ISSN: 2473-3857
v2 n4023

Testing Frequency in an Introductory Computer Programming Course

Joni K. Adkins
jadkins@nwmissouri.edu

Diana R. Linville
dianar@nwmissouri.edu

School of Computer Science and Information Systems
Northwest Missouri State University
Maryville, MO 64468, USA

Abstract

This paper reports the findings of a study done to determine if increasing the number of exams in a course had an effect on student grades. Some studies have found that more frequent exams positively influence scores while other studies have found more frequent exams do not make a difference in student achievement. This study examines the impact of adding two additional exams to an introductory computer programming course taken by undergraduate computer science, information systems, and other STEM majors. The findings did not show any significant differences in student performance between the fall classes that took three exams and the spring classes that took five exams. In addition a survey was given to discover student attitudes and preferences regarding exam frequency and scheduling. The survey results revealed students want more exams in courses to reduce anxiety and increase confidence and motivation to study.

Keywords: exam frequency, testing frequency, number of exams, computer programming, information systems, computer science

1. INTRODUCTION

Educators often consider and modify the assessment plan in an attempt to improve student learning and achievement. Testing frequency and its impact on student performance have been studied for years. The number of exams and quizzes administered is generally limited due to faculty resources (Kuo & Simon, 2009). The additional work for an instructor to conduct frequent testing in their courses can be daunting. In addition, administering exams consumes valuable instruction time that could be used for classroom learning (Roediger & Karpicke, 2006; Mines, 2014; Leeming, 2002). Therefore, fewer exams, perhaps a midterm and final, are common in many college classrooms. Some believe that students will study less when given more exams because the overall weight of each exam on the overall class grade is lower (Mines, 2014).

Studies have provided conflicting results as some show support that frequent testing in the classroom improves student performance (Leeming, 2002; Kling, McCorkle, Miller, & Reardon, 2005; Gholami & Moghaddam, 2013) while others find that there is no statistical difference in student performance with less frequent testing (Murphy & Stanga, 1994; Mines, 2014). The relevant literature does not come to a definite consensus on the impact frequent testing has on student performance in a course.

2. TESTING FREQUENCY AND STUDENT PERFORMANCE

Faculty can look to the literature for evidence about the appropriate number and frequency of exams. A seminal article by Bangert-Drowns, Kulik, and Kulik (1991) analyzed several other studies and found positive effects of frequent testing in 29 of the 35 studies; 13 of the positive studies and 1 of the negative studies were statistically significant. Another meta-analysis revealed that more frequent exams and student performance were not correlated in a linear fashion (Kuo & Simon, 2009). In other words, adding another exam to a course becomes less significant as the number of exams increases.

More frequent exams means that there is less material for students to learn (Bangert-Drowns, et al., 1991), students may prepare better instead of procrastinating, and students are able to receive more feedback (Mines, 2014). Kling, et al. (2005) investigated the impact of frequent testing on student performance in a marketing course. Their study included 2 sections of a marketing course where 1 section was given 12 quizzes during the semester while the other section was given 3 exams. Both sections were given the same final exam at the end of the semester. Their findings suggested that students retain information better in frequent testing environments with high content overlap (Kling, et al., 2005).

Leeming (2002) conducted a study in a psychology course to determine if a student's performance improved when given an exam every day in class. He proposed that students who did poorly in the course had the ability to learn but just did not study enough (Leeming, 2002). The results showed that grades were significantly higher when students were given an exam every day. Leeming (2002) also found that students in the exam-a-day course outperformed students in the traditional course on a retention test and that fewer students withdrew from the course. A similar experiment to analyze the effect of weekly quizzes on final achievement tests in high school students was conducted by Gholami and Moghaddam (2013). The study included 70 students in different classes taught by the researchers. The classes were split up into an experimental group who received weekly quizzes, and a control group who only received a mid-term and a final. The results indicated the experimental group who took weekly quizzes did significantly better than the control group on the final achievement test (Gholami & Moghaddam, 2013).

The increased performance may be explained by the testing effect. The testing effect is the "phenomenon of improved performance from taking a test" contending that testing both measures and changes knowledge, leading to increased performance (Roediger & Karpicke, 2006, p. 181). In a study by Butler and Roediger (2007), participants watched a lecture and then either studied a lecture summary, took a multiple choice test, took a short answer test, or did nothing. One month later, the participants took a comprehensive exam. Butler and Roediger (2007) found that all review methods improved the participant's score on the final exam with the short answer exam having the most impact. An examination of several studies on testing memory concluded that "repeatedly studying material is beneficial for tests given soon after learning, but on delayed critical tests with retention intervals measured in days or weeks, prior testing can produce a greater performance than prior studying" (Roediger & Karpicke, 2006, p. 189). Thus the testing effect may play a role if the students have more exams in a class, leading them to study and learn from the exams.

College faculty realize that students often wait until right before an exam to begin their studying. These intense "cramming" sessions are encouraged by less frequent exams while more frequent exams may lead to more continuous studying (Roediger & Karpicke, 2006). Michael (1991) called this the procrastination scallop, where students wait until they have a test to begin their studying. Therefore, if more frequent exams were given, students may study more (Michael, 1991).

More frequent testing is not always found to have a significant effect on student performance. Murphy and Stanga (1994) examined the effects of frequent testing in an introductory income tax course. Their experiment used four sections of a single course taught by the same instructor where two sections were given six exams prior to the final and two sections were given three exams before the final. The questions on the exams and the final exam were exactly the same. There was no significant difference in final exam scores between the two groups (Murphy & Stanga, 1994). In another study, Mines (2014) examined the relationship between testing frequency and the final grade in an

ISSN: 2473-3857
v2 n4023

environmental engineering course. The study looked at data from ten course offerings between the years of 2001 and 2012. The statistical data showed that testing frequency had little effect on a student's final grade (Mines, 2014). Due to conflicting evidence, there is still a need for more research to confirm or refute the effect of frequent testing (Ramshe, 2014).

3. TESTING FREQUENCY AND STUDENT ATTITUDES

Another factor to consider when determining the right number of exams is student preference. While the literature reports inconsistent results about student achievement and frequent testing, studies regularly show that students prefer frequent testing (Leeming, 2002; Kling, et al, 2005; Kuo & Simon, 2009). Regardless of these findings, a common pedagogy in college courses remains a midterm exam and a final exam (Roediger & Karpicke, 2006).

The findings are fairly consistent across research areas regarding student's attitude towards the course and instructor (Kuo & Simon, 2009). Bangert-Drowns, et al. (1991) completed a meta-analysis on the effects of frequent classroom testing. Student attitude towards frequent testing was evaluated in four of the studies they examined, and the results showed that students who had more frequent exams rated their instruction more positively. They concluded that by frequently testing students there is a positive effect on the classroom environment (Bangert-Drowns, et al., 1991). Leeming (2002) had students complete a questionnaire at the end of the term regarding specifically the exam-a-day procedures. One question posed was "Given a choice, I would choose this procedure over just a few exams." with the overwhelming majority agreeing (Leeming, 2002). Students indicated they studied and learned more (Leeming, 2002). Attendance may also be positively affected by more frequent exams (Leeming, 2002) while also reducing test anxiety (Kuo & Simon, 2009; Kling, et al, 2005; Gholami & Moghaddam, 2013).

Student evaluations of the instructor is another way to access student attitude. Murphy and Stanga (1991) used the end-of-term student evaluations to assess if there was an adverse effect on instructor evaluations when frequent exams were given. Students in their experimental group who took more exams during the semester, felt stronger about the benefits of the course and the effectiveness of the instructor's teaching (Murphy & Stanga, 1994). While these students also indicated that they felt less anxiety before taking an exam, course evaluations for both groups were favorable. In addition, comments given as feedback on the evaluations supported their conclusion that students prefer frequent testing (Murphy & Stanga, 1994). Kling, et al., (2005) also used the instructor evaluations at the end of the term to assess student attitude. They hypothesized that frequent testing would improve the instructor evaluations at the end of the term. Their results revealed that instructor feedback was higher when more quizzes were given throughout the semester (Kling, et al., 2005).

The researchers in this study were interested in finding if adding exams to an introductory computer programming class would improve learning but were unable to find studies regarding testing frequency in computer programming courses. Another reason for questioning the exam frequency was growing enrollments have made finding empty classrooms for evening exams a challenge. Therefore, changes were made to the testing plan for the spring classes. Keeping in mind the research on performance and student attitudes in regards to testing frequency, the researchers tested to see if more exams in an introductory computer programming course would lead to improved average scores on individual exams, the final exam, and the overall course. In addition, student attitude towards frequency of exams was also assessed.

4. METHOD

Data from two instructors who taught Computer Programming I in both the spring and fall were used in the study. Course materials used in all sections in the study were consistent with the same assignments, projects, and the same or similar quizzes and exams. All students were given a textbook and had access to the same instructor-generated materials including notes, videos, and exercises. All sections of the course enforced the same attendance policy where students lost points after three absences. Students enrolled in the course were mostly Missouri Academy students or undergraduate freshmen with majors in computer science, management information systems, interactive digital media, or another STEM field.

ISSN: 2473-3857
v2 n4023

All students in the fall introductory computer programming course sections met in the evening for three 90-minute exams, approximately five weeks apart. The students in the spring course sections took five exams during their regularly scheduled 50-minute class. The exams were given approximately every three weeks. The total number of exam points in the class did not change. The three exams in the fall were worth 100 points each, and the five exams in the spring were worth 60 points each. The students in the fall course took the same instructor-generated exams. The instructor-generated exams in the spring course were slightly changed so that a student taking the exam later in the day would not have an advantage from learning about exam questions. An example of a modification is shown below.

Question 1 Section 1:
s = "hello"
r = "world"
What is the output of the following operation:
print(s * 3)

Question 1 Section 2:
s = "python"
r = "programming"
What is the output of the following operation:
print(r * 2)

All fall students took the same 200-point comprehensive final exam while the spring students took similar versions of the final exam, again to prevent later sections from having an advantage of learning exam answers.

At the end of the spring term, students who took the course in either the fall or spring were invited to participate in a survey that asked them about their exam frequency preference.

5. RESULTS

The exam scores were averaged to get an overall exam percentage for comparison. The final exam percentage and the final course percentage were also compared. An independent-samples t-test was conducted to compare average exam score, final exam score, and final course grade for fall students who took three exams and spring students who took five exams. While the exam average and the final grade for the spring class with five exams was higher than the fall class, there were no significant differences found in average exam score, final exam score, or final grade between the fall and spring groups. Table 1 shows the relevant statistics.

	Fall (n= 96)		Spring (n =111)		df = 205	
	M	*SD*	*M*	*SD*	*t*	*p*
Exam Avg.	.81	.18	.83	.11	.69	.49
Final Exam	.78	.24	.78	.21	.14	.89
Final Grade	.81	.19	.82	.13	.22	.83

Table 1: Results of t-tests

In addition to examining student performance, the survey allowed the authors further insight into student perceptions of testing frequency. See Appendix A for the survey questions.

Frequency and distribution statistics were calculated on the survey questions with 5 representing strongly agree and 1 representing strongly disagree. Seventy percent of the students (n = 106) agreed or strongly agreed they preferred to have content broken into smaller and more frequent exams with a mean score of 3.89 (*SD* = 1.17). Fifty-five percent of the students (n = 106) agreed or strongly agreed they experienced increased anxiety with fewer exams in a course (*M* = 3.36, *SD* = 1.39). Eighty-one percent of the students (n = 106) preferred having more tests to provide frequent feedback so they could adjust their study skills (M = 4.01, *SD* = 1.01). Students were more confident in courses with multiple exams (n = 106, *M* = 3.85, *SD* = 1.12), and 69 percent indicated they were motivated to study more when there were frequent exams (n = 106, *M* = 3.71, *SD* = 1.17).

Seventy-seven percent of the students in the spring class (n = 96) thought their final grade in Computer Programming I would be higher due to having more frequent exams. Students overwhelmingly (n = 106) preferred taking an exam during the regularly scheduled class period instead of a scheduled evening exam with 88 percent selecting the class period.

6. DISCUSSION OF RESULTS

The results of this study were parallel to Mines (2014) and Murphy and Stanga (1994) as they

ISSN: 2473-3857
v2 n4023

also did not find a relationship between number of exams and student performance. One factor that might explain the absence of significant differences is the impact of more frequent testing decreases with each additional exam (Kuo & Simon, 2009; Bangert-Drowns, et al., 1991). For instance, adding one exam for a total of two exams is found to have a greater impact than adding a fourth exam for a total of five exams. Adding two more exams to the spring classes did not make an impact, perhaps because the fall group had three exams already.

The student survey results showed that the majority of students preferred more exams to fewer exams and thought fewer exams added anxiety. These findings about student exam frequency preference mirror other studies where students have indicated they preferred more exams (Kling, et al., 2005; Gholami & Moghaddam, 2013).

The analysis of written comments had one major theme: the difficulty of attending evening exams due to other commitments, primarily a job. A common evening exam has been used in this course for many years. It was established so all students would take the same exam at the same time and the time allotment could be longer than a regular class period. Many college students have evening commitments including college social or academic activities, part-time jobs, and athletic practices and events making evening exams difficult to attend. Given both the survey statistics and the written comments, more frequent exams will likely be given during the regular class period.

7. LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The final exam was worth 200 points which was approximately 25 percent of the course grade. Students know exactly how many points they need to earn on the final exam to get a specific grade in the class so they may only complete enough of the exam to earn those points. The final exam may not be the best indicator of overall student performance if they do not answer all questions they know. Student evaluations of the instructors could not be compared as the fall evaluations were not available due to an unsuccessful pilot of electronic evaluations. Exam scores and final grades in the data set were not associated with a student grade so analysis on different majors, gender, nationality, or GPA could not be performed. Other limitations include day versus night testing, limited sample size, minor differences in exams, and differences in instructor teaching styles and experience. Future research could study the number of exams given in other levels of programming or information systems courses, the impact of daily or weekly quizzes on performance, or test to find a better measure of overall student learning.

8. CONCLUSION

The number of exams to give in a course will continue to be explored since there is no conclusive evidence to determine whether frequent or infrequent exams have a greater impact on student learning. This study supports other research that shows that students prefer more frequent exams (Bangert-Drowns, et al., 1991; Leeming, 2002). The results showed students like more frequent exams due to decreased test anxiety, higher confidence in knowing the material, and increased motivation to study. College faculty relying on only a mid-term and final exam should reflect on these factors and consider adjusting the number of exams given.

9. REFERENCES

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1991). Effects of frequent classroom testing. *Journal Of Educational Research, 85(2),* 89.

Butler, A. C., & Roediger, H. I. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*4-5), 514-527. doi:10.1080/09541440701326097

Gholami, V., & Moghaddam, M. M. (2013). The effect of weekly quizzes on students' final achievement score. *I.J.Modern Education and Computer Science, 1*, 36-41.

Kling, N., McCorkle, D., Miller, C., & Reardon, J. (2005). The impact of testing frequency on student performance in a marketing course. *Journal Of Education For Business, 81(2),* 67-72.

Kuo, T., & Simon, A. (2009). How many tests do we really need?. *College Teaching*, *57*(3), 156-160.

ISSN: 2473-3857
v2 n4023

Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology Classes. *Teaching Of Psychology*, *29*(3), 210-212.

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavior Analyst*, *14*(2), 229–239.

Mines Jr, R. O. (2014). The impact of testing frequency and final exams on student performance. *American Society for Engineering Education Southeast Section Conference.*

Murphy, D. P., & Stanga, K. G. (1994). The effects of frequent testing in an income tax course: An experiment. *Journal of Accounting Education, 12*(1), 27-41.

Ramshe, M. H. (2014). A review of the studies on the frequent administrations of englich tests. *Journal of Language Teaching and Research, 5*(6), 1412-1416.

Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing memory basic research and implications for educational practice. *Perspectives On Psychological Science*, *1*(3), 181-210. doi:10.1111/j.1745-6916.2006.00012.x

ISSN: 2473-3857
v2 n4023

Appendix A

Computer Programming I Exam Frequency Survey

For each item, please select the value that most closely represents your opinion at this time.

1. Strongly disagree
2. Somewhat disagree
3. Neutral
4. Somewhat agree
5. Strongly agree

1. Please rate your level of agreement with each statement in comparison to other courses that use a 3 exam and final exam format.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I prefer to have the content broken down into smaller more frequent exams throughout the semester.	○	○	○	○	○
I prefer to have more content covered on fewer exams throughout the semester.	○	○	○	○	○
I experience more anxiety when I have fewer exams, such as just a midterm and a final in the course.	○	○	○	○	○
I prefer fewer exams to more exams.	○	○	○	○	○
I prefer having more tests which provides more frequent feedback so that I can adjust my study habits.	○	○	○	○	○
I am more confident in a course with multiple exams (more than 3) than my courses that have fewer (1-3) exams.	○	○	○	○	○
I am more motivated to study because I have more frequent exams.	○	○	○	○	○

2. Do you anticipate that your final grade will be higher due to having more frequent exams?

○ Yes
○ No
○ I am not currently in the course.

3. What is your current grade in the course?

4. What do you think your final grade will be in the course or what was your final grade if you have completed the course?

5. When do you prefer to take exams?

○ During a regularly scheduled class time
○ At a scheduled time in the evening

6. Please share any additional comments regarding testing frequency in college courses?