Levels of Active Learning in Programming Skill Acquisition: From Lecture to Active Learning Rooms

Patrick Seeling
Department of Computer Science
Central Michigan University
Mount Pleasant, MI 48859
pseeling@ieee.org

Jesse Eickholt
Department of Computer Science
Central Michigan University
Mount Pleasant, MI 48859
eickh1jl@cmich.edu

***Abstract*—In this paper, we perform a high-level comparison between different modifications of an entry-level programming course in computer science. The course iterations feature an online tutoring component paired with (i) textbook/traditional lecture/lab, as well as (ii) active learning in a specially designed active learning room, and (iii) active learning in a computer laboratory, both featuring an online textbook. We find that shifting the course format to active learning alone does not necessarily improve learner performance across all commonly graded course activities and might even result in lower performances in certain categories. Simply replacing traditional classroom settings with a laboratory in which to perform hands-on exercises might not be sufficient, as performance overall declined. Interestingly, we do not notice a very pronounced impact of lower performance and increased failing rates on student attitudes, as measured by student opinion scores.**

***Index Terms*—Computer science education; Integrated textbooks; Programming; Learner performance**

I. Introduction

STEM education has attracted significant research efforts in the recent past, predominantly focusing on an increase in the number of students successfully entering the workforce. One of the reasons behind this focus is the predicted shortage of adequately educated members of the workforce [1]. Computer Science as a discipline within STEM commonly encounters students from a variety of backgrounds [2]. The course under consideration here falls into this category, as it is offered to students campus-wide as a general education elective in addition to computer science and information technology students within the department. Introductory level courses in computer science can be particularly challenging, as new, domain-specific thinking, approaches, and tools have to be mastered within short time [3]. This initial skill acquisition phase can be challenging, and approaches to alleviate the cognitive load [4] could be effective in increasing retention rates. Within cognitive load theory, learners can be aided through worked-out examples, which provide an opportunity of assimilating the underlying problem solution schemata before having to apply them. In turn, following the stages of problem formulation, finding solution steps, and formulating the final solution, different strategies can be applied to help learners. One example avenue to aid learners is to follow a more interactive instructional approach [5], [6], [7]. However, it is difficult to achieve this environment of timely feedback, especially for larger class sizes [8]. This course augmenting strategy leads, ultimately, to active learning environments [9], such as reviewed in, e.g., [10]. These, however, only apply to the time spent in class, while the road to subject mastery typically requires a significant amount of additional training outside of the classroom. As individualized tutoring is commonly not a feasible option due to the number of learners and tutors available, intelligent tutoring systems have emerged as a potential remedy for this problem [11], [12]. While software-driven online tutoring systems have become available, active learning environments generally require significant remodeling to enable students to work in smaller groups on specific problem sets.

The contribution of this paper is the evaluation of an introductory course in computer programming that compares (i) traditional, (ii) active learning, and (iii) computer laboratory environments with respect to academic achievement and student attitude outcomes. In the following section, we review the course under consideration and the differences between offerings. We describe the impacts of these on academic achievements and attitudes in Section III for the students in general, as well as by high and low effort groups. We conclude in Section IV.

II. Course Overview

The Department of Computer Science at Central Michigan University has major programs in Computer Science and Information Technology, for which this course represents the initial requirement. Furthermore, the course under consideration is offered as a general education elective open to all students on campus.[1] The programming language employed is JAVA

[1] We note that due to the freshman placement of this course, it is generally not possible to differentiate majors within the student population, as many remain undeclared.

(imperatives-first) and typically ends at the complexity level of multi-dimensional arrays. A reference textbook utilized is [13], with content selected to fit the overall instructional approach.

The Spring 2015 offering represents a traditional lecture-based format complemented by additional laboratory sessions with guided hands-on exercises. The Fall 2015 course iteration abandoned this separation and the class sessions were held in a specially remodeled active learning classroom where students had to bring their own device (BYOD). Short instructional content presentations at session beginnings were followed by small student group hands-on exercises and intermittent class discussions to enable all groups to achieve the same level of exercise completion. The Fall 2016 course iteration followed this approach, however the class sessions were entirely held in a computer laboratory classroom with one dedicated computer for each learner. The arrangement in the room, however, was row-based and did not facilitate grouping in a straight-forward fashion compared to the individual tables in the active learning classroom. The hands-on exercise content was kept as similar as possible, making the different iterations comparable.

The Spring 2015 semester offering featured a physical text-book requirement, which included access to the publisher's on-line tutoring environment, MyProgrammingLab (MPL). MPL enables students to practice at their own pace and receive feedback for typical online quiz questions, but also for short programs. Submitted snippets of code are evaluated and real-time feedback is provided to the individual learner. The other two course iterations featured an integrated environment by the same publisher (REVEL), which combines an online version of the textbook with in-line hands-on exercises that match those of the MPL environment.

The course's main grading categories included mid-terms and one final course examination (each consisting of three individual parts covering underlying concepts, code understanding, and code writing). Chapter-based online quizzes (underlying concepts and code understanding commonly examined with multiple-choice questions and two tries per quiz), programming homework assignments (employing a grading rubric communicated in the syllabus, typically with a one week completion time), and exercises in the MPL/REVEL environments (covering underlying concepts, code understanding, and code writing, with unlimited tries) were assigned at the end of each corresponding chapter.

III. Results

All learner evaluations were re-scaled to relative achievement scores to enable direct comparisons. We initially present the overall course characteristics in Table I. We note that in Spring 2015 a large lecture was flanked by several smaller hands-on laboratories, while in Fall 2015 the capacity for the active learning room was the limiting factor. A single section was offered by the instructor in Fall 2016. The overall course grade average (GPA) was fairly comparable for the initial two offerings, with the Fall 2016 one showing a significantly lower average GPA. We separate the students in each offering by the median course GPA into top and low performers. We observe that the top performers in the first two iterations are highly comparable, whereas the top performers in the last iteration attained a lower score. This trend is worsened when considering the low achieving students, which in this final iteration of the course ranked only at a D- level. For the switch from a traditional class to an active learning format, however, we notice a general achievement increase, which leads to the postulation that the active learning format could be of high value especially to lower performing students.

TABLE I
Overview of course iterations and student performances.

Item	Spring 2015 (Traditional)	Fall 2015 (Active Learn.)	Fall 2016 (Laboratory)
Enrollment	69	42	21
Avg. GPA	2.72 (B-)	2.69 (B/B-)	2.14 (C+/C)
Top 50%	3.81 (A/A-)	3.82 (A/A-)	3.5 (A-/B+)
Low 50%	1.66 (C-)	1.76 (C-)	0.77 (D-)
SOS Resp.	42	22	13
Exam (M, SD)	0.593, 0.269	0.699, 0.236	0.658, 0.296
Progr. (M, SD)	0.783, 0.197	0.735, 0.237	0.842, 0.148
Quiz (M, SD)	0.740, 0.144	0.640, 0.154	0.681, 0.172
Tutor (M, SD)	0.869, 0.172	0.827, 0.158	0.739, 0.156
Course (M, SD)	0.845, 0.170	0.850, 0.171	0.746, 0.203

A. *Detailed Differences Between Offerings*

Levenes evaluation of the homogeneity of variances indicates no significant differences for the course performances in Table I. In turn, we consider ANOVA with Tukey's honestly significant difference (HSD) post-hoc test to compare the student performances across semesters. We find that the relative quiz and online tutor performances exhibit significant differences, while the overall course performance approaches significance between the different course modifications. Specifically, we note that the overall significant difference [$F(2, 136)$=6.085, p=.003] can be further evaluated by semester. The almost 10% difference for the quiz performance between the traditional and active learning format offerings is significant (based on Tukey HSD with p=.002). Furthermore, we find that the online tutoring system usage differences between the traditional class setting and the hands-on only setting were significant as well with post-test p=.007 [$F(2, 136)$=4.85, p=.009]. The result is that the ANOVA reveals an approach of significance for the total course performance, $F(2, 136)$=2.743, p=.068. A further investigation employing the post-hoc test reveals that both the traditional course offering (p=.072) and the active learning format offering (p=.083) approach significance here, with an approximate 10% drop in overall course performance for the hands-on only course revision.

Overall, we note that the shift from the traditional format to the active learning format results in an overall mixed result set. While exam performance significantly increased, other course evaluations yielded lower attainments of subject mastery levels. However, overall course performance remained comparable on average. Replacing the active learning environment with a regular computerized room and performing similar activities seems to again increase certain course evaluation performances, but approaches an overall significant drop in total semester performance. Considering that these are averages per course offering across learners, we now shift the view to different learner populations within the course offerings.

B. Evaluation by Student Persistence and Course Offering

We now focus on the additional viewpoint of different student achievements employing the online tutoring component utilization as indicator. The average tutor utilization continuously dropped from the traditional to the active learning, and further to the hands-on only course revisions. This particular course component required 80% of assigned problems to be completed, but with unlimited repetitions. In turn, students that might not have achieved subject mastery through the different other activities had sufficient opportunity to correct misconceptions and achieve the required assignment score needed. We split the students in the individual course offerings into two separate groups based on the median of the relative online problem completions in each respective offering. We present the overall means (M) and standard deviations (SD) in Table II. With significant deviations in the variances between groups, throughout this comparison we employ ANOVA paired with a Games-Howell post-hoc test to accommodate the different group variabilities.

We initially notice a significant difference in exam performance by persistence, $F(5,133)$=3.299, p=.008. Particularly, the post-hoc test reveals the low and high group exam scores from the standard course offering are significantly different from the high performing students from the laboratory only offering (p=.005 and p=.035, respectively). Interestingly, the large difference between the high and low effort students in the laboratory offering merely approaches significance (p=.053).

Evaluation of the relative programming performance has shown significant differences between groups as well, $F(5,133)$=4.529, p=.001. Students demonstrating a lower level of persistence in the traditional setting exhibit differences to their higher performing peers (p=.004) as well as to the high performing students in the hands-on only offering ($p <$.001), but not to those in the active learning class. Interestingly, the learners in both more interactive classes do not exhibit significant differences by student determination other than low-high interactions between both offerings. While the difference between low performers is not significant at all (p=.988), that for high persistence approaches significance (p=.079). However, rubric-based grading has to be regarded carefully, as grader bias and differences can have a significant impact on these values.

The administered quizzes similarly exhibited significant differences amongst the groups by online tutor efforts, $F(5,133)$=13.429, $p <$.001. Within each of the semesters offered, significant interactions were found between the low and high effort learners. For the active learning student group, we additionally found significant interactions between low performers and high performers in the traditional setting ($p <$.001) and the laboratory setting ($p <$.001) as well as their peer group (p=.018).

Lastly, the course performance differences based on student efforts, as indicated by their online tutoring completion level, exhibits significant differences as well, $F(5,133)$=12.727, $p <$.001. While we find interactions between the traditional settings' two groups ($p <$.001) and the laboratory setting's two groups (p=.026), we only notice an approach of significant difference within the active learning group (p=.064). While additional cross-interactions between low-high effort student groups exist, such as for high performers in both the traditional and active learning class with their lower effort peers in the other course configurations, they fall into line with our earlier observations and are omitted here due to space constraints.

Overall, these individual findings corroborate earlier observations that student effort, here measured by completion of the online tutoring assignments with unlimited retries, has a significant impact on the individual and overall performances. Moreover, we note that as a whole, the active learning approach ultimately resulted in non-significant differences amongst these two learner groups, which could mean that such approach can effectively help the otherwise struggling students to perform better, especially in specific course evaluation categories. The similar laboratory environment, on the other hand, did not yield those benefits and student performances effectively bottomed for low performing students.

C. Impact on Student Attitudes

Given these overall differences in course performances, we note that the traditional offering resulted in a fail rate (as given by D and E semester grades) of 13%, the active learning class resulted in 15%, whereas the hands-on only offering resulted in 32%. Table III provides the student feedback at the end of each semester, administered as Student Opinion Surveys (SOS)[2]. We initially notice that the overall instructor effectiveness found wide approval, but the disapproval was highest for Fall 2016. This indicates that the overall student experience in the course (with several failing students) might very well be reflected in the overall attitudes towards the course. Overall, students agreed that the instructor was enthusiastic about the course, with the Fall 2016 group having the highest level of disagreement. We notice a difference in the negative view on instructor preparation for the Fall 2016, however on a low level. Interestingly, learners in all three semesters seem to have the same general view on the effectiveness of course material presentation, with only the active learning Fall 2015 configuration having a larger number of agree over strongly agree. The overarching trend continues for the course organization. However, the traditional lecture has a significant fraction of students that are neutral, while the active learning class exhibits the highest agree responses. Considering that

[2]Due to the anonymous nature of the SOS mechanism, the evaluation results illustrated in Table III can only be qualitatively evaluated.

TABLE II
LEARNER PERFORMANCE BY PERSISTENCE AS DETERMINED THROUGH ONLINE EXERCISE COMPLETION.

Group		Exam		Progr.		Quiz		Tutor		Course	
		M	SD	M	SD	M	SD	M	SD	M	SD
Traditional	Low	0.568	0.256	0.702	0.225	0.669	0.152	0.759	0.189	0.751	0.181
Sp. '15	High	0.617	0.283	0.861	0.125	0.809	0.094	0.976	0.018	0.936	0.092
Active learning	Low	0.658	0.226	0.702	0.232	0.566	0.115	0.716	0.161	0.777	0.165
F '15	High	0.738	0.245	0.767	0.243	0.711	0.155	0.933	0.034	0.920	0.150
Laboratory	Low	0.463	0.298	0.749	0.166	0.560	0.169	0.606	0.122	0.602	0.196
F '16	High	0.833	0.157	0.926	0.050	0.789	0.080	0.858	0.044	0.875	0.095

TABLE III
OVERVIEW OF ANONYMOUS STUDENT OPINION SCORE FEEDBACK AT THE END OF RESPECTIVE COURSE IMPLEMENTATION SEMESTERS.

Item	Sem.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Instructor's teaching	F. 16	0.154	0.538	0.077	0.077	0.154
helped me	F. 15	0.227	0.318	0.318	0.091	0.045
learn	S. 15	0.310	0.310	0.190	0.167	0.024
Treated students	F. 16	0.692	0.154	0.077	0.000	0.077
with respect	F. 15	0.773	0.182	0.000	0.045	0.000
	F. 15	0.667	0.214	0.071	0.048	0.000
Accessible	F. 16	0.462	0.385	0.077	0.000	0.077
to students	F. 15	0.591	0.364	0.045	0.000	0.000
	S. 15	0.524	0.238	0.214	0.024	0.000
Organized	F. 16	0.462	0.231	0.154	0.077	0.077
course well	F. 15	0.227	0.500	0.136	0.091	0.045
	S. 15	0.405	0.262	0.286	0.048	0.000
Presented course	F. 16	0.308	0.308	0.231	0.000	0.154
material well	F. 15	0.182	0.409	0.273	0.136	0.000
	S. 15	0.333	0.262	0.238	0.143	0.024
Seemed well	F. 16	0.462	0.385	0.077	0.000	0.077
prepared	F. 15	0.500	0.409	0.091	0.000	0.000
	S. 15	0.452	0.238	0.286	0.024	0.000
Was enthusiastic	F. 16	0.615	0.308	0.000	0.000	0.077
about subject	F. 15	0.773	0.136	0.045	0.045	0.000
	S. 15	0.667	0.238	0.071	0.024	0.000
Overall instructor	F. 16	0.417	0.250	0.167	0.083	0.083
effectiveness	F. 15	0.318	0.455	0.182	0.045	0.000
	S. 15	0.310	0.357	0.286	0.048	0.000

potentially access to the instructor could pose a significant challenge, we observe in Table III that this is not the case, as students agreed to high levels of accessibility. Similarly, potential other differences could be the case, which, however, do not find reflection in the student view of respect from the instructor, which they agree to in broad majority, with the active learning format again providing the highest level of student agreement.

The evaluation of the teaching and impact on learning by the instructor in these cases inherit the individual classroom configurations and could be seen as an additional indicator for best practices. Interestingly, the majority of students in the traditional course offering and the hands-on laboratory offering were both rating the teaching highest for their own benefit. It is interesting that students in the Fall 2016 offering rated the instructor teaching highly, whereas the better performing course iterations rated it lower. One particular reason could be that especially in the active learning classroom the overall format was new for a large fraction of learners, and they did not know how to evaluate this particular configuration, reflected in the undecided category of survey answers.

IV. CONCLUSION

Compared to traditional as well as hands-on only laboratory utilization with the same course features results in overall strong support for active learning classroom utilizations. Learners exhibit at least comparable levels of overall performance, but the active format seems to increase the overall achievement of lower performing students.

In currently ongoing works, we are extending on the modalities of course offerings by including an open-source approach to active learning classroom conversions, Practical Active Learning Stations (PALS). This approach would have enabled the laboratory course offering to convert to an active learning format, with all potential benefits.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1608043.

References

[1] T. Beaubouef and J. Mason, "Why the high attrition rate for computer science students - some thoughts and observations." *SIGCSE Bulletin ()*, vol. 37, no. 2, p. 103, 2005.

[2] M. Rizvi and T. Humphries, "A Scratch-based CS0 course for at-risk computer science majors," in *2012 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2012, pp. 1–5.

[3] L. Porter, M. Guzdial, C. McDowell, and B. Simon, "Success in introductory programming," *Communications of the Acm*, vol. 56, no. 8, pp. 34–36, Aug. 2013.

[4] J. Sweller, "Cognitive Load Theory." Elsevier, 2011, pp. 37–76.

[5] B. Simon, J. Parris, and J. Spacco, "How we teach impacts student learning: peer instruction vs. lecture in cs0," in *Prof. of ACM SIGCSE*. Denver, CO, USA: ACM, Mar. 2013, pp. 41–46.

[6] M. Lou Maher, C. Latulipe, H. Lipford, and A. Rorrer, "Flipped classroom strategies for cs education," in *Proc. of ACM Technical Symposium on Computer Science Education (SIGCSE)*, Kansas City, MO, USA, Feb. 2015, pp. 218–223.

[7] M. N. Giannakos, J. Krogstie, and N. Chrisochoides, "Reviewing the flipped classroom research: reflections for computer science education," in *Proc. of the Computer Science Education Research Conference (CSERC)*. Berlin, Germany: ACM, Nov. 2014, pp. 23–29.

[8] C. Ott, A. Robins, and K. Shephard, "Translating Principles of Effective Feedback for Students into the CS1 Context," *ACM Transactions on Computing Education (TOCE)*, vol. 16, no. 1, pp. 1–27, Jan. 2016.

[9] J. L. Jensen, T. A. Kummer, and P. D. d. M. Godoy, "Improvements from a Flipped Classroom May Simply Be the Fruits of Active Learning," *CBE-Life Sciences …*, vol. 14, no. 1, pp. ar5–ar5, Feb. 2015.

[10] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, "Active learning increases student performance in science, engineering, and mathematics," *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8410–8415, Jun. 2014.

[11] K. Buffardi and S. H. Edwards, "Responses to adaptive feedback for software testing," in *Proc. of the Conference on Innovation & Technology in Computer Science Education (ITiCSE)*, Virginia Tech. Uppsala, Sweden: ACM, Jun. 2014, pp. 175–170.

[12] A. N. Kumar and L. C. Kaczmarczyk, "Programming tutors, practiced concepts, and demographics," in *Proc. of the ASEE/IEEE Frontiers in Education Conference (FIE)*, Oklahoma City, OK, USA, Nov. 2013, pp. 773–778.

[13] Y. D. Liang, *Intro to Java Programming*. Pearson Education Limited, 2014.