Teach Students to Study Using Quizzes, Study Behavior Visualization, and Reflection: A Case Study in an Introduction to Programming Course

Marcia C. Moraes*
Computer Science Department, Colorado State University
marcia.moraes@colostate.edu

Albert Lionelle
Khoury College of Computer Science, Northeastern University
a.lionelle@northeastern.edu

Sudipto Ghosh
Computer Science Department, Colorado State University
sudipto.ghosh@colostate.edu

James E. Folkestad
School of Education, Colorado State University
james.folkestad@colostate.edu

ABSTRACT

Due to a long history of using rote memorization and rereading as the primary means to study, students are coming to the University with misconceptions about study strategies that are beneficial for their performance and long-term learning. Techniques such as spaced retrieval practice, interleaving, and metacognition are proven by cognitive and educational researchers as strategies that greatly improve learning. They focus on helping students to own responsibility for their learning and retention of information. Considering their benefits, quizzes were re-branded to be formative low-stakes retrieval practice activities (RPAs) in an Introduction to Programming Course (CS1), meaning that students would use the quizzes as learning tools, testing themselves in a spaced and interleaved manner as many times as they want during the semester. Additionally, the U-Behavior learning and teaching method was used. This method applies visualizations of student's study habits and self-reflections to help students to be aware of their study practices, reflect on them, and change their study routine to improve performance and long-term learning. Study behaviors were analyzed and the final Canvas exam, final coding exam, and final course grades were compared for students who spaced and interleaved their practice with students who did not. Results showed a statistically significant increase in all grades evaluated for students who practiced using this novel combination of spacing and interleaving integrated with U-Behavior visualizations and RPA reflection activities for learning.

CCS CONCEPTS

• **Social and professional topics**; • **Computing Education**;

KEYWORDS

Study behavior visualizations, spaced practice, interleaving, CS1

*Corresponding author: marcia.moraes@colostate.edu

ICETC 2023, September 26–28, 2023, Barcelona, Spain

ACM ISBN 979-8-4007-0911-1/23/09.
https://doi.org/10.1145/3629296.3629362

ACM Reference Format:
Marcia C. Moraes, Albert Lionelle, Sudipto Ghosh, and James E. Folkestad. 2023. Teach Students to Study Using Quizzes, Study Behavior Visualization, and Reflection: A Case Study in an Introduction to Programming Course. In *The 15th International Conference on Education Technology and Computers (ICETC 2023), September 26–28, 2023, Barcelona, Spain.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3629296.3629362

1 INTRODUCTION

Introductory Computer Science courses are known to have high failure rates [1–3]. The most frequent reasons to fail CS1 were found to be a lack of time and motivation, frustration, ineffective study strategies, and course prioritization conflicts [4–6]. Moreover, students mentioned the need for using new techniques for learning instead of just rereading their review notes and lectures, such as more hands-on practice, and increasing their efforts in order to succeed in computer science.

Several studies [5, 7, 8] point out that students arrive at the university without taking responsibility for their own learning. They also come with unrealistic views about the number of hours needed for each course, and have misconceptions about study strategies, such are rereading, memorization, and cramming before exams, which they mistakenly think are beneficial for their learning. Research in learning sciences, both in laboratory [9–11] and classroom settings [12–15], have shown the benefits of testing as a learning tool to improve students' performance and long-term retention. To be effectively used, testing needs to occur as a retrieval practice (via self-testing) activity that is spaced and interleaved [16]. Instead of massing practice, often called "cramming", students should study for shorter intervals spaced out over time, and they should perform interleaving, i.e., intermix topics when studying. Also, testing needs to be understood as a formative assessment [9] where students can learn from their mistakes. According to Roediger III and Pyc [16] testing can be easily implemented in the classroom as frequent low-stakes quizzes.

Considering these benefits, the U-Behavior learning and teaching method [17, 18], which is a combination of spaced and interleaved practice with visualizations and reflections of such practices, was integrated in the CS1 course to encourage students to be aware of their study habits, reflect on them, and change their study routine to improve performance and long-term learning. U-Behavior method emphasizes spaced and mixed practice as a learning tool, so CS1 quizzes were re-branded to be formative low-stakes Retrieval Practice Activities (RPAs). Unlike previous work in Computer Science education [6] that evaluated the use of ungraded quizzes in students' performance, the goal of this case study was to use U-Behavior, visualizations of study habits, and self-reflection to support self-regulated learning [9], and respond to the following research questions.

- Q1: How do students practice their RPAs during the semester? Do they change their study behaviors after being made aware of undesirable behaviors and reflecting on them?
- Q2: Do students who demonstrate desirable study behaviors improve their performance on final Canvas exam, final coding exam, and final course grade compared with students who did not change their behaviors?

2 RELATED WORKS

Lyle et al. [13] examined how retention of classroom content is affected by increasing the amount and spacing of retrieval practice in a precalculus course for engineering students. Their findings support the use of spacing practice in real-world mathematics education. Hartwig and Malain [11] examined students' decisions about spacing their study in an undergraduate course and whether self-selected spacing predicted course performance. Study could be done by spaced retrieval practice (via quiz tool) or spaced reading (via textbook). The authors found that, even controlling for time spent quizzing, when students opted for more frequent and spaced quizzing instead of spaced reading, they earned higher grades on exams. Similar results were found also by Harindranathan and Folkestad [12].

O'Malley and Aggarwal [6] analyzed how students' engagement with ungraded quizzes, done as retrieval practice in an introductory programming course, contributed toward students' better performance on final exams. YeckehZaare et al. [14] developed a learner-centered retrieval practice tool for an introductory Python course that spaces practice and interleaves topics. They also designed a grading scheme to incentivize students to space their practice over the semester instead of cramming around exam's due dates. The authors found that 32% of the students used the tool more than the required number of times to gain full points, and use of the tool correlated with higher exam grades. In another work, YeckehZaare et al. [15] proposed the use of retrieval practice and spacing in an introductory programming classroom and called this technique "retrieval-based teaching." The strategy consisted of encouraging active retrieval by asking ungraded rapid-fire questions, where each question moves from one student to another until the correct answer is reached. The instructor provided feedback for every incorrect answer and only presented the correct answer if students could not find it collaboratively. Results showed that students who were taught with this strategy earned an average of 2.36 percentage points higher in course grades compared with students who were not exposed to this strategy.

Interleaved practice has been demonstrated to improve learning compared with massive practice in several different context [16]. Lionelle et al. [19] designed a spiral CS1 curriculum which consisted of spacing and interleaving the teaching of CS1 topics. Results shown that students that used the spiral model outperformed students who learned via traditional method. In order to use the same concepts, but in a next course, Lionelle et al. [20] proposed one-credit "booster" course to students taking CS2, the course aimed to teach students study techniques such as spacing and interleaving. In another study, Lionelle et al. [21] designed a formative/summative grading system for CS0 and CS1 classes to support a structured growth mindset where formative assignments were composed by low-stake quizzes and labs.

U-Behavior [17, 18] is a method of teaching and learning that reimagines and redesigns online quizzes as Retrieval Practice Activities (RPAs), which are quizzes designed for learning practice. Specifically, students learn how to apply the power of testing, distributed practice, retrieval practice, and interleaved practice. U-Behavior provides online instruction about why to practice (study) using this technique. The method reinforces this instruction with personalized visual-form learning analytics that allow students to view their learning behavior and reflect upon it. This visualization is then paired with ongoing guidance and coaching to improve behavior over time [18]. U-Behavior automatically generates feedback containing guidance and coaching, which is presented to students when they are asked to reflect about their study behavior.

Unlike previous work [6, 14, 15, 21] which applied spacing and interleaving in CS courses, this proposal intends to use the U-Behavior teaching and learning method by incorporating RPA reflection assignments in an introductory computer science curriculum based on the spiral model [19] that uses formative quizzes [9, 16]. The RPA assignments make use of spacing and interleaving strategies as well as study habit visualizations and self-reflections that aspire to promote change in student's study behaviors. Different from [20] who aimed to have a CS2 course to teach students study strategies such as space and interleave, this work aims to have those strategies embedded in the CS1 curriculum itself as RPA assignments, that teaches students about effective study behaviors by applying self-reflection based on student's study behaviors visualizations.

3 METHODS

This study was conducted in a CS1: Java for prior programming experience group as part of a Computer Science Undergraduate program at a Research 1, land-grant university, in the Fall 2021 semester. A convenience sample of 88 students self-registered for the course with 75 consenting to be included in the research. The course was taught in-person, implemented in 16 weeks, and organized in the form of a spiral curriculum proposed by [19] with three lectures and two labs per week. The instructor of the course is part of the research team and already implemented U-Behavior in other courses. Canvas was the Learning Management System (LMS) used. The introduction module presented the overview of the course as well as the explanation about the research and the consent survey. The explanation consisted of (1) a short animated interactive video describing the benefits of retrieval practices activities for long-term learning and student performance and the U-Behavior App, (2) a quick survey about the video, and (3) an RPA guide document. The guide explained the four levels of practice behaviors that students would use to self-reflect upon their practice. The levels are based on the following two behavior scores that are presented to the students when they generate their RPA graphs using the U-Behavior App, which was available in the course's navigation bar.

- Behavior Score 1 - RPAs practiced on three different days: Represents the percentage of RPAs that the student practiced on three different days out of the total number of RPAs available in the course.
- Behavior Score 2 - RPAs Mixed (practice): Represents the percentage of RPAs that the student used in a mixed/interleaved manner out of the total number of RPAs available in the course.

Based on these behavior scores, students were presented the following four levels of practice behaviors:

- Level 4 - Highly Effective Practice Behavior: Students at this level practiced at least 70% of the RPAs on three different days, based on Behavior Score 1, and also interleaved their practice between RPAs (at least 40% on Behavior Score 2).
- Level 3 - Effective Practice Behavior: Students at this level practiced at least 70% of the RPAs on three different days, but interleaved practice between fewer than 40% of the RPAs.
- Level 2 - Less Desirable Practice Behavior: Students at this level practiced between 40-69% of the RPAs on three different days. Behavior Score 2 is not considered in this level.
- Level 1 - Low Practice Behavior: Students at this level practices less than 40% of the RPAs on three different days. Behavior Score 2 is not considered in this level.

Students completed eleven RPAs introduced as low-stakes formative learning-based quizzes, which they had the option to take an unlimited number of times and retain their highest score. The U-Behavior App was available throughout the semester, and students could access it at any point to view their RPA graphs.

One week before the mid-term, students were asked to complete the following reflection assignment:

- Review RPA graph, download and submit: Students were asked to access the U-Behavior App on Canvas to see their score and visualization of their practice behaviors regarding their RPAs and download a pdf file containing that graph submitting to their instructor.
- Review feedback on your practice behavior: Students were presented with four different RPA graphs representing undesirable practice behavior (minimal practice), good practice behavior, massing practice around due dates, and optimal practice behavior, and needed to select the one that looked the most similar to their own RPA graph. Students were reminded that there was no single correct answer but that they should be as accurate as possible in order to receive customized feedback. The feedback was customized for the kind of behavior selected by the student.
- Planning your future practice: Students were asked to respond to the following reflection question: "After reviewing the feedback on your RPA practice behavior (in Step 2), do you plan to change how you use the RPAs in the upcoming weeks (please describe why and how)?".

This reflection assignment was purposely done one week before mid-term to provide students with an opportunity to visualize their study behaviors, analyze their study behaviors, and plan future actions. By making students self-reflect about their own study behaviors, students were allowed to self-regulate and work towards study practices that would boost their learning and performance. The idea was that students would use the customized feedback to drive their practice during the reminding of the semester. In other words, students who were already practicing desirable study behaviors, spacing and interleaving, would be encouraged to keep up the good work. However, if they were not practicing using spacing and interleaving, they would be encouraged to change their behaviors towards proven good study practices.

At the end of the semester, one week before the final exam, students were asked to complete a closing reflection assignment. This reflection had only two steps:

- Review RPA graph, download and submit: This was the same process as the first reflection.
- Retrieval practice activity graph - reflection questions: Students were prompted to answer three questions.
- Compare the RPA graph with a desirable RPA graph in which a student spaced and interleaved their practice during the entire semester.
- Explain whether they demonstrated desirable learning behaviors using the RPAs during the semester. Desirable behaviors were defined as repeated practice (practicing each RPA at least three times), spacing out the practice of RPAs over days, and interleaving the practice between the RPAs.
- Rank themselves in terms of the percentage of effective practice behaviors, according to the four levels previous presented: Level 4 - Highly Effective Practice Behavior; Level 3 - Effective Practice Behavior; Level 2 - Less desirable practice behaviors; and Level 1 - Low practice behaviors.

Students did not receive any other instruction regarding study habits besides these self-reflections assignments.

3.1 Data Collection

Quantitative and qualitative data was collected from all 75 students who consent to participate in the study. Quantitative data was collected using Canvas course gradebook and qualitative data was collected from the two RPAs reflection assignments. All data was anonymized after collection.

3.2 Data Analysis

Data analysis was conducted on students who completed both reflections to measure growth throughout the semester. From the 75 consented students, 47 students fulfilled that requirement and had their data analyzed. Students where categorized into two groups: Group 1 - students who demonstrate desirable behaviors towards learning, and Group 2 - students who did not demonstrate desirable behaviors towards learning. Two of the authors with experience in U-Behavior based studies placed the students into groups by analyzing the RPA graphs submitted by the students for each one of RPAs reflection assignments.

For the first reflection, grouping desirable behavior was represented by student repetition and spacing practice as demonstrated by Behavior Score 1. Students with 27% or more on Score 1 were placed into Group 1; if they had less than 27% they were placed into Group 2. The 27% was calculated considering that at the moment

Table 1: Number of students per group per reflection

	Group 1 (Desirable)	Group 2 (Undesirable)
First RPA Reflection	10	37
Second RPA Reflection	18	29

Table 2: Percentage of average of spacing and interleaving/mixing for group 1 and group 2 in the second reflection

	Group 1 (Desirable)	Group 2 (Undesirable)
Average of spacing	88%	33%
Average of mixing	95%	56%

of the first RPAs reflection assignment, students should have practiced at least three out of the eleven total RPAs on three different days, corresponding to 27%. For the second reflection, students who were at Levels 3 and 4 were classified as Group1, while students at Levels1 and 2 were classified as Group2. After definitions were set, the two authors coded/categorized the RPA graphs from the first reflection and met to discuss the results. Since there were no differences in the categorizations, one of the authors finalized the categorization by coding the RPA graphs for the second reflection.

In order to verify if there was a statistical significance in performance between those groups, final Canvas exam, final coding exam, and final grades were analyzed. Final Canvas exam consisted of a summative final exam individually generated for each student based on questions randomly selected from Canvas question banks. Final coding exam was a summative programming assignment that encompassed all content learned during the semester. For the coding exam, students are allowed to use their common tools (IDE, Javadoc, etc), but they are restricted on submission attempts.

4 RESULTS

Student behavior was assessed by looking at their categorization (Group 1 or Group 2), and also by determining if there was a change between groups. Students had a Negative Change if they went from Group 1 to Group 2 because that would indicate that they practiced desirable behaviors until the mid-term but then fell out of good practice. Students who changed from Group 2 to Group 1 had a Positive Change, since after the first reflection they changed their practice by spacing and interleaving their RPAs.

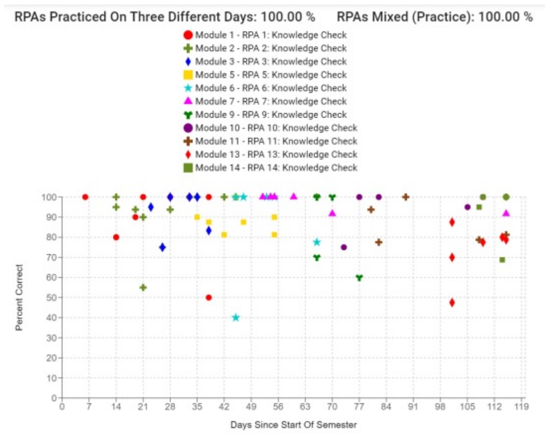
Table 1 presents the total number of students for Group 1 and Group 2 for the first and second reflections. Group 1 had an 80% increase in students by the second reflection. However, not all students who were in Group 1 for the first reflection remained on Group 1 for the second reflection. Three students had a Negative Change, meaning that they were in Group 1 at the middle of the semester and went to Group 2 at the end of the semester. Correspondingly, eleven students had a Positive Change, as they migrated from Group 2 to Group 1.

Table 2 presents the percentage of average for spacing (Behavior Score 1) and interleaving/mixing (Behavior Score 2) for the second reflection. Group 1 spaced more than twice the percentage of Group 2's spacing, and the percentage for interleave practice was extremely high for Group 1. Although research suggested that interleaving is a difficult practice to perform [1], for both groups the interleaving average was greater than the spacing average, indicating that students were coming back and retaking previous RPAs in both groups, but not all students were spacing out those retaking.

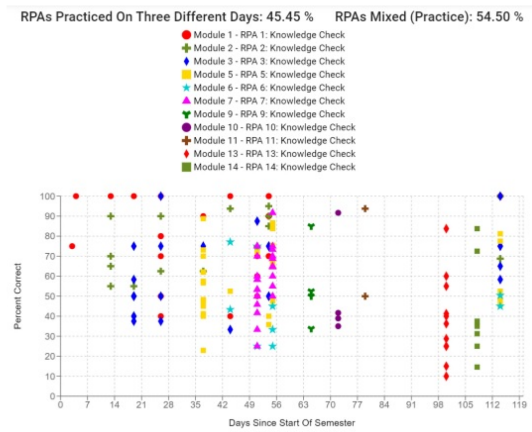
Additionally, all 18 students from Group 1 were assigned Level 4 - Highly Effective Practice Behaviors, meaning not only did they practice desired behavior, they did it with a high degree of consistency in spacing and interleaving. Figures 1a and 1b present two examples of RPA graphs from students in Group 1, to show two different patterns of study found in this group. The X-axis specifies the days since the semester started and the Y-axis represents the percentage of correctness obtained in each attempt. Each attempt is represented by a colored image and each color signifies one of the RPAs offered during the class. After accessing U-Behavior App in Canvas, students had access to an interactive RPA Graph. They could then interact with the graph or download as a pdf file. Students could interact with the App in two ways. When they mouse over the title of a specific RPA in the legend of the graph, all attempts for that RPA were highlighted, so it was easy for the students to see if they are spacing and interleaving/mixing their practice. When they mouse over an attempt, they could access information about that attempt, as percentage of correctness, date and time of the submission.

RPA practice behavior of students similar to those represented in Figure 1a is more constant during the entire semester. They practiced all RPAs at least three different days during the semester and mixed all the RPAs. The practice behavior of students similar to those represented in Figure 1b is less intense. They had an interval of two weeks without practice, and by the end of the semester (two weeks before finals) the student resumed the spaced and mixed practice. Eight of the 18 students had a practice behavior similar to Figure 1a, while ten had a practice behavior more similar to Figure 1b, meaning that most of the students in Group 1 did not practice their RPAs every week during the semester. That makes sense, since they needed to practice each RPA on at least three different days during the semester.

For Group 2, the majority (20 out of 29, which corresponds to 69%) of the students were assigned Level 1 - Low Practice Behavior, while the remaining of students (9 students, 31%) were assigned Level 2 - Less Desirable Practice Behavior. Figures 2a and 2b present one example of a student at Level 1 and Level 2.

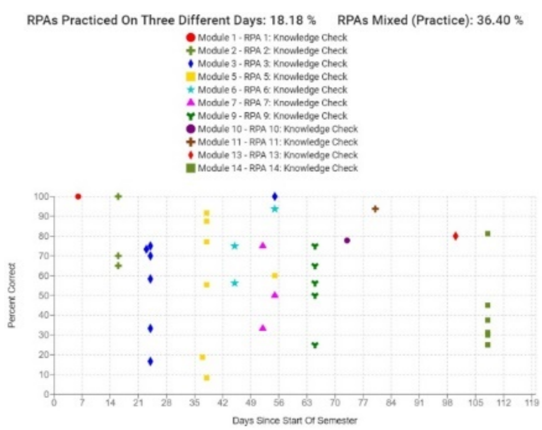


(a)

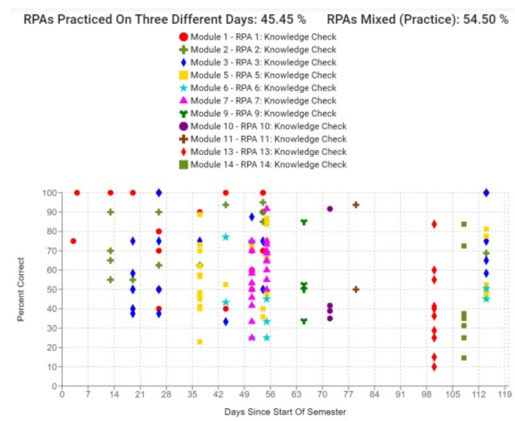


(b)

Figure 1: Examples of RPA graphs for students in Group 1



(a) Level 1



(b) Level 2

Figure 2: Examples of RPA graphs for students in Group 2

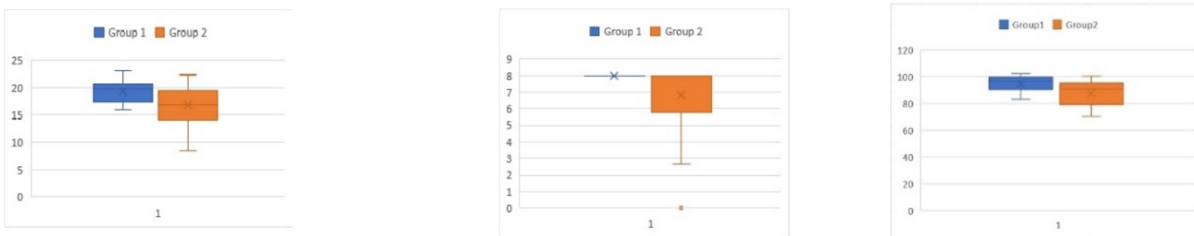
Comparing the four figures, it is observed that there were differences between practice behaviors. Students representing Group 1 (Figures 1a and 1b) have their practice spaced and interleaved during the semester. Students representing Group 2 (Figures 2a and 2b) have their attempts stacked on the top of each other, meaning that they massed their practice on the same day, instead of spacing it on different days.

Figure 3a presents the box plot for the final Canvas Exam for both groups. Group 1 had higher final Canvas exam grades than Group 2, with a median grade increase being 5.5 points. The maximum grade for final Canvas exam was 24 points. Although no students reached the maximum points in any group, we can observe that Group 1 had less variation on their grades than Group 2. In addition to a final Canvas exam, students take a separate final coding exam. Looking at the scores on the coding exam, the same statistical test was also applied verifying Group 1 and Group 2 coding exam scores. Students in Group 1 had a significantly higher mean final coding exam grade (8) than the students on Group 2 (6.84) with $p = .028$, $g = .68$ which is considered a medium effect size.

Figure 3b presents the box plot for the final coding exam for both groups. All students in Group 1 reached the maximum grade for that exam, which was 8 points. This was an unusual result because is somewhat rare for a group to get the full score. Group 2 had much higher variability with most scores being between 6 to 8 points, and all the way down to 0 points. We also examined final course grades for Group 1 and Group 2 (Figure 3c) by doing an Independent Samples t Test. The 18 students in Group 1 had a significantly higher mean final course grade (94.92) than the 29 students on Group 2 (87.37) with $p = .003$, $g = .92$ which is considered a large effect size.

5 LIMITATIONS

This study was only done in one class, and with a small sample size (N=47). Moreover, as pointed out by previous research on spaced practice [9], it is not known if only the best students are doing the RPAs and practicing the desirable behaviors. However, as showed



(a) Final Canvas Exam (b) Final Coding Exam (c) Final Grade

Figure 3: Box plots graphs

in [9, 11], simply doing tests improve students' performance. More research should be done to test the results in other classes and contexts. Another limitation is that students who practiced their RPAs and used U-Behavior self-reflection assignments may have also practiced other study habits, such as study groups. More studies should be done to verify that the correlation with final grades is not at least partially due to unmeasured external practice.

The RPAs have the same types of questions that are asked in the final Canvas exam. However, since RPAs do not involve coding assignments, the use of the final coding exam score could be considered as a threat to construct validity. Moreover, the final course grade includes both the final Canvas exam score and the final coding exam grade plus other graded assignments. Due to a small population, from the 47 students, only 6 students had self-identified as female, it wasn't possible to conduct a gender analysis of the data.

6 DISCUSSIONS AND CONCLUSION

Regarding Q1, comparing Group 1 students (Figures 1a and 1b) with Group 2 (Figures 2a and 2b) students it is observed a difference in their practice behavior. Group 1 students have their practice spaced and interleaved during the semester, while Group 2 students massed their practice on the same day. Eleven students moved from Group 2, the group that focused on studying by cramming for the exam, and switched to Group 1 after their first reflection. When presented with the visualization of their study habits and given a recommended model to follow, those students chose to make a change in their habits. According to Zakrajsek [22], behavior change is challenging and changing the behavior of almost ¼ of the students is an encouraging result.

Regarding Q2, Group 1's final Canvas exam score was 12.5% greater than students on Group 2 which crammed their practice to study for the exams. This shows potentially improved long-term recall, and test taking ability for the students who practice with spacing and interleaving. It is also worth noting that this isn't just because Group 1 took more tests. Both groups equally used the weekly quizzes to study, but instead the difference was when and what they choose to study. Group 1's final coding exam score was 15% greater than students on Group 2 that crammed for exams. This could be an indication that students could improve not only their test taking ability but also their ability in coding. This result is particularly interesting, as there isn't a clear mapping towards test taking capability and coding capability. Group 1's final course grade was 3.03% greater than students on Group 2. Students who spaced and mixed their practice had an overall better performance than students who crammed their RPAs. Although this is a small sample size, the results are promising.

The results for Q2 are in accordance with previous studies [6, 11, 13–15] that encountered an increase in students' performance for those students who practiced space and interleave practices. When students reflect upon their study habits with the support of a visualization and choose to practice in a spaced and mixed manner, they had an improvement in their performance. Combining the results of Q1, it is possible to observe that not only does practicing and interleaving help, but that students can be trained to use visualizations about study habits to reflect and adjust such habits to those that are proven to improve retention of knowledge and learning.

Given the difficulty of CS1 and its importance as a gateway course into Computer Science, there is a benefit to students when study habits are incorporated as part of CS1. Furthermore, when those habits can be visualized, and students can reflect upon them and change their habits to those that will improve their learning, the importance of studying for Computer Science is reinforced. CS1 is an intense course for students as they learn computational thinking, and does teaching study habits help with that process? This question can be explored further, but initial results from this case study show that not only is possible to teach these habits, it is also possible to build them into a course, so it isn't adding additional content on to the student. Instead, they are able to reflect upon their own habits with visualization, and focus on correcting them while using the built-in assignments (RPA) in the course. Despite of the limitations presented in Section 5, this case study presented positive results that aims to start a conversation about building study habits into a course in order to improve student success. Further studies will be conducted to verify the outcomes of this approach in other courses.

ACKNOWLEDGMENTS

This work was supported in part by funding from NSF under Award Number OAC 1931363 and funding from the Center for the Analytics of Learning and Teaching (C-ALT) at Colorado State University.

REFERENCES

[1] Simon Andrew Luxton-Reilly, Vangel V. Ajanovski, Eric Fouh, Christabel Gonsalvez, Juho Leinonen, Jack Parkinson, Matthew Poole, & Neena Thota. 2019.

Pass Rates in Introductory Programming and in other STEM Disciplines. Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education. ACM, pp. 53-71, ACM Conference on Innovation and Technology in Computer Science Education, Aberdeen, United Kingdom. https://doi.org/10.1145/3344429.3372502

[2] Jens Bennedsen and Michael E Caspersen. 2019. Failure rates in introductory programming: 12 years later. ACM inroads 10, 2, 30–36.

[3] Lauren E. Margulieux, Briana B. Morrison & Adrienne Decker. 2020. Reducing withdrawal and failure rates in introductory programming with subgoal labeled worked examples. IJ STEM Ed 7, 19 (2020). https://doi.org/10.1186/s40594-020-00222-7

[4] Nabeel Alzahrani, Frank Vahid, Alex Daniel Edgcomb, Roman Lysecky, & Susan Lysecky. 2018. An Analysis of Common Errors Leading to Excessive Student Struggle on Homework Problems in an Introductory Programming Course Paper presented at 2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah. 10.18260/1-2—29771

[5] Nabeel Alzahrani, Givargis, Tony, Salloum, Mariam, and LePendu, Paea. 2022. Detecting Students of Concern in Introductory Programming Classes: Techniques to Indicate Potential Struggle or Cheating. Ph.D. Dissertation. University of California, Riverside, USA. Advisor(s) Vahid, Frank. Order Number: AAI29213159.

[6] Caleb O'Malley and Ashish Aggarwal. 2020. Evaluating the Use and Effectiveness of Ungraded Practice Problems in an Introductory Programming Course. In Proceedings of the Twenty-Second Australasian Computing Education Conference (Melbourne, VIC, Australia) (ACE'20). Association for Computing Machinery, New York, NY, USA, 177–184. https://doi.org/10.1145/3373165.3373185

[7] Christie A. Lawson, Stephanie McGuire, Russ Hodges, Rosianna Gray, Saundra Y. McGuire, Mark Killingbeck, & Jawn SegoviaLawson. 2021. Recipe for Success: Teaching Students Metacognitive and Self-Regulatory Learning Strategies. Learning Assistance Review, 26(2), 149-178.

[8] Saundra Y. McGuire, & Stephanie McGuire. 2022. The High School Students' Studying and Learning Handouts. Wise Action.

[9] Robert Ariel, & Jeffrey D. Karpicke (2018). Improving self-regulated learning with a retrieval practice intervention. Journal of Experimental Psychology: Applied, 24(1), 43–56. https://doi.org/10.1037/xap0000133

[10] Robert A Bjork, John Dunlosky, and Nate Kornell. 2013. Self-regulated learning: Beliefs, techniques, and illusions. Annual review of psychology 64, 417–444.1

[11] Marissa K Hartwig and Eric D Malain. 2022. Do students space their course study? Those who do earn higher grades. Learning and Instruction 77, 101538.

[12] Priya Harindranathan and James Folkestad. 2019. Learning Analytics to Inform the Learning Design: Supporting Instructors' Inquiry into Student Learning in Unsupervised Technology-Enhanced Platforms. Online Learning 23, 3, 34–55.

[13] Keith B Lyle, Campbell R Bego, Robin F Hopkins, Jeffrey L Hieb, and Patricia AS Ralston. 2020. How the amount and spacing of retrieval practice affect the short- and long-term retention of mathematics knowledge. Educational Psychology Review 32, 1, 277–295.

[14] Iman YeckehZaare, Paul Resnick, and Barbara Ericson. 2019. A Spaced, Interleaved Retrieval Practice Tool That is Motivating and Effective. In Proceedings of the 2019 ACM Conference on International Computing Education Research (Toronto ON, Canada) (ICER '19). Association for Computing Machinery, New York, NY, USA, 71–79. https://doi.org/10.1145/3291279.3339411

[15] Iman YeckehZaare, Chloe Aronoff, and Gail Grot. 2022. Retrieval-based Teaching Incentivizes Spacing and Improves Grades in Computer Science Education. In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1. 892–898.

[16] Henry L Roediger III and Mary A Pyc. 2012. Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. Journal of Applied Research in Memory and Cognition 1, 4, 242–248.

[17] Kelly McKenna, Beth Pouska, Marcia C Moraes, and James E Folkestad. 2019. Visual-form learning analytics: a tool for critical reflection and feedback. Contemporary Educational Technology 10, 3, 214–228.

[18] Kelly McKenna, James Folkestad, and Marcia Moraes. 2020. U-Behavior: Visual-Form Learning Analytics to Enhance Teaching and Learning. In 6^{th} International Conference on Higher Education Advances. Editorial Universitat Politècnica de València, 697–705.

[19] Albert Lionelle, Sudipto Ghosh, Benjamin Say, and J. Ross Beveridge. 2022. Increase Performance in CS 2 via a Spiral Redesign of CS 1. In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education - Volume 1 (Providence, RI, USA) (SIGCSE 2022). Association for Computing Machinery, New York, NY, USA, 502–508. https://doi.org/10.1145/3478431.3499339

[20] Albert Lionelle, Sudipto Ghosh, Shannon Ourada, and Westin Musser. 2022. Increase Performance and Retention: Teach Students How To Study. In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education - Volume 1 (Providence, RI, USA) (SIGCSE 2022). Association for Computing Machinery, New York, NY, USA, 349–355. https://doi.org/10.1145/3478431.3499340

[21] Albert Lionelle, Sudipto Ghosh, Marcia Moraes, Tran Winick, and Lindsey Nielsen. 2023. A Flexible Formative/Summative Grading System for Large Courses. In Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1 (Toronto ON, Canada). Association for Computing Machinery, New York, NY, USA, 624–630. https://doi.org/10.1145/3545945.3569810

[22] Todd D. Zakrajsek. 2023. The New Science of Learning: How to Learn in Harmony With Your Brain. Third Edition. Routledge.