Comparing Outcomes Across Different Contexts in CS1

Bruce A. Maxwell
Colby College
Waterville, ME 04901
bmaxwell@colby.edu

Stephanie R. Taylor
Colby College
Waterville, ME 04901
srtaylor@colby.edu

ABSTRACT

Context-based CS1 courses focusing on Media Computation, Robotics, Games, or Art have been shown to improve outcomes such as retention and gender balance, both important factors in CS education. Colby College has offered a Visual Media focused CS1 course since 2008, and in response to faculty and student feedback, we expanded our curriculum to include a second context-based CS1 course focused on Science applications. Our goal was to have completely different projects but teach the same fundamental concepts. In order to measure whether students in each version were learning the same concepts, and to reduce confounding factors, the same professors co-taught both versions of CS1 and students completed the same homework, quizzes, and final exam.

Our analysis of the quiz, final exam, and final overall performance showed no statistically significant difference by context or by gender. There was also no difference by context or gender in whether students took additional CS courses in the following two semesters. Furthermore, as a percentage of the students eligible to take the next offering of CS2, Data Structures and Algorithms, 48% of the students in these two offerings of CS1 registered for CS2, with no significant difference between contexts. Our conclusion is that we were successful in achieving similar outcomes, and the benefits of context-based CS1 courses, in both the Visual Media and Science versions of the course.

Keywords

Computational Thinking, CS 1, Context-based learning

1. INTRODUCTION

Context-based learning has been shown to be a good strategy for increasing both retention and gender diversity in introductory computer science courses [CS1][8]. The definition of context-based learning is typically characterized by a focus on the types of programming projects students undertake to gain hands-on experience with fundamental CS concepts. While the content and underlying concepts of these projects must also be part of lectures, the context is provided primarily by a single focus that ties the projects together and makes them meaningful to students.

SIGCSE '17, March 08 - 11, 2017, Seattle, WA, USA

ISBN 978-1-4503-4698-6/17/03...$15.00

DOI: http://dx.doi.org/10.1145/3017680.3017757

Guzdial introduced Media Computation, a popular context strategy, into CS1 and was able to show significant improvements in both retention and gender balance [6][7][8]. Rebelsky and colleagues showed that their Media Scripting variation also had benefits in retention, gender balance, and student confidence [14]. Other contexts evaluated in CS1 include robotics, games, science, biology, creative art, and other interdisciplinary topics [19][15][11][3][10][9][4][5][12][1][2]. The primary control in these studies has generally been a "traditional" CS1 course that, presumably, did not have a contextual focus to the projects or other course content.

While there is clear evidence of improved retention and gender balance, an important question for context-based CS1 courses is whether students are learning the same content as non-contextualized courses or courses with different contexts [8]. Tew at. al. looked at this problem and developed a standard assessment tool to help answer the question [18][16]. They found positive correlation with the FCS1 assessment tool and exams given in multiple different sections of CS1 that used different languages, but only one of those sections, a Media Computation section, was clearly in the context-based category [17].

The growth in contextualized courses is leading to more schools having multiple versions of CS1, each with a different context. For example, Union College developed multiple contexts for CS1 [2]. Harvey Mudd has developed interdisciplinary CS1 courses [4], and Bryn Mawr has offered multiple versions of CS1 using art, games, and robotics [19].

One question for programs offering multiple contextualized CS1 courses is whether the context matters for goals like retention, gender balance, and persistence (students taking more CS courses). Are students getting the same preparation in each course? Do the benefits in gender balance occur across different contexts? Do students take more CS courses at the same rate regardless of context? One reason these are difficult questions to answer is that there are many factors differentiating courses. Normal practice would be to have a separate professor teach each context, implying different pedagogical strategies, different personalities, different levels of teaching experience, different grading practices, and different styles of relationships with students. Separating these factors from the course context in any kind of statistical analysis would be extremely challenging.

Our CS program, situated in a small liberal arts college, has been using a Visual Media context for CS1 since the spring of 2008. Like other programs with contextualized CS1 courses, we observed benefits in retention, gender balance, and student engagement in the course compared to the prior CS1 course, which was not contextualized. In the fall of 2015 we took the opportunity to develop a new section of our CS1 course with a Science context. The new version was developed based, in part, on input from faculty across other disciplines, particularly the natural sciences and social sciences. An important goal in designing the new course was to ensure that students taking either CS1 version learned the same fundamental material and were prepared to continue in Computer Science, if they wished to do so. We also wanted to understand whether different context had differing effects on gender balance and persistence.

In addition to designing the new Science focused projects with the same fundamental concepts as the Visual Media projects, we decided to have Author 1 teach the lectures for both versions of the course, and to have Author 2 teach the lab sections for both versions of the course. Furthermore, we gave the students in both versions the same homework, quizzes, and final exam. Therefore, the only major difference between the two courses was the set of weekly projects and their context. Thus, we were able to directly examine whether these two different contexts–Visual Media versus Science–resulted in differences in student learning, gender balance, or enrollment in future CS courses.

2. EXPERIMENTAL DESIGN

As part of a decision to expand and vary our introductory CS curriculum, the two authors co-taught two sections of our CS1 course with differing contexts in the fall of 2015. Section A, Visual Media, was a version of the course we have taught since the spring of 2008. Section B, Science, was a new version we developed with input from faculty across disciplines regarding course context. The programming language for both sections was Python 2.7.

The Visual Media version of the course covers the fundamental concepts of programming and computer science, including: variables, functions, loops, conditionals, lists, stacks, classes, inheritance, polymorphism, recursion, references, grammars, and interpretation. There are eleven projects. The first three projects are based on turtle graphics, and students build up increasingly complex scenes using functions and hierarchical design. The next two projects involve pixel-based image manipulation using Zelle's graphics package [20]. Project six is an animation project, also using the Zelle package. Projects seven through eleven return to turtle graphics, and the students use L-system grammars and their own interpreter to create more complex images, finishing with a 3D turtle package developed in-house [13]. These five projects also introduce students to classes, recursion, and inheritance [**?**].

The Science version of the course covers all of the same fundamental topics except grammars and interpreters, ideas necessary for the L-system projects. Instead, those topics are replaced with file parsing, particularly CSV files, and a select set of more advanced Unix commands such as grep, cut, and curl. Just like the Visual Media section, the Science version has eleven projects. Projects one to three involve accessing live web data from a buoy on a local lake and calculating statistics or derived values from it. Projects four through seven deal with agent-based population modeling using penguins for project four and elephants for projects five to seven. Finally, projects eight through eleven build up a 2-D physics-based modeling system with gravity, collisions, rotation, and interaction, making use of the same Zelle graphics package as the Visual Media section. The 2-D geometry used in the Science section for physics simulation is very similar to the 2-D geometry used in the Visual Media section for drawing. However, the Science section has students implement position, velocity, an acceleration equations of motion for controlling objects.

Most of the time students spend on either section of the course is spent on the projects. Students are in lecture for 2.5 hours per week, and they are in lab, working on the projects, for 1.5 hours per week. In addition, we expect students to be working 5-10 hours per week outside of class on the projects. Therefore, having different projects for the two sections means that at least two-thirds of the time students spend working on course activities will be different when comparing the two sections.

In the fall of 2015, the two authors co-taught the two different sections. The course format was three 50 minute lecture sections (MWF) plus an 80 minute lab section. There were two lab sections per lecture section. 37 students completed the Visual Media focused course, and 44 students completed the Science focused course. Each week students completed a short homework assignment, a short quiz (20-minutes), and a programming project. The final exam was the equivalent of six quizzes. The projects and homework constituted 50% of a student's grade, with the quizzes (25%), final (20%) and participation (5%) making up the balance.

Author 1 gave the lectures for both sections and wrote the quizzes, homework, and final exam. Author 2 taught all of lab sections and co-designed the lab exercises with Author 1. Both the Visual Media and Science sections received the same quizzes, homework, and final exam. While the lectures had some project-specific content, the primary difference between the two sections was the weekly programming project. All of the course materials are available at the following URLs.

VM: http://cs.colby.edu/courses/F15/cs151
Sci: http://cs.colby.edu/courses/F15/cs151s

Our hypothesis was that the students were learning the same fundamental material in either section, regardless of the contextual focus of the projects, as measured by performance on quizzes, exams, and overall final grades. Our second hypothesis was that the lack of a difference was consistent across gender. Our third hypothesis was that there was no difference between the two sections in the number or gender distribution of students moving on to the subsequent Data Structures and Algorithms course.

3. ANALYSIS

To measure differences between the two courses, we analyzed twelve quizzes, the final exam, and the overall grade percentages. The quizzes and final exam made up 45% of the overall grade percentages. Each quiz and the final exam followed the same format. Students were given a piece of Python code and then asked a series of questions about the code and its functionality. In some cases, students had to describe the state of the computer–e.g. what variables were being maintained by the computer and what were their val-

Table 1: Quiz means and T-test results.

Quiz	VM Mean	Science Mean	p < X
1	18.35	18.64	0.550
2	17.16	17.30	0.787
3	17.14	17.43	0.627
4	14.19	16.25	0.043
5	17.41	17.73	0.699
6	16.24	15.84	0.584
7	17.32	18.48	0.125
8	16.11	16.25	0.865
9	17.35	16.98	0.508
10	14.54	15.00	0.595
11	18.27	18.05	0.721
12	17.54	17.45	0.870

Table 2: Final and overall grade means and T-test results.

Activity	VM Mean	Science Mean	p < X
Final	86.22	87.73	0.438
Overall	88.09	88.77	0.708

ues. The quizzes were graded on a 20 point scale by either Author 1 (last quiz) or one of two student graders based on a rubric. To minimize grader bias, halfway through the semester the students switched which section they graded. The final exams for both sections were graded on a 100 point scale by Author 1.

Table 1 shows the means and the two-tailed probability that the values are the same given Welch's T-test (unequal variances), with $N_{VM} = 37$ and $N_S = 44$. Looking at the data, only one quiz, week 4, shows shows a significant difference ($p < 0.043$) that the scores are different between the two sections. All other quiz results show no significant difference. It is also the case that one section does not dominate the other in score averages. The Science section has the high score in eight out of twelve weeks and the Visual Media section has the higher score in four out of twelve weeks. Assuming the probability of one section having the higher score is 0.5, there is a roughly 20% chance of one section having the higher score 4 or fewer times, another indication that there is no statistically significant difference in quiz scores between the two sections.

The final exam and overall grades show a similar result, as shown in Table 2. The Science section obtained slightly higher means in both categories, but the difference was not statistically significant.

We further analyzed the data by dividing the two sections by gender. There were 18 women (49%) and 19 men ((51%) in the Visual Media section and 16 women (36%) and 28 men (64%) in the Science section. We compared different genders within each section and the same gender between sections. There were only two comparisons that produced any close to significant differences. In the week 4 quiz, the women in the Science section earned a mean of 17.00 versus the 13.78 by the women in the Visual Media section ($p < 0.038$). The other difference was on the final exam where the women in the Science section earned a mean of 91.09 on the final exam and the men in the Science section earned a mean of 85.81 ($p < 0.016$).

Table 3: Number of students registered for Data Structures and Algorithms [DSA] or another CS course [OCS].

*Percents calculated relative to number of non-seniors [NS].

Group	#	# NS	DS&A	(%)*	OCS	(%)*
VM	37	30	14	47%	11	37%
Sci	44	34	17	50%	12	35%
VM Women	18	14	6	43%	5	36%
VM Men	19	16	8	50%	6	38%
Sci Women	16	10	6	60%	4	40%
Sci Men	28	24	11	46%	8	38%
Women	34	24	12	50%	9	38%
Men	47	40	17	43%	14	35%

Table 4: Registrations for Fall 2015, Spring 2016, and Fall 2016.

Course	#	Women	%
F15 VM	37	18	49%
F15 Science	44	16	36%
S16 VM A	33	15	45%
S16 VM B	26	9	35%
F16 VM	45	13	29%
F16 Science	37	13	35%
Total VM	141	55	39%
Total Sci	81	29	36%

The breakdown of quiz grades also showed that the women in the Science section earned the high quiz score for nine out of twelve weeks; the women in the Visual Media section earned the high quiz score for two out of twelve weeks, and the men in the Visual Media section earned the high score for one week. The women in the Science section also had the highest mean for the final (91.09) and overall grade (90.56). Except for the differences identified above, however, none of these differences were statistically significant, given the population sizes and variances.

Table 3 shows how many students took at least one additional course in computer science. It shows the number of students who have registered for the Data Structures and Algorithms [DS&A] course in Fall 2016 (the next possible offering after Fall 2015). It also shows the number of students who took a different CS course in either the January term or spring term of 2016. Overall, approximately half of the non-seniors in each section are registered for DS&A. While the percentages vary from 60% (Science section women) to 43% (VM section women), in all cases a change of one student more or less would bring the percentages to exactly 50%. Therefore, we see no significant difference either between the sections or between genders in the number of students taking more CS courses beyond the introductory course. It is worth noting that the women from the Fall 2015 semester both took more additional CS courses and registered for DS&A at a slightly higher rate than the men.

It may be worth noting that there were fewer women, as a percentage, in the fall 2015 Science version of the course (36% versus 49%). Looking at the single instance in the fall of 2015, this appears to be a real difference. Given a binomial distribution and assuming gender is balanced (50%), there is a 50% chance of having 18 or fewer women out of

37 (Visual Media section). However, given the same assumption of balanced interest, there is only a 5% chance of having 16 or fewer women out of 44 (Science). It is important, however, to look at other offerings of the Visual Media version of the course, and we can also examine registrations for both sections in the Fall of 2016. Table 4 shows the number of total students and the number and percentage of women in the fall 2015 (both versions offered), spring 2016 (only visual media), and fall 2016 (both versions offered) semesters. In the spring of 2016 we offered only the Visual Media version of the course, so results are shown for the two sections: VMA, and VM B. It turns out that the Fall 2015 Visual Media course had the highest ratio of women in the history of the course. By contrast, the percentage of women in the two versions of the course averaged across three semesters (39% for VM, and 36% for Science) is not significantly different when considered as either a binomial distribution or when executing a Student's T-test directly on the ratios ($p < 0.59$).

Prior discussions of contextual CS1 courses have also looked at retention, the number of students who register for a course and then do not complete it. The Visual Media version of the course had five students (12%) drop the course who had attended at least three lectures. Of those five, three dropped within the first two weeks (the regular add-drop period) and the other two dropped in the third week. This figure was similar to previous iterations of the Visual Media course. All other students completed the course with a passing grade. In the Science version, one student (2%) dropped the course within the first two weeks and all other students completed the course with a passing grade. These results are also similar to those reported in other studies of contextualized courses [8].

4. DISCUSSION

When we designed the new Science version of CS1, we intentionally designed the projects to match the concepts introduced in the Visual Media version, even though the content was very different. Our hope was to ensure that students learned the same fundamental concepts throughout the course. The analysis indicates that student performance on quizzes, exams, and overall course grade was not significantly different. It is important to note that having the students in both versions complete the same homework, quizzes, and final exam means that much of their learning process was identical. However, based on observation and self-reporting on course evaluations, the majority of student time invested in the course outside of class is spent on the projects. If there was a differential effect caused by different contexts on how well students learned the fundamental course concepts, we believe it would have appeared in the analysis.

Our second hypothesis, that the context had no impact on gender balance in CS1, appears to be substantiated as we gather more data. Certainly, the Science version gender balance is similar to our long-term average, and we still see fairly large fluctuations from semester to semester in the Visual Media course. However, this is an area where we need to continue to be aware of potential differences in the impact on gender balance of having different contexts.

As with all analyses of this type, we have to be careful about seeing an effect based on small numbers. Given discussions with students in the course, some of the choice of which section to pursue was due to scheduling conflicts, and there were some women in the Visual Media section who expressed a preference for the Science section. If just two women had switched sections from Visual Media to Science, the percentages would have been almost equal (43% versus 41%).

Finally, the data clearly show no significant difference in the numbers of students pursuing additional CS courses based either on context or gender. 44% of the students registered for the fall 2016 Data Structures and Algorithms course are women. Of all of the goals we have for CS1, having gender parity in subsequent courses is one of the most important outcomes.

Overall, the message of this analysis is that context matters, but primarily when compared with no-context. Which context you choose for a CS1 course is an important decision from the point of view of course preparation and the enthusiasm of the professor teaching the course. However, we have shown that assignments in two completely different contexts that make appropriate use of fundamental CS concepts and introduce them in an intentionally designed order can result in the same overall understanding of fundamental CS concepts and properly prepare students for subsequent CS courses.

Having a variety of contexts for CS1 gives students the ability to choose a context of direct interest to them. While designing the Science version of the course, we believed there would be students who would be excited about the content and might choose to take CS1 specifically because they felt it was relevant to them, where they might otherwise decline to take the course. What we forgot, was that there were also students excited by the Visual Media topics. While we have no data to back this up, the authors both felt more excitement from the students in both sections of the course because (most) students were able to intentionally choose which version they wanted to take. Thus, giving students a choice of contexts may increase student engagement (a possible future hypothesis to study).

5. SUMMARY

Context-based CS1 courses have been shown to improve outcomes such as retention and gender balance, both important factors in CS education. We were interested in creating multiple versions of contextualized CS1 courses in response to feedback from students and faculty, but we wanted to ensure students were learning the same fundamental concepts. We also wanted to ensure that we were seeing similar benefits in retention, gender balance, and persistence across different contexts.

Our experiment, based on giving Visual Media and Science focused versions of the course the same homework, quizzes, and final exam, showed that we did achieve statistically similar outcomes. In addition, students in each version, and across gender, took additional CS courses at similar rates. Therefore, we have confidence that, regardless of context, our CS1 students achieved similar levels of understanding of fundamental CS concepts and interest in pursuing further CS courses. In addition, we believe that offering students choices increased student engagement and made the classes more fun to teach.

6. REFERENCES

[1] V. Barr. Create two, three, many courses: An experiment in contextualized introductory computer science. *J. Comput. Sci. Coll.*, 27(6):19–25, June 2012.

[2] V. Barr. Disciplinary thinking, computational doing: Promoting interdisciplinary computing while transforming computer science enrollments. *ACM Inroads*, 7(2):48–57, May 2016.

[3] Z. Dodds, R. Libeskind-Hadas, and E. Bush. When cs 1 is biology 1: Crossdisciplinary collaboration as cs context. In *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '10, pages 219–223, New York, NY, USA, 2010. ACM.

[4] Z. Dodds, R. Libeskind-Hadas, and E. Bush. Bio1 as cs1: Evaluating a crossdisciplinary cs context. In *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '12, pages 268–272, New York, NY, USA, 2012. ACM.

[5] I. Greenberg, D. Kumar, and D. Xu. Creative coding and visual portfolios for cs1. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, SIGCSE '12, pages 247–252, New York, NY, USA, 2012. ACM.

[6] M. Guzdial. A media computation course for non-majors. In *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '03, pages 104–108, New York, NY, USA, 2003. ACM.

[7] M. Guzdial. Does contextualized computing education help? *ACM Inroads*, 1(4):4–6, Dec. 2010.

[8] M. Guzdial. Exploring hypotheses about media computation. In *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research*, ICER '13, pages 19–26, New York, NY, USA, 2013. ACM.

[9] S. Lee and E. Hildreth. An introductory computational course for science students. *J. Comput. Sci. Coll.*, 25(6):60–65, June 2010.

[10] S. A. Markham and K. N. King. Using personal robots in cs1: Experiences, outcomes, and attitudinal influences. In *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '10, pages 204–208, New York, NY, USA, 2010. ACM.

[11] P. Mawhorter, E. Shaver, Z. Koziol, and Z. Dodds. A tale of two platforms: Low-cost robotics in the cs curriculum. *J. Comput. Sci. Coll.*, 24(4):180–188, Apr. 2009.

[12] M. M. McGill. Learning to program with personal robots: Influences on student motivation. *Trans. Comput. Educ.*, 12(1):4:1–4:32, Mar. 2012.

[13] P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants*. Springer, 1990.

[14] S. A. Rebelsky, J. Davis, and J. Weinman. Building knowledge and confidence with mediascripting: A successful interdisciplinary approach to cs1. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, SIGCSE '13, pages 483–488, New York, NY, USA, 2013. ACM.

[15] J. Summet, D. Kumar, K. O'Hara, D. Walker, L. Ni, D. Blank, and T. Balch. Personalizing cs1 with robots. In *Proceedings of the 40th ACM Technical Symposium on Computer Science Education*, SIGCSE '09, pages 433–437, New York, NY, USA, 2009. ACM.

[16] A. E. Tew and M. Guzdial. Developing a validated assessment of fundamental cs1 concepts. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, SIGCSE '10, pages 97–101, New York, NY, USA, 2010. ACM.

[17] A. E. Tew and M. Guzdial. The fcs1: A language independent assessment of cs1 knowledge. In *Proceedings of the 42Nd ACM Technical Symposium on Computer Science Education*, SIGCSE '11, pages 111–116, New York, NY, USA, 2011. ACM.

[18] A. E. Tew, W. M. McCracken, and M. Guzdial. Impact of alternative introductory courses on programming concept understanding. In *Proceedings of the First International Workshop on Computing Education Research*, ICER '05, pages 25–35, New York, NY, USA, 2005. ACM.

[19] D. Xu, D. Blank, and D. Kumar. Games, robots, and robot games: Complementary contexts for introductory computing education. In *Proceedings of the 3rd International Conference on Game Development in Computer Science Education*, GDCSE '08, pages 66–70, New York, NY, USA, 2008. ACM.

[20] J. Zelle. *Python Programming: An Introduction to Computer Science*. Franklin, Beedle & Associates, 1st edition, 2003.