assistant for Dr. Barbara Moskal in the Mathematical and Computer Sciences Department at the Colorado School of Mines. Her research interests include engineering education and assessment.

Barbara Moskal, Colorado School of Mines

Barbara M. Moskal (bmoskal@mines.edu <mailto:bmoskal@mines.edu>) received her Ed.D. in Mathematics Education with a minor in Quantitative Research Methodology and her M.A. in Mathematics from the University of Pittsburgh. She is an Associate Professor in the Mathematical and Computer Sciences Department at the Colorado School of Mines. Her research interests include student assessment, K-12 outreach and equity issues.

Wanda Dann, Ithaca College

Dr. Wanda Dann is an Associate Professor of Computer Science at Ithaca College. Her research has encompassed program visualization and object-oriented and event-driven programming. Dr. Dann has provided leadership in the international computer science education community, serving as SIGCSE 2004 Program co-Chair and SIGCSE 2005 Symposium co-Chair.

Stephen Cooper, Saint Joseph's University

Stephen Cooper is an Associate Professor of Computer Science at Saint Joseph's University. He taught previously at Rivier College, serving as Computer Science program director. He has also worked at IBM as a systems programmer. Dr. Cooper's research interests lie in the semantics of programming languages as well as in program visualization. He has been the principal investigator for several National Science Foundation and private grants.

The Alice Curriculum: Impact on Women in Programming Courses

Abstract

The recruitment and retention of women to the field of Computer Science is a concern that has received a great deal of research attention. In particular, the number of women majoring in Computer Science has decreased over the last fifteen years, and remained low over the past five years. These continuing trends are contributing to a shortage of qualified professionals in computer science. One cause may be that traditional introductory programming courses have been found to have a greater appeal to men than to women. In addition, many women who enroll in introductory computing courses often have less previous programming experience than men. These factors negatively impact many women's attitudes towards programming, leading them to drop a computer science major. This paper discusses an effort to combat the problem of female attrition at the community college level by offering an innovative introductory programming course. Course implementation and data collection occurred during the fall 2004. Paired t-tests and analysis of covariance were used to evaluate whether changes had taken place for men and women with respect to learning and attitudes. This work was supported in part by the National Science Foundation (DUE-03020542).

I. Introduction

The *Java-based Animation: Building virtual Worlds for Object-oriented programming in Community colleges* (JABRWOC) project is a three year effort, funded by the National Science Foundation (NSF).[1] JABRWOC utilizes the Alice software in an innovative introductory programming course. Alice is a 3D animation programming environment developed by Dr. Randy Pausch at Carnegie Mellon University. The Alice software is freely available online.[2] Drs. Stephen Cooper, Saint Joseph's University, and Wanda Dann, Ithaca College, recognized the potential of the Alice software as a teaching tool and created a curriculum (henceforth known as the Alice curriculum), pedagogy, and instructional materials for the innovative approach.

The Alice curriculum was developed with the intent of effectively teaching introductory programming skills, while maintaining positive attitudes towards the subject. This curriculum simplifies first year programming by focusing on problem solving and algorithmic thinking before introducing syntax. In traditional first year courses, students must master algorithmic thinking and syntax simultaneously. By using a drag and drop editor in the Alice software, students are freed from the difficulties of syntax errors, such as missing commas or semicolons, allowing them to focus on problem solving and algorithmic thinking. Early on in the curriculum, the concept of object oriented programming is introduced.[1] A feature of the Alice software is the use of three dimensional graphics (3D) to create virtual worlds. Using the Alice software, students may test their programs in 3D virtual worlds, an environment that has great appeal to a generation of students that grew up in a multimedia environment. In traditional programming curricula, it can take several semesters before students develop the programming skills necessary to create and test 3D animations. Alice curricular materials include a textbook and classroom implementation aides such as lecture notes, slides, solutions, sample projects and tests.[3,4]

The effectiveness of the Alice curriculum is currently being measured by evaluating student learning and attitudes at three community colleges. This paper discusses the results of the assessment efforts during the fall 2004. Specifically, the results are discussed in terms of the differing effects on men and women. These results constitute the first of four semesters of summative data collection.

II. Research Questions

This paper seeks to examine the impact that the Alice curriculum has on male and female aptitudes and attitudes in a community college setting. The specific research questions are as follows:

1. Does exposure to the Alice curriculum support the improvement of both male and female students' skill levels and attitudes in an introductory programming course?

2. Does exposure to the Alice curriculum allow men and women to reach the same skill levels and attitudes by the end of an introductory programming course?

III. Methods

Data collection for this investigation took place during the fall 2004. .The three participating community colleges introduced the Alice curriculum and administered the assessment instruments. These instruments were validated for the study during the 2003-2004 academic year and this process has been discussed in prior papers.[5,6] The following sections describe the assessment instruments, participating schools, and the implementation process at each school.

A. Instruments

Three instruments were used for the purpose of this investigation. The first was a Demographics Survey, which all students completed at the beginning of their instruction with the Alice curriculum. This provided the study with information regarding students' gender and ethnicity, as well as the college and course in which each student was enrolled.

Improvement in student skill level was measured by means of the Concepts Exam. The Concepts Exam uses a multiple choice format, is delivered on-line, and is comprised of two sections, Basic Concepts and Alice Specific questions. Students' scores for each portion are calculated as the fraction of correct answers out of the total. The instrument was designed using guidelines for constructing quality multiple-choice questions.[7,8] The first section of the exam, Basic Concepts, covers concepts in introductory programming that are platform independent. In other words, this section is designed to measure the extent to which students develop algorithmic thinking and an understanding of programming concepts as a result of the given course. Algorithmic thinking is a desired outcome in first year programming course regardless of platform. The second section, Alice Specific, covers concepts specific to using the Alice software.

Changes in students' attitudes towards programming was measured by a computer attitude survey created by Loyd and Gressard [9].This survey, henceforth known as the Attitude Survey, was selected based on prior research that supports its validity for the age group of interest[10,11]. A concern with using this instrument is that it is more than ten years old and thus, has some dated questions. Unfortunately, there is not currently a more recent, validated survey available for these purposes. The Attitude Survey employs a Likert rating scale with selected responses ranging from Strongly Agree to Strongly Disagree. The survey was administered to students online and includes questions concerning attitudes towards computers and programming courses. Student responses were scored such that a higher score indicates a more positive attitude.

B. Implementation

Three community colleges, which in combination have a diverse student population, are participating in the JABRWOC project: Camden County College (CCC), Community College of Philadelphia (CCP), and

Tompkins Cortland Community College (TC3). At each institution, sections of introductory programming courses were selected to use the Alice curriculum. These course sections will henceforth be known as treatment groups. In some instances, different course sections or similar courses were chosen as control groups, in which students did not receive exposure to the Alice curriculum.

As is the case in many educational research investigations, it was not possible here to use a double blind, random experimental design.[12] Due to human subjects concerns and the practices of the participating school registrar's offices, students had to be informed as to the nature of the experiment and given a choice as to whether they would complete the treatment or control version of the course. Therefore, a Quasi-Experimental Design with Non-Equivalent Control Group Design, was used.[6] Furthermore, participating instructors read a set of prepared directions to the students before administration of the assessment instruments. These instructions informed students that a new curricular approach was being tested, and feedback concerning their attitude towards and knowledge of the subject was necessary. Not only did this comply with human subject requirements, but also it provided the students with a reason for completing the instruments.[5] Each treatment and control group completed the aforementioned assessment instruments.

Table 1 displays the different courses, the number of sections within each course and the duration of the Alice curriculum intervention. CCC implemented the Alice curriculum in a semester-long course, and designated another course as a control group. There were differences in the student population between these treatment and control groups. The students in the control group were computer science majors; the students in the treatment group were a mix of computer science and other majors. CCP chose two different courses in which to implement the curriculum. The first introduced a modified version of the Alice curriculum as a module in a course in which the primary focus was computer literacy. The module lasted for two to three weeks on average, and focused primarily on the tutorial sessions available in the Alice software. Certain sections of this course were designated as treatment, others as control. In addition, CCP integrated the Alice curriculum into an introductory programming course as a five-week module. No control groups were selected for this course. TC3 had only treatment groups, and the Alice curriculum was integrated into two separate courses as five-week modules. Variations in the manner in which the courses are offered allows for the comparison across levels of implementation, as will be discussed in the results section.

Table 1. Treatment and Control Courses

	Treatment	Control	Duration
CCC	CSC 105 – 7 sections	CSC 111 – 5 sections	Full semester
CCP	CIS 103 – 8 sections	CIS 103 – 5 sections	Two to three weeks
	CIS 106 – 6 sections	N/A	Five weeks
TC3	CSCI 160 – 2 sections	N/A	Five weeks
	CIS 108 – 4 sections	N/A	Five weeks

In each instance of treatment classrooms, the Demographics Survey, the Concepts Exam, and the Attitude Survey were completed at either the beginning of the semester or the beginning of the five week module. Control students completed these instruments at the beginning of the semester. Control students were not asked to complete the Alice Specific portion of the Concepts Exam, as they did not receive instruction with the Alice software. Data collected at the beginning of the semester or module will henceforth be known as the Pre Data. At either the end of the semester or the five week module, treatment students were again asked to complete the Concepts Exam and the Attitude Survey. Control students completed these instruments at the end of the semester. These data will henceforth be known as the Post Data. It is noted that the instruments used to collect the Pre Data and Post Data are identical. Only those students who have provided written consent to participate in this study are included in the datasets. [1]

C. Population and Sub-Categories

As this paper seeks to address the question of impact of the Alice curriculum on men and women, all statistical tests divide the population and sub-categories into male and female. Table 2 defines these categories.

Table 2. Population Subsets

Main Category	Sub-Category	Definition
Total Population	Women	Includes all participating students in the treatment group
	Men	
Computer Literacy	Women	Includes all participating students at CCP in CIS 103
	Men	
Semester Long Alice Course	Women	Includes all participating students at CCC
	Men	
Five Week Alice Courses – CCP & TC3	Women	Includes all participating students at TC3 and CCP CIS 106
	Men	
Five Week Alice Courses – TC3 Only	Women	Includes all participating students at TC3
	Men	

The Total Population was tested, and then separated by institutions. The Computer Literacy course data was examined separately, as this class is not meant to be programming intensive, and results were expected to differ. The Semester Long Alice course is also examined separately, so that we can determine if the length of exposure to the Alice curriculum is a factor in success. The Five Week Alice modules at CCP and TC3 are combined for an analysis, and finally all courses at TC3 are analyzed alone to determine if implementation techniques affected student performance.

IV. Results

This section begins with a discussion of the instrument response rate. Next, the results of the paired t-tests and the Analysis of Covariance (ANCOVA) are presented. The datasets were analyzed using the MINITAB statistical software with statistical significance defined to be $\alpha = 0.05$.

A. Response Rates

Community college courses in general suffer from high attrition rates.[1] Many students must commute to school and/or are working full-time, which may lead them to drop a course. Therefore, the study must take into account the fact that the dataset for Pre Data may be larger than that for the Post Data. Since dependence must be assumed between Pre and Post Data, only paired data is used in this study. Consistent data is defined here as students who completed the Demographics Survey and both Pre and Post versions of one assessment instrument. The total number of students across institutions that provided consistent data responses for treatment and control groups are listed in Table 3, separated by assessment instrument.

Table 3. Number of Students Across Institutions to Provide Consistent Data

	Treatment	Control
Attitude Survey	154	30
Concepts Exam	150	18

B. T-tests

As suggested by Table 3, the treatment responses provided the study with enough data to separate the total population into the subcategories defined in Table 2. Results will be discussed in terms of the Total Population, as well each subcategory. The response rate for the control data was not large enough to support statistical analysis within these subcategories. Therefore, the analysis that follows only examines these subcategories for the treatment group. Furthermore, when the datasets were not normally distributed results from the t-tests are not valid, and therefore statistical tests were not completed.[13]

In order to determine if men and women in each subgroup were improving either their programming knowledge or attitudes towards programming, paired t-tests were run on the consistent data. In each case, difference scores were determined by subtracting each treatment and control consistent data Pre score from the Post score. For both portions of the Concepts Exam, these mean differences were always positive, and therefore tested against the alternate hypothesis H_1: $\mu_{Difference} > 0$. Scores from the Attitude Survey differed slightly. In cases where the mean difference was greater than zero, we conducted a test similar to the one for the Concepts Exam. However, if the mean difference was negative, it was tested against the alternate hypothesis H_1: $\mu_{Difference} < 0$. P-values from these t-tests are listed in Table 4 and Table 5.

Table 4. Treatment Group: Concepts Exam and Attitude Survey

	p-value: Basic Concepts	p-value: Alice Specific	p-value: Attitude Survey
Total Population	0.000*	0.000*	0.467
Total Population – Women	0.000*	0.028*	0.188
Total Population – Men	0.000*	0.000*	0.197
Computer Literacy	0.007*	0.132	0.341
Computer Literacy – Women	0.058	0.283	0.178
Computer Literacy – Men	0.033*	***	0.246
Semester Long Alice	0.000*	0.000*	0.206
Semester Long Alice – Women	***	***	0.328
Semester Long Alice – Men	0.000*	0.000*	0.235
Five Week Alice – CCP & TC3	0.000*	0.000*	0.418
Five Week Alice – CCP & TC3 – Women	0.008*	***	0.355
Five Week Alice – CCP & TC3 – Men	0.000*	0.000*	0.489
Five Week Alice – TC3	0.000*	0.000*	0.268
Five Week Alice – TC3 Women	0.002*	0.044*	0.151
Five Week Alice – TC3 Men	0.000*	0.000*	0.416

*indicates significance at $\alpha = 0.05$

*** indicates non-normally distributed data

Table 5. Control Group: Concepts Exam and Attitude Survey

	p-value: Basic Concepts	p-value: Attitude Survey
Total Population	0.090	0.155
Total Population – Women	0.081	0.200
Total Population - Men	0.303	0.280

As these tables suggest, there was a statistically significant change for the Total Population from Pre to Post on both sections of the Concepts Exam for the treatment group, but not for the control group. This was further found to be true for males and females within the Total Population. Table 4 also suggests that in the Computer Literacy course at CCP, there was no statistically significant change in females' knowledge with respect to Basic Concepts or Alice Specific content. Also, within the Computer Literacy course at CCP, in general, no statistically significant change was found for the Alice Specific section. All other courses and populations in the treatment group displayed statistically significant positive changes from Pre to Post on the two sections of the Concepts Exam. With respect to student attitudes, neither treatment nor control students displayed a statistically significant change in attitudes toward programming over the semester. This result does not differ across gender.

C. Analysis of Covariance

Previous research suggests that women and men begin introductory programming courses with different skill levels and attitudes.[15,16] In order to examine whether women or men who begin with different average Pre scores acquire greater gains as a result of the treatment, an Analysis of Covariance (ANCOVA) was completed on the treatment group using men and women as factors. The ANCOVA adjusts for initial differences between men and women in the Pre Data and then tests for statistically significant difference in the adjusted Post scores.[14] Results for the ANCOVA test are displayed in Table 6.

Table 6. Treatment Group: ANCOVA for Concepts Exam and Attitude Survey

	Adjusted Post Score: Women	Adjusted Post Score: Men	F-value
Basic Concepts	0.408	0.492	5.003*
Alice Specific	0.220	0.332	10.110*
Attitude Survey	2.410	2.379	0.373

*indicates significance at $\alpha = 0.05$

The outcome of the ANCOVA for the Basic Concepts and Alice Specific portions of the Concepts Exam indicates that the adjusted Post scores for men are greater than those of women. This suggests that men are acquiring greater gains as a result of the intervention. With respect to attitudes, however, no statistically significant difference was found between males and females.

V. Conclusions

Based on the outcomes of the t-tests, the Alice curriculum appears to be an effective method for teaching basic programming concepts, as both treatment men and women enrolled in programming intensive courses using the Alice curriculum showed improvement with these skills after exposure to Alice. This result does not differ across the various classroom implementations. This is in contrast to the control groups who did not display improvement from Pre to Post assessment. As was discussed earlier, the basic concepts portion of the exam was designed to measure the platform independent understanding of programming concepts and algorithmic thinking. The development of algorithmic thinking skills is an anticipated student outcome in any first year programming course. Therefore, we would expect that both the treatment and control students would display statistically significant increases in this skill. However, the results presented here suggest that in the control group, there was no statistically significant increase in this important skill from pre to post assessment. Given the additional attention that is given to algorithmic thinking in the Alice curriculum, growth in this area was anticipated and found in the Alice classrooms.

All of the treatment subgroups with the exception of the Computer Literacy course displayed improvements in their content knowledge. The lack of improvement in Computer Literacy was not a

surprise, given that this course only has a minimal emphasis on programming and instructors chose to implement a modified version of the Alice curriculum. Results of the ANCOVA test suggest that both men and women are learning necessary skills with the Alice curriculum, although men develop these skills and proficiency with the Alice software more rapidly than women. A concern with this result is the attention that has been given to the development of the Alice curriculum with the purpose of appealing to women and of improving female skills in programming. Based on the described results, the Alice curriculum in 2004 did not level the playing field for women who are taking an introductory programming course. Although the achievement of women in the treatment group was greater than women in the control group, this was also true of the men. Women in the treatment group as compared to men were still at a disadvantage by the end of the course. Since the spring of 2005, the designers of the Alice curriculum have been working to improve the Alice curriculum, based on the results of this study. Future plans include a more in-depth qualitative analysis of the impact of Alice on male and female performances.

The results regarding student attitudes after exposure to the Alice curriculum are contradictory to traditional findings in programming courses. Research suggests that many students, women in particular, exhibit more negative attitudes toward the subject after an introductory course and this results in female attrition.[16] While treatment students do not display any improvement in attitudes, they also do not demonstrate a decline; the same is true for students in the control groups. In other words, neither the treatment nor the control curriculum negatively impacted students' attitudes. When comparing male and female attitudes, a surprising result was found. There is no evidence to suggest that treatment men and women are completing the introductory programming course with different attitudes. As with the t-tests for attitudes, this ANCOVA result conflicts with traditional findings that suggest women have a less positive attitude than men at the end of an introductory programming course.[15]

The Alice curriculum was developed to provide introductory programming students the basic skills necessary to succeed in a programming course, while at the same time improving student attitudes towards the subject. In particular, we were interested in whether the Alice curriculum was a means to end the disparity between men's and women's skill levels and attitudes at the conclusion of an introductory programming course. Results from the data analysis suggest that exposure to the curriculum allows both men and women in programming intensive courses to improve their skill levels, while maintaining their attitudes towards programming. However, while the Alice curriculum does appear to effectively teach introductory programming skills to both genders, the ANCOVA results suggest that the increase in knowledge is greater for men than for women. Concerning the implementation of the curriculum in programming intensive courses, neither the length of course nor specific institution appears to impact the students' level of learning. Students enrolled in courses whose primary focus is not programming did not show significant improvement in skill levels.

The completion of the data collection for the fall 2004 marks the midpoint for the three year JABRWOC study. As previously mentioned, the first year of the study was dedicated to formative assessment. During this time, instruments were selected, validated and then tested through a pilot implementation. The pilot study, which took place during the spring 2004, allowed us to make necessary adjustments to assessment instruments and administrative procedures.[5] The fall 2004 data collection is the first effort in summative assessment.

Based on the results of this study, there are several open questions that remain for future research. For example, it is currently unclear as to why a curriculum that is designed to appeal to women was not effective in increasing female skill levels to that of their male counterparts, even when adjustments were made for initial differences. It is further unclear as to why innovative software and curriculum designed to appeal to a multimedia generation maintained student attitudes but did not result in improved student attitudes with respect to computer science. Responding to these questions will provide necessary

information to the Alice research team that will support the refinement of the curriculum to better match our goals. As noted above, the designers of the Alice curriculum are currently working to determine ways to modify the Alice curriculum to further improve results with women. Future efforts also include the examination of the impact of the Alice curriculum in introductory computer science courses at four year colleges. Recently, the NSF funded the *Program Visualizations Using Virtual Worlds* project (PVUVW) (NSF DUE-0339734), which is designed for this purpose.[17] To determine the effectiveness of the curriculum at higher learning institutions, similar attitude and aptitude data are being collected. Efforts are also underway to examine whether the Alice software can be used to introduce middle school and high school students to programming concepts.

Acknowledgments:
We would like to thank Dr. William Navidi from the Colorado School of Mines for his assistance with the statistical analysis. We would also like to thank Dr. Tracy Camp from the Colorado School of Mines for her consultation regarding computer science education.

References

1. Cooper, S., Dann, W., & Moskal, B. *Java-Based Animation in Building viRtual Worlds for Object-oriented programming in Community colleges*. NSF-DUE-0302542.

2. Alice v2.ob Learn to Program Interactive 3D Graphics, http://www.alice.org (accessed December 2004)

3. Curricular Materials for Learning to Program with Alice: Please contact author.

4. Cooper, S., Dann, W., & Pausch, R. (2005) *Learning to Program with Alice Beta Version*. Prentice Hall.

5. Hutchinson, A., Moskal, B., Cooper, S., & Dann, W. *Formative Assessment: An Illustrative Example Using Alice.* 2005 ASEE Annual Conference & Exposition.

6. Gay, L.R. *Educational Research: Competencies for Analysis and Application (3rd Ed.).* New York: Macmillan Publishing Company, 1987.

7. Kehoe, J. Practical Assessment, Research and Evaluation, *Writing Multiple-Choice Test Items, 1995,* Online: http://www.edresearch.org/scripts/seget2.asp?db=ericft&want=http://www.edresearch.org/ericdb/ED398236.htm (accessed December 2004)

8. Frary, R.B. Practical Assessment, Research and Evaluation, *More Multiple-Choice Item Writing Do's and Don'ts, 1995,* Online: http://www.edresearch.org/scripts/seget2.asp?db=ericft&want=http://www.edresearch.org/ericdb/ED398238.htm (accessed December 2004)

9. Loyd, B.H. & Gressard, C.P. Computer Attitude Scale. *Journal of Computing Research,* 15(3), 241-259.

10. SearchERIC, http://searcheric.org (accessed October 2003)

11. ETS Test Link, http://www.ets.org/testcoll (accessed October 2003)

12. Olds, B., Moskal, B. & Miller, R. (2005). "Assessment in engineering education: Evolution and trends". *Journal of Engineering Education,* 94 (1), 13-25.

13. Navidi, W. *Statistics for Engineers and Scientists*. New York: McGraw-Hill, 2006.

14. Glass, G.V, Hopkins & K.D. *Statistical Methods in Education and Psychology.2nd edition*. Prentice Hall, Inc. Englewood Cliffs, New Jersey, 1984.

15. Treu, K., Skinner, A. (2002). "Ten Suggestions for a Gender Equitable CS Classroom". *SIGCSE Bulletin – Inroads Special Issue: Women and Computing,* 34(2), 165-167.

16. Duplantis, W., MacGregor, E., & Klawe, M., Ng, M. (2002). "'Virtual Family': An Approach to Introducing Java Programming". *SIGCSE Bulletin – Inroads Special Issue: Women and Computing,* 34(2), 40-43.

17. Cooper, S., Dann, W., & Moskal, B. *Program Visualizations Using Virtual Worlds.* (NSF-DUE-0339734).