



Analysis of the Impact of Testing Frequency on Student Performance in a Basic Thermodynamics Course

Dr. John R. Reisel, University of Wisconsin, Milwaukee

John R. Reisel is an Associate Professor of Mechanical Engineering at the University of Wisconsin, Milwaukee (UWM.) He serves as Associate Director of the Center for Alternative Fuels, and Co-director of the Energy Conversion Efficiency Lab. In addition to research into engineering education, his research efforts focus on combustion and energy utilization. Reisel was a 2005 recipient of the UWM Distinguished Undergraduate Teaching Award, the 2000 UWM-College of Engineering and Applied Science Outstanding Teaching Award, and a 1998 recipient of the SAE Ralph R. Teeter Educational Award. Reisel is a member of ASEE, ASME, the Combustion Institute, and SAE. Reisel received his B.M.E. degree from Villanova University in 1989, his M.S. degree in mechanical engineering from Purdue University in 1991, and his Ph.D. in mechanical engineering from Purdue University in 1994.

ANALYSIS OF THE IMPACT OF TESTING FREQUENCY ON STUDENT PERFORMANCE IN A BASIC THERMODYNAMICS COURSE

Abstract

Many engineering courses, such as Thermodynamics, have topics which build upon the material previously learned in the course. For example, students will have difficulty learning the Second Law of Thermodynamics if they have not mastered the First Law. Unfortunately, many students delay studying material in courses until an exam is drawing near. This can be a particular problem in a course which does not inherently interest a student, such as a non-Mechanical Engineering student required to take Thermodynamics as a course outside their major. As a result, they may find themselves well behind in a course and struggling with the material currently being taught because they had not spent enough time learning earlier material while it was being covered in class.

One technique which has been used to motivate students to learn the course material promptly is to test students more frequently, rather than waiting a month or more to do so. The author used this more-frequent-testing technique for many years, using shorter (30-45 minute) quizzes every 2 to 3 weeks in a Basic Thermodynamics course. Before using this method, the author had used a more traditional approach of giving the students 2 mid-term exams during the semester. While the frequent-quiz technique generally received positive feedback from the students and appeared to aid in their learning of the material, the two techniques had not been directly compared to quantifiably measure their relative impact.

In the Fall 2011 semester, the author taught two sections of Basic Thermodynamics, and used the frequent-quiz technique in one and the 2 mid-term exam technique in the other. Other than the testing frequency, the two sections were kept as similar as possible. The lecture content and homework assignments were identical. Results of the final exam in the course were used to judge which technique was more successful in aiding the students' learning. Yet to be determined is the impact of each technique on student retention of the material in a second Thermodynamics course.

In this paper, a thorough discussion of the study methodology and results is presented. A discussion of the benefits and detriments of both techniques is provided, and recommendations for teachers on testing frequency in Thermodynamics courses are made.

INTRODUCTION

Most engineering courses have material which builds upon material taught earlier in the course. In order to succeed in such a course, students usually need to keep current in their learning of the course material as the course progresses. For example, if the material being taught in Week 7 builds upon material that was taught in Week 5, students will have needed to gain at least a basic understanding of the material covered in Week 5 by Week 7 in order to have success in learning the material in Week 7. If the students do not understand the earlier material, they will need to spend time and effort learning that material even after the course has progressed to more

advanced topics – as a result, the students will not be learning the more advanced topics as they are taught and will need to further catch up on that material later.

As a concrete example of this, consider a typical Basic Engineering Thermodynamics course, built around a textbook such as Moran *et al.*¹ or Cengel and Boles². In such a course, students are often first taught basic thermodynamic definitions and concepts, followed by fundamental property relations, concepts of heat and work, the First Law of Thermodynamics, and then the Second Law of Thermodynamics. While additional topics may also be covered, this sequence will suffice for this illustration. If students do not quickly learn how to find Thermodynamic properties, students will need to spend extra time figuring out how to find Thermodynamic properties while learning the First Law concept; this will delay their gaining of an understanding of the First Law. Later, when they encounter the entropy balance concept in the Second Law of Thermodynamics, they may need to be spending time learning how to calculate the heat transfer from the First Law, thereby impeding an attempt to gain a fundamental understanding of the Second Law. While students can still pass a class in such a fashion, they will not be gaining the mastery of the subject matter desirable for use in future courses or their future careers.

It is important for teachers to encourage and motivate their students to learn the course material as it is taught so that the students can fully master the subject. This encouragement can become particularly difficult to achieve in fundamental courses taken by a wide range of students who may not perceive the relevance of the subject matter to their major or their career aspirations. For example, a one semester course on Basic Engineering Thermodynamics is often taken by most engineering students, regardless of major, as a required course. Students in non-Mechanical Engineering disciplines often are less interested in this course, as it is out of their major and appears to have little connection to their career goals. Furthermore, even some mechanical engineering students who are more interested in mechanical design may see the course as a burden that must be endured rather than an opportunity to learn the fundamental principles of energy. Other courses, such as courses in Electrical circuits, Statics, and Dynamics will garner such reaction from various sets of students as well. Therefore, it may be necessary for instructors in such courses to find methods that force students to keep up with the course material, rather than allowing students to put off studying the material until a test is imminent.

One technique that has been proposed and used as a means to keep students more involved with a course is the use of frequent quizzes in a course³⁻⁸. This is often encouraged to be a daily activity in class, as a means to assure students prepare consistently for class.³ Daily quizzing can also provide a means of quickly correcting misunderstandings of material, and encourages active engagement by the students in the course material.⁴ Mahwhinney, *et al.*⁵ found that there was greater variability in class preparation by students when the length of time between tests or quizzes increased. As such, they found that more frequent quizzing increased student preparation. Wesp⁶ found that students who had daily quizzing received higher course grades and discouraged an attitude of delaying studying the material until a large test approached. Connor-Greene⁷ determined that students create their own study patterns in response to the format of a course – thus students who knew that they needed to prepare for daily quizzes were more apt to study regularly. On the other hand, Connor-Greene did not note a correlation between students' grades and the testing frequency they encountered. Bluman *et al.*⁸ also noted that student grades did not show that frequent quizzing improved the students' performance on

concepts on exams, but also did not hurt student morale and students voted to keep the quizzes in place.

Regarding this concern, Lowman⁹ cautioned against frequent quizzing in that it can damage student morale by increasing student anxiety. Furthermore, if students are distracted by a quiz at the beginning of a course, they may not be able to quickly shift their focus to a learning mode for the lecture material after the quiz.

To strike a balance between the possible negative impacts of daily quizzing and the potential positive impacts of improving student learning by increasing their motivation to keep abreast of the course material, the author implemented a student evaluation plan based around 4-6 quizzes a semester, as opposed to 2 large exams, in a Basic Engineering Thermodynamics course. While having students take a quiz every 2-3 weeks requires the students to keep up with course material as it is taught, it does not lead to potentially increasing student anxiety levels through daily quizzing, and provides students with some flexibility in arranging their weekly study schedules. The quizzes were supplemented with weekly homework assignments. In the Fall 2011 semester, the instructor taught two sections of the course, providing an opportunity for the direct assessment of the effectiveness of the quizzes versus exams model. Below, the quiz and exam evaluation methods are described, the design of the experiment to compare the methods is presented, and the results are discussed.

DESCRIPTIONS OF EVALUATION METHODS EMPLOYED

When evaluating student performance in a Basic Engineering Thermodynamics course, the author has used two methodologies. While there is some semester-to-semester variation in the methods, these two methods retain similar features each time they are employed. The two methods will be referred to as the “Quiz Method” and the “Examination Method”. Common to each method is that some portion of the student grade (10%) is determined through student performance on weekly homework assignments, some portion (5%) of the grade is based on a short essay assignment, and that some portion of the final grade is determined by the course’s final exam (30%). The remaining 55% of the grade is based upon in-semester testing of the students.

The Examination Method is the more commonly used method for assessing student performance during the course of the semester. When using this method, the author uses two 75-minute mid-term exams, usually around Week 8 and Week 13 of the 15-week semester. Each exam consists of approximately 4 calculation-based problems, and is designed to test the knowledge of the students on the material that had been covered in the class. The first exam covers Thermodynamic property relations, heat and work, and the First Law of Thermodynamics, applied to both closed and open systems. The second exam is on the Second Law of Thermodynamics, including entropy production. Testing on basic Thermodynamic cycles usually is not performed until the Final Exam, although occasionally a question on Thermodynamic cycles may be used on the second exam.

The Quiz Method employs between 4 and 6 quizzes, depending on the format of the course sessions. When this method was originally used by the author, the course was taught in three 50-

minute sessions each week, while now the course is usually taught in two 75-minute sessions each week. With three course meetings a week, the author employed 6 quizzes during the semester, with one being given every other week starting in the third or fourth week of the semester. Each quiz was based on the material of the 2 previous weekly homework assignments. When the course was shifted to a two-session-a-week format, the author initially employed a 4-quiz format with each quiz lasting 45 minutes. Each quiz was every 3rd week covering the 3 previous homework assignments. This was done to try to reduce the work burden on the students. However, the author later shifted this format to a 5-quiz format, with each quiz lasting 30 minutes and covering the previous two weeks' homework assignments. The first quiz is in Week 5, covering material from the first 3 weeks of the semester. Each subsequent quiz is administered every two weeks, and covers the two weeks of material since the last quiz coverage. Quizzes typically consist of 2 calculation problems, although short problems (such as calculating a Carnot Efficiency of a heat engine) may result in 3 problems. This format was found to be acceptable to the students.

It can be noted that in the present format of each method, there are 150 minutes of class time devoted each semester to in-semester testing of student knowledge. The difference between the methods is in the distribution of the time and frequency of evaluation. The idea behind using the quiz format is that it will help student learning by (a) prompting students to keep up with material as it is being taught in class and (b) reduce text anxiety / pressure on the students, by distributing the evaluation into smaller parts. Each quiz is worth 11% of the grade in the Quiz Method, and each examination is worth 27.5% of the final grade in the Examination Method. By making each testing instrument worth a lower percentage, it is thought that students will feel less pressure when taking each quiz. However, this may also reduce the level of effort that students place on preparing for each quiz. In addition, the Examination Method may prepare students better for the longer (2 hour) final examination, although students should be familiar with taking final exams in college by the time they take Basic Engineering Thermodynamics.

METHODOLOGY OF STUDY

In the Fall 2011 semester, the author was presented with an opportunity to directly compare the performance of students using both the Examination Method and Quiz Method for testing student performance. The author taught two sections of Basic Engineering Thermodynamics at the University of Wisconsin-Milwaukee (UWM), each on a Tuesday/Thursday schedule (Section A was from 11:00-12:15, and Section B was from 3:30-4:45), and each with a comparable class size (31 students in Section A and 27 students in Section B). As the author had previously taught the course over 20 times, the lecture notes were well developed. While there was some variation in what was actually said in each section (as the author uses an extemporaneous lecture style), the material taught in each section was essentially the same each lecture.

Section A was chosen for using the "Quiz Method" of evaluation, and Section B was chosen for the "Examination Method" of evaluation. (It should be noted that some difference in the daily lecture topical coverage did result at times during the semester due to Section A having used some lecture time on their quizzes before Section B took an exam. However, this difference was less than an hour of lecture time, and each class did have some common material every lecture day.) The same homework assignments were given to each class, and were due on the same day

each week. Homework solutions were not posted for student inspection until Section B had turned in their homework for that assignment.

To judge the effectiveness of each method, the performance of each class on the final exam was used for comparison purposes. It should be noted that the final exam was not taken at the same time by the two classes, as Section A took the final exam one day before Section B. To guard against students passing on test questions, small changes to the nature of the problems were made. For example, a First Law of Thermodynamics problem was changed from an air compressor to an air turbine, and another problem had oxygen compressed in a piston-cylinder device on one version of the final exam but having it expand in the device on the other final exam. Doing this maintained the level of difficulty of the exams as being identical, without giving one section a particular advantage on the test.

Figure 1 presents a timeline for how the in-semester testing and assessment was performed for each course, as well as the basic topical coverage on a weekly basis for the two sections. In Section A, Quiz 1 covered the material on Assignments 1-3, Quiz 2 covered the material on Assignments 4-5, and each subsequent quiz covered the material on the next two successive quizzes. In Section B, Test 1 covered the material on Assignments 1 through 7, and Test 2 covered the material on Assignments 8-11. It can be noted that Quiz 5 in Section A did cover material on the Rankine cycle, while there was no evaluation of cycle analysis for Section B until an Otto Cycle question on the final exam.

Week of	9/6	9/13	9/20	9/27	10/4	10/11	10/18	10/25
Topic	Basics		Properties		Heat/Work	First Law		
Sect. A					Quiz 1		Quiz 2	
Sect. B								Test 1
Assn. Due		1	2	3	4	5	6	7

Week of	11/1	11/8	11/15	11/22	11/29	12/6	12/13
Topic	Second Law				Rankine Cycle	Otto Cycle	Vap. Ref. Cycle
Sect. A	Quiz 3		Quiz 4			Quiz 5	
Sect. B					Test 2		
Assn. Due	8	9	10	11	12	13	(14)

Figure 1: Weekly timeline of the topical coverage, and assessment instruments for both sections of the Basic Engineering Thermodynamics course in the Fall 2011 semester. Note, Assignment 14 was not collected, but solutions were available after the due date.

One factor which could not be controlled between the two courses was the actual students who registered for each section. While the students were being drawn from the same body of engineering students, the academic quality and studiousness of the students could vary between the two sections. To gauge the studiousness of the class, it can be noted that 85.9% of the possible homework assignments were submitted by Section A, and 77.8% of the possible homework assignments were submitted by Section B. It should be noted that one student who was struggling in Section A did not take the final exam, and that three students who were performing the worst in the class in Section B did not take the final exam. In addition, one student in Section B was called to active military duty in the middle of the semester, and while he completed the course the following semester, he will not be considered in the analysis. If we consider the 30 remaining students in Section A and the 23 students who did take the final exam in Section B, 88.5% and 84.9% of the possible homework assignments in the class were submitted by Section A and Section B respectively. Therefore, the level of commitment of the students in Section A and the students in Section B who took the final exam should be considered comparable, with Section A possibly being slightly higher. For the purposes of this study, it will be assumed that the academic quality of the students in Section A and the students in Section B who took the final exam are similar.

RESULTS AND DISCUSSION

Table 1 summarizes the statistics for the grades of the students on the final exam in each section. Figure 2 presents the percentages of the students in each section taking the final exam who received scores within defined ranges on the final exam.

In Figure 2, it can be seen that the majority of the grades in Section A were between 70 and 90%, while the grade distribution was shifted higher in Section B, with the majority of grades being between 80 and 100%. From Figure 2, it can also be seen that students in Section B avoided the low grades that did impact 10% of the students in Section A. It should be remembered that all but one (3.2%) of the lower achieving students in Section A took the final exam, while 3 (11.5%) of the lower-achieving students in Section B chose to not take the exam. One of the students in Section A did even acknowledge that, going into the final exam, he had no chance to pass the class but wanted to take the final exam for practice for the following semester. Such individual student decisions do affect the overall analysis of the results, but can not be avoided in this study.

Table 1: Summary statistics for the two sections' grades on the final exam. The scores are out of 100 points.

	Section A	Section B
Students	30	23
Average	74.1	80.8
Standard Deviation	15.1	11.5
Median	79.3	82
High	95	100
Low	27	51

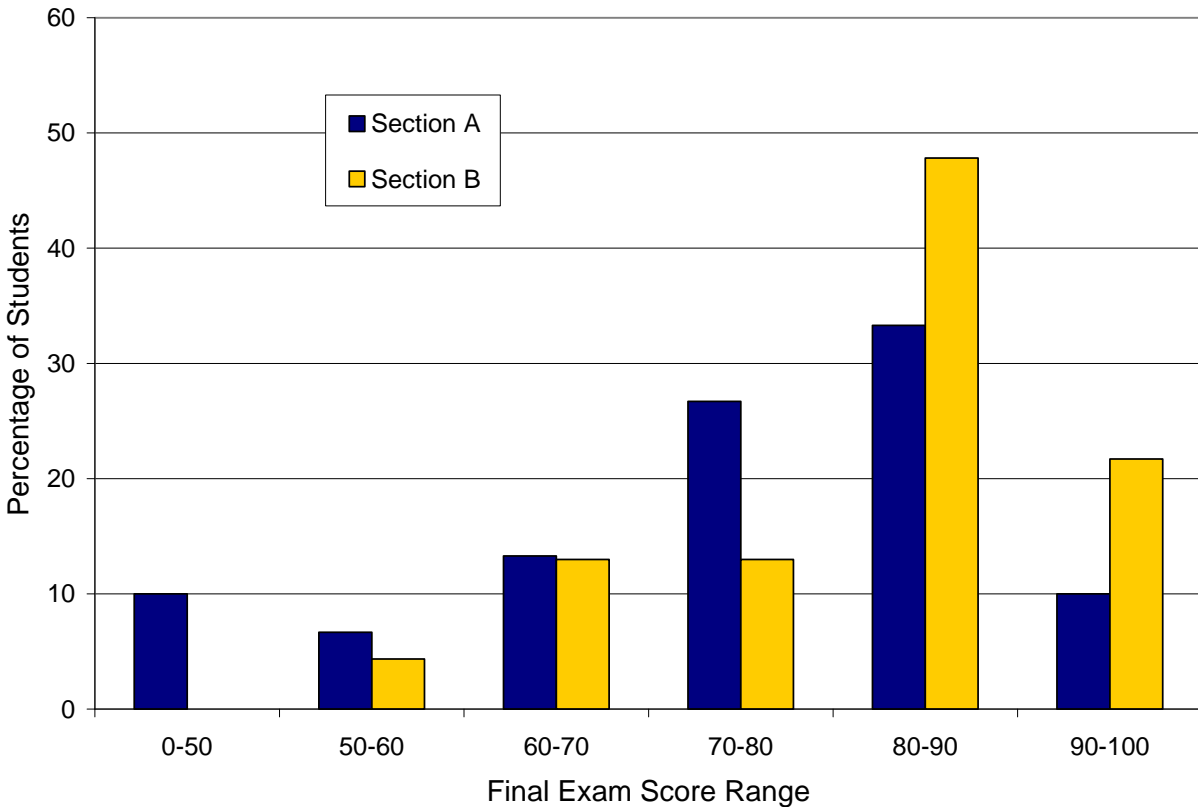


Figure 2: Percentages of students in each section whose final exam score fell inside the ranges indicated.

On the surface, it appears that the frequency of evaluation does not impact the student grades significantly. In fact, if anything, the students who were quizzed more frequently did not do as well as those who were tested infrequently. But when the factor of failing students choosing to not take the final in Section B is considered, it is more easily interpreted that there is no statistically significant difference in performance between the two groups of students. The data presented in Table 1 has an ANOVA p-value of 0.479, indicating no statistical significance in the results.

The percentages of homework assignments submitted by each section were discussed above. It was noted that a slightly higher percentage of possible assignments were submitted in Section A than in Section B. While these numbers are again close enough to be influenced by one or two students in Section B choosing to ignore homework assignments, there is also a possible influence of the more frequent quizzing evaluation method in Section A leading to students making some additional effort to keep up with the material. In either case, it can be noted that the slightly higher percentage of assignments submitted did not eventually lead to better performance on the final exam. Although any individual student likely benefitted from doing the homework, there just wasn't enough of a difference between the class homework submission rates as a whole to impact the comparison of the final exam results.

One item of importance to this study which can not be determined with this sample group is how well the students in the two evaluation methods retain material for use in future classes. While both evaluation methods may produce similar performance on that particular course's final exam, it is possible that the quizzing method may instill a deeper understanding of the material that lasts until used in the future. There is a second semester course in Thermodynamics that is offered, but as the first semester course is taken by all engineering students at UWM and only mechanical engineering students take the second semester course, the number of students who will take the second semester course is only a fraction (~25%) of the students in this initial study – this makes the sample size available for analysis quite low. In addition, as the second semester course is an elective course, some students will take the course 2 semesters after the first semester course, some 3 semesters, some 4 or more, making comparison between the smaller sample size of students even more problematic. It is recommended that this type of study could be better performed at an institution where most of the students in the first semester course take the second semester course at the same time.

As there does not appear to be any particular advantage to in-course student learning from either technique, one needs to weigh the benefits and detriments with regards to other factors, such as instructor workload. The Quiz Method, as it requires creation and grading of testing instruments at more points during the semester, does require a greater time commitment from an instructor than the Examination Method. However, the Quiz Method also offers an opportunity for the instructor to more quickly judge the level of understanding of a topic among the students. This can lead to more prompt correction of misunderstandings or reinforcement of concepts in a class, potentially reducing student confusion or misconceptions from propagating longer through the semester. For the students, the Quiz Method offers the potential to learn from their mistakes more rapidly, but such an advantage only works if the students are taking the quiz results and attempting to identify their knowledge gaps. As mentioned, the Quiz Method may help reduce student test anxiety by reducing the grade impact of each testing instrument, but conversely the Examination Method reduces the number of times that any test anxiety occurs in the semester. Finally, the Quiz Method, by diluting the percentage of the final grade that each testing instrument is worth, may work against increasing the diligence of students' studying by placing less importance on each quiz. Whereas a student may put in considerable effort studying for a test that is worth ~25% of their grade, they may put in less effort to prepare for a quiz worth ~10% of their grade, as they may view it as less importantly individually.

CONCLUSIONS

Taking advantage of an unusual opportunity at UWM, a study was performed to compare the performance of students in a Basic Engineering Thermodynamics course while being exposed to two different evaluation methods. During the course of the semester, one section was given five quizzes, with each quiz approximately every 2 weeks, and the second section was given two longer mid-term exams. The amount of time spent on quizzing/testing was the same for the two sections. Frequent quizzing was considered as an alternative to the more common mid-term exam method because it was thought that it would pressure students into keeping up with learning the course material promptly and that it would reduce test anxiety by reducing the percentage of the final grade dependent on each evaluation instrument.

The results on the final exam for the two sections were very similar, suggesting that there is no immediate student-knowledge benefit for either technique. While the results were slightly better for students who took the two mid-term exams rather than the five quiz format, the results were skewed slightly by a larger percentage of students who were likely to fail the course choosing to not take the final exam in the Examination Method section. It can be noted that the Quiz Method section did experience greater homework submission rates, but again the difference is small and can be impacted by the choices or habits of one or two students.

As there are advantages and disadvantages to both evaluation methods, and as there does not appear to be an immediate advantage to one method or the other for in-course learning of the material, ultimately the choice of evaluation method should rest upon instructor preference. While the Quiz Method offers a number of advantages in theory, unless an instructor finds that those advantages indeed carry over to their own courses, there does not appear to be enough of an advantage of either method to warrant preference in its adoption.

Acknowledgements

The author would like to thank the students in Section 1 and Section 4 of MechEng 301: Basic Engineering Thermodynamics, in the Fall 2011 semester for their willingness to participate in this study.

Bibliography

1. M.J. Moran, H.N. Shapiro, D.D. Boettner, and M.B. Bailey. *Fundamentals of Engineering Thermodynamics*. John Wiley & Sons, Hoboken, NJ. 7th ed. (2011)
2. Y.A.Çengel and M.A. Boles. *Thermodynamics: An Engineering Approach*. Mc-Graw Hill, New York, NY. 7th ed. (2008)
3. R.J. Liebler, "The five-minute quiz." *J. Accounting Ed.*, **21**: 261-265. (2003)
4. P. Kugel, "Improving learning without improving teaching." *Comp. Sci. Ed.*, **1**: 145-152. (1989)
5. V.T. Mahwhinney, D.E. Bostow, D.R. Laws, G.J. Blumenfeld, and B.L. Hopkins. "A comparison of students studying behavior produced by daily, weekly, and three-week testing schedules." *J. Appl. Behavior Analysis*, **4**: 257-264. (1971).
6. R. Wesp. "Reducing procrastination through required course involvement." *Teaching of Psychology*, **13**: 128-130. (1986)
7. P.A. Connor-Greene. "Assessing and promoting student learning: blurring the line between teaching and testing." *Teaching of Psychology*, **27**: 84-88. (2000)
8. J.E. Bluman, K. Purchase, and C.T. Duling. "Daily review quizzes: a hindrance or a help?" *Proceedings of the 2011 American Society for Engineering Education Annual Conference & Exposition*, Paper AC 2011-1146. (2011)
9. J. Lowman, *Mastering the Techniques of Teaching*. Jossey-Bass, San Francisco, CA. (1995)