Declining Test Scores among Introductory Accounting Students: A Comparison of Mean Test Scores in Multi-Day Examination Periods

Matthew Mouritsen, PhD
Associate Professor, School of Accountancy
MBA Program Director
Goddard School of Business & Economics
Weber State University
Ogden, Utah, USA

Jefferson Davis, PhD
Professor, School of Accountancy
Goddard School of Business & Economics
Weber State University
Ogden, Utah, USA

Abstract

The purpose of this paper is to examine the effect that a multi-day exam period has on average daily exam scores. The results suggested that the mean exam score on the first day of the exam period was higher than the mean exam score on the last day of the exam period. Likewise, the mean exam score on the first two days of the exam period was higher than the mean exam score of the last two days of the exam period. Multiple-day exam periods offer flexibility for students in scheduling when they will complete an exam. However, multiple-day exam periods result in lower mean scores at the end of the exam period when compared to the first of the period. Students should be advised of the potential impact on their performance by delaying the completion of an exam. This is important information for students since the authors found that more students take exams later in the exam period.

Keywords: Multiple-day exam periods; accounting; testing; exams.

1. Introduction

Students in introductory financial accounting courses are subjected to a variety of testing approaches. The exams vary in the types of questions asked (i.e., multiple choice, work out problems, and essay) and they vary in the way they are administered in terms of the exam being given in-class or in a proctored testing center. Exam availability also varies in that tests are administered over multiple days or at a scheduled block of time. Additionally, exams may or may not have a designated time limit. The objective of this paper describes how student performance on introductory financial accounting exams can vary based on how the exam is administered, particularly in terms of the number of days it is available in a proctored testing center.

A study done by Reed and Holley (1989), in which an analysis of final exam scores by day taken, found that the day the final exam is taken had a significant impact on performance on the exam. The exams in the Reed and Holley study were scheduled over a four-day period and imposed two-hour time limits on students. They concluded that introductory accounting "A" students (prior to the final exam) scored 9.6% lower on day four when compared to similar "A" students who took the exam on day one. Similarly, they found that "A" students (prior to the final exam) in Intermediate Financial Accounting II who took the exam on day four of the exam period scored 14.2% lower than similar "A" students who took the exam on day one.

Reed and Holley noted "at the introductory accounting level, potential A and F students alike experienced deterioration in exam performance as the week progressed."

They further commented that lower scores during exam periods for introductory accounting students may have been the result of fatigue, prolonged stress, lack of adequate rest, and increased end-of-semester distractions. They noted that A and B students who scored lower near the end of the exam period may not have properly budgeted their time for study. However, D and F students (prior to the final exam) in upper division accounting courses who were assigned to take the exam on day four did score better than their peers who took the exam on day one. They suggest that the additional time available between day one and day four may have been "put to good use." They do not propose two additional factors that may be relevant: First, students in upper division classes may have a stronger relationship with fellow students. This relationship may facilitate the sharing of test content among students who took the exam on day and those taking it on day four. Second, D and F students may be more desperate as a results of their failing or near failing grades and be more willing to cheat to improve their grades.

2. The Problem

By making exams available over several days, students will have the option of using the additional time to study for the exams. Moreover, those students who take the exams early in the testing period will "share" content and answers with those who take the exams later in the testing period. Therefore, mean scores from tests taken at the end of the exam period will be higher than the mean scores from exams taken at the beginning of the exam period. Thus, the overall question is this, "are exams scores in the latter part of a testing period impacted positively by making exams available over a multi-day period?"

3. The Method of Data Collection and Analysis

Students' exam scores in introductory financial accounting taught at an AACSB accredited public university are the subject of this paper. The course, Accounting 2010, is required for all business majors. It is one of five classes comprising the Business Foundations core. Members of the accounting faculty at the institution have flexibility in determining how to administer their exams. The exam scores for this study were collected from one professor's classes taught from 2003 to 2009, during fall and spring semesters. Although the school's introductory accounting courses have a variety of delivery models, (day, evening, online), the scores for this paper were from day-time sections that were taught face-to-face. The classes in the study had an average enrollment of 40.6 students, with a range of 25 to 66 students. Over the time frame of the study, the textbook required for the class was written by three nationally recognized co- authors. The textbook changed from the 3rd to the 4th edition, but the changes to introductory level accounting content were minimal.

The results in this study are derived from midterm exams administered over a seven year period. Each exam was paper-based, worth 100 points (12% of the course grade), and covered two chapters of textbook content. The types of questions on each exam were consistent across time (primarily multiple choice questions with some workout problems). The following table (Exhibit 1) shows the exam number and the content tested on each:

Exhibit 1: Exam Content & Coverage

Exam	Content	Chapters
1	Financial Statements	1 & 2
2	The Accounting Cycle	3 & 4
3	Merchandising and Cash	5 & 6
4	Receivables and Inventory	7 & 8
5	Long-term assets and Liabilities	9 & 10

Every exam in the study was administered in a central testing center. There was no instructor-imposed time limit on any exam. However, students were asked to allow plenty of time, given the hours of operation at the testing center. All of the exams were available over a multi-day period and students were given the flexibility to determine the day on which they would take the exams. The reason for making exams available over a multi-day period was that working students, which represent a large portion of the university's enrollment, need flexibility in being able to meet the demands of learning activities that occur outside of regular class time. Thus, nearly all of the multi-day periods included a Saturday, when no classes were scheduled.

From fall semester, 2003, to spring semester, 2009, data from actual exams were collected and summarized. The following exams were included in the study as shown in Exhibits 2 and 3:

2

Exhibit 2: Number of Times Exams were given in the Study
By Exam Number

Exam #	Count
1	11
2	11
3	9
4	12
5	11
Total Exams	54

Exhibit 3: Number of Times Exams were given in the Study
By Number of Days in the Exam Period

Number of Days	Count
2	2
3	20
4	14
5	6
6	11
7	1
Total Exams	54

In order to determine if the exams scores in the latter part of a testing period were impacted positively by making exams available over a multi-day period, a comparison of means on examination days was conducted. Specifically, two primary comparisons were made:

1. The mean exam score on the first day of the exam period and the mean exam score on the last day of the exam period.
2. The mean exam score of the first two days of the exam period and the mean exam score of the last two days of the exam period.

4. *The Results and Discussion*

As described above, the results of two main comparisons are presented below. First, the mean exam score on the first day of the exam period is compared to the mean exam score on the last day of the exam period. Second, the mean exam score of the first two days of the exam period is compared to the mean exam score of the last two days of the exam period.

4.1.   Comparing First Day and Last Day Means

The following Exhibit (4) illustrates the range of differences between the class mean on the first day and the last day of the exam period on all 54 exams in the study.

Exhibit 4: Range of Differences Between Class Means on the First Day and the Last Day of the Exam Period (by range of difference)
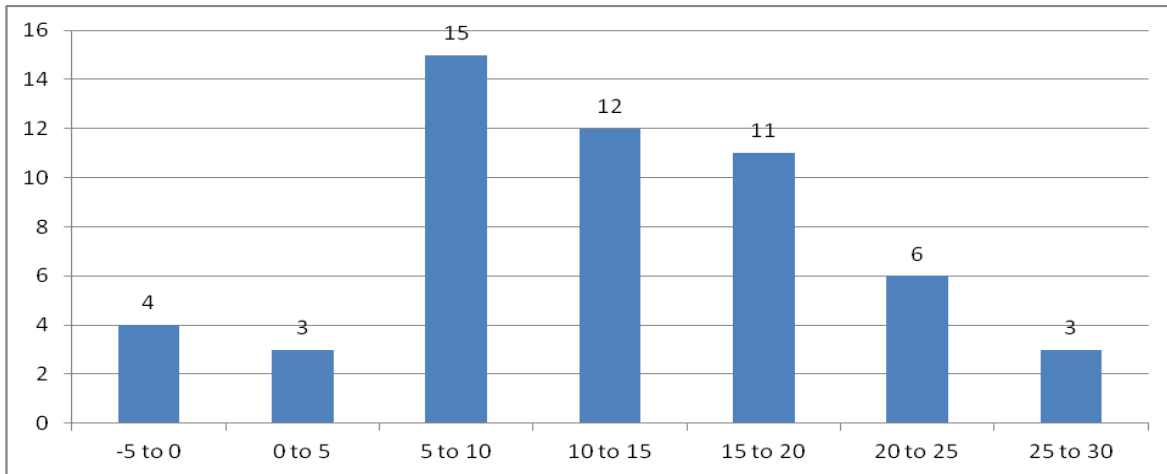

Exhibit 5 shows that four exams had mean scores on the last day of the exam that were in a range of 0 to 5 points higher than the exam scores on the first day of the exam period. All other exams (50) yielded first day means that were higher than the last day means, with three exams having means that were at least 25 points higher on the first day of the exam period. On average, mean exam scores on the first day of the exam period were 12.47 points higher than mean scores on the last day. Specifically, the average first day and last day exam scores and their differences are shown in Exhibit 5.

Exhibit 5: Average Exam Score on the First Day and Last Day of the Exam Period and the Difference between the Two Scores

Exam Number	Exams Given	First Day Mean	Last Day Mean	Difference
1	11	91.9	86.2	5.68
2	11	93.0	80.2	12.83
3	9	89.0	74.5	14.52
4	12	91.9	78.5	13.40
5	11	86.1	70.3	15.74

As illustrated in Exhibit 6, when comparing first day exam means with last day means by exam number (exams with the same content coverage), on average each exam number had higher mean scores on the first day than on the last day of the exam period. For example, on Exam 5, which was administered 11 times, mean scores on day one of the exam period were 86.1, 15.74 points higher than the mean scores of 70.3 on the final day.

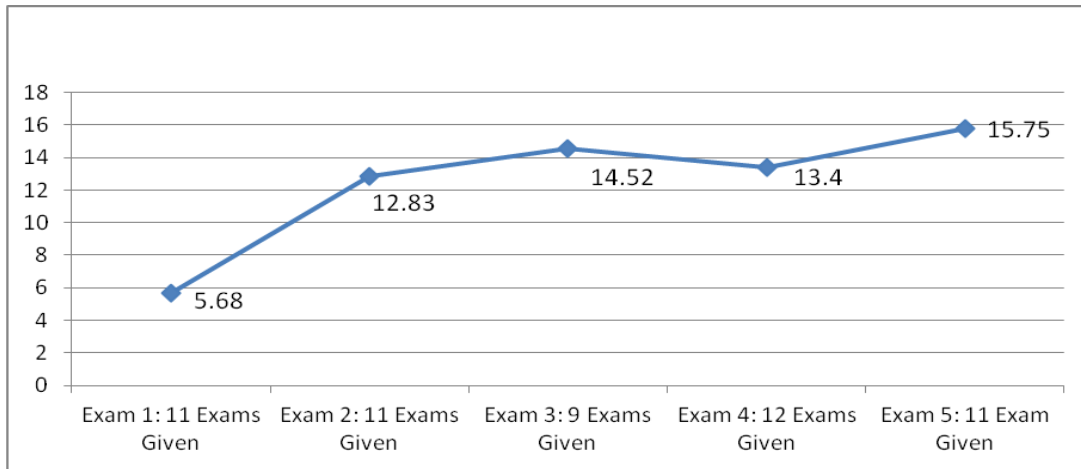**Exhibit 6 Mean Difference between Class Means on the First Day and the Last Day of the Exam Period (by exam number)**

Exhibit 7, below, lists the first day mean, the last day mean, and their difference for exams based on the number of days in the exam period.

Exhibit 7: Mean Difference between Class Means on the First Day and the Last Day of the Exam Period (by days in the exam period)
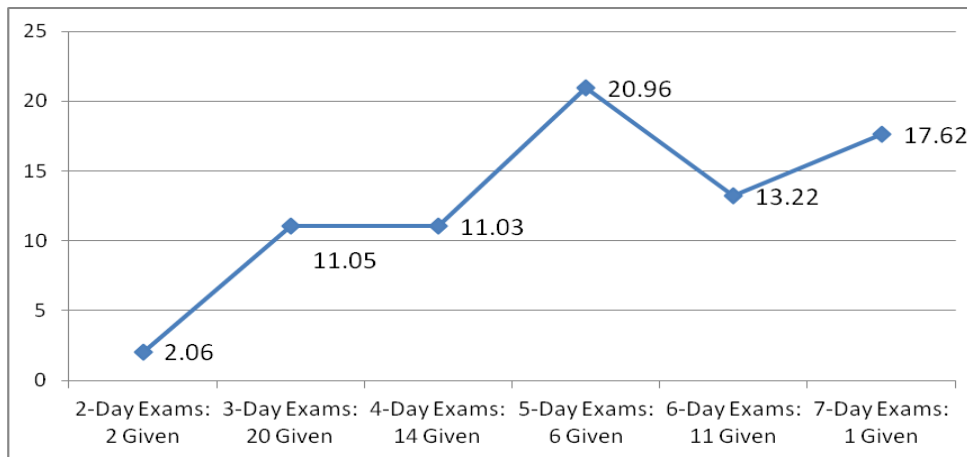

Exhibit 8 depicts the comparison of first day exam means with last day means for exams with the same number of days in the exam period. On average, each exam period, no matter how many days in the exam period, had higher mean scores on the first day than on the last day of the exam period. To clarify by example, there were fourteen 4-day exams administered in the study. Overall, mean scores on the first day of the four day period were 91.99, 11.03 points higher than the mean scores of 80.96 on the last day of the period.

Exhibit 8: Average Exam Score on the First Day and Last Day of the Exam Period and the Difference between the Two Scores, By Number of Days in the Exam Period

Exam Days	Exams Given	First Day Mean	Last Day Mean	Difference
2	2	86.8	84.7	2.06
3	20	90.2	79.2	11.05
4	14	91.99	80.96	11.03
5	6	93.0	72.04	20.96
6	11	88.26	75.04	13.22
7	1	89.0	71.4	17.62

4.2 Comparing the Mean of the First Two Days with the Mean of the Last Two Days

Exhibit 9 below charts the range of differences between the class mean on the first two days and the last two days of the exam period on 32 exams in the study. Only exams with four to seven days in the exam period are included in this comparison. Exams with two and three day exam periods were excluded due to not having two "first" days that were different than the "last" two days in the testing period.

Exhibit 9: Range of Differences between Class Means on the First Two Days and the Last Two Days of the Exam Period (by range of difference)
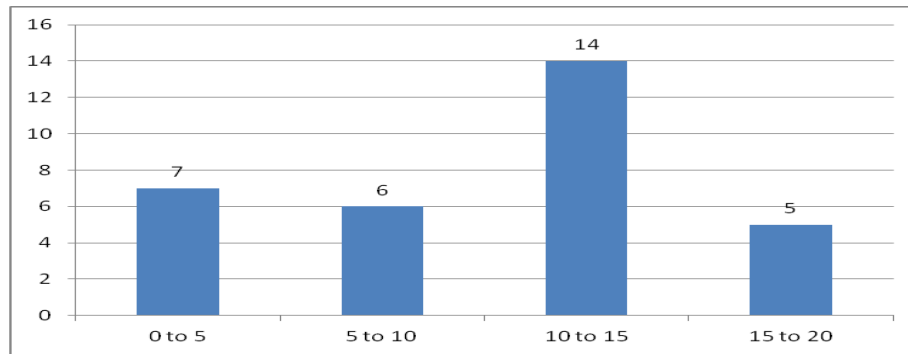

Exhibit 9 shows that all 32 exams administered had mean scores on the first two days of the exam that were higher than the mean exam scores on the last two days of the exam period. Specifically, 14 exams had a mean on the first two days of the exam period that were 10 to 15 points higher than the mean on the last two days. Overall, the mean exam score on the first two days of the exam period was 9.8 points higher than the mean score on the last two days.

As illustrated in Exhibit 10, when comparing the exam mean for first two days of the exam period with the mean of the last two days of the exam period by exam number (exams with the same content coverage), each exam number had a higher mean score on the first two exam days than on the last two exam days. For example, exam 2, which tests the accounting cycle and debits and credits, was administered seven times. The mean score from the first two days of the exam period was 10.3 points higher than the mean score on the last two days.

Exhibit 10: Mean Difference between Class Means on the First Two Days and the Last Two Days of the Exam Period (by exam number)
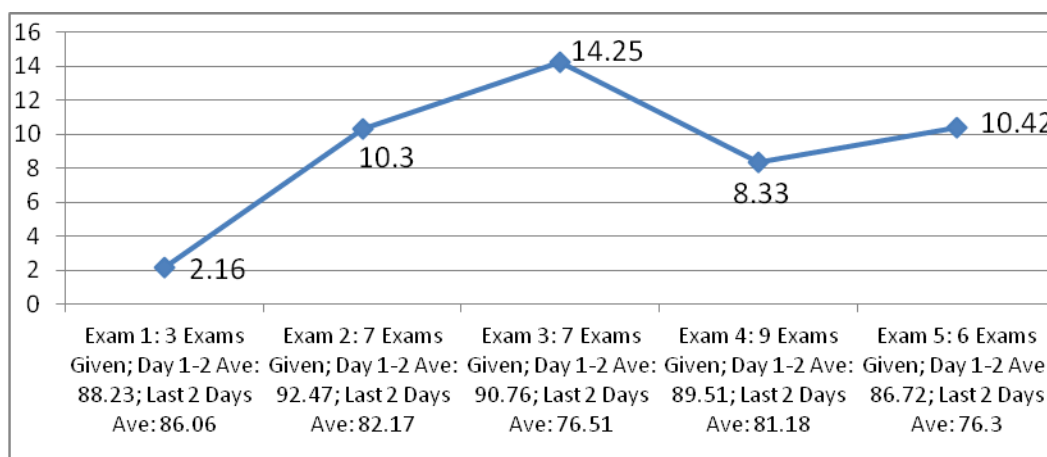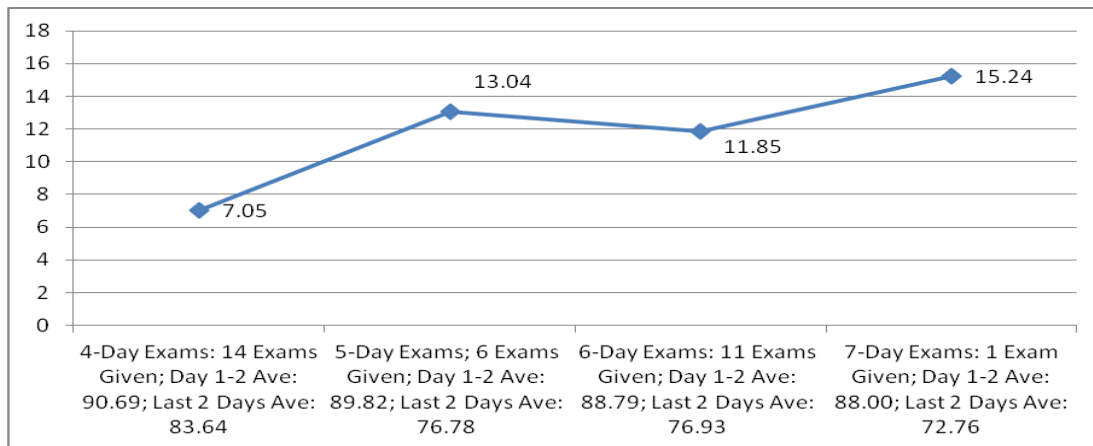

Exhibit 11 depicts the comparison of exam means on the first two days in the exam period with the mean scores of the last two days for exams with the same number of days in the exam period. Overall, each exam period, whether four days or seven days, had a higher mean score on the first two days than on the last two days of the exam period. In particular, there were 14 exams administered to students that consisted of a four-day exam period. The mean score on the first two days of those 14 exams was 90.69, 7.05 points higher than the mean score of 83.64 on the exams taken by students on the last two days of the testing period.

6

Exhibit 11: Mean Difference between Class Means on the First Two Days and the Last Two Days of the Exam Period (by days in the exam period)

	4-Day Exams: 14 Exams Given; Day 1-2 Ave: 90.69; Last 2 Days Ave: 83.64	5-Day Exams; 6 Exams Given; Day 1-2 Ave: 89.82; Last 2 Days Ave: 76.78	6-Day Exams: 11 Exams Given; Day 1-2 Ave: 88.79; Last 2 Days Ave: 76.93	7-Day Exams: 1 Exam Given; Day 1-2 Ave: 88.00; Last 2 Days Ave: 72.76
Value	7.05	13.04	11.85	15.24

5. *Summary and Conclusions*

The results of this study indicate that scores do not increase during a multi-day period for exams given in introductory financial accounting. On the contrary, scores, on average, are higher in the early stages of the exam period than they are in the latter stages of the period. Although offering students flexibility as to when an exam can be completed appears in their favor, the data analyzed in this study suggests otherwise. In other words, to take an exam when it is convenient may actually lead to a lower score, if the convenient time is in the latter part of the exam period. So, if given the flexibility to take the exam, when do students take them? Exhibit 12 provides a view that depicts the percentage of students taking tests organized by the day of the exam period.

Exhibit 12: Percent of students taking 3-day through 6-day exams by day taken

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
3-Day Exams	18%	24%	58%			
4-Day Exams	7%	15%	34%	43%		
5-Day Exams	8%	8%	16%	16%	52%	
6-Day Exams	6%	10%	16%	15%	19%	33%

From Exhibit 12 it is clear that the largest percentage of students preferred to wait to the last day of the exam period to complete the exam, which was also the day on which the lowest mean scores occurred.

The results of this study suggest two findings that are contrary to the hypothesis. First, while many students prolong taking the test to the last day of the exam period, they may not be using that additional time to study, thus achieving lower scores on that day of the testing period. Second, the results also suggest that students taking the exam early in the exam period may not be "sharing" content with students taking the exam at the end of the testing period.

The implications for accounting faculty seem clear. Multiple-day exam periods do offer flexibility for students in terms of providing the time that they consider most conducive for taking the exam. However, multiple-day exam periods also result in lower mean scores at the end of the exam period when compared to the first of the period. Students should be advised of these results and the potential impact on their performance.

The results suggest that additional analysis is warranted in the following areas:

1. Are the descriptive statistics and mean comparisons presented in this study statistically significant?
2. What influence does the availability of the exam on Saturday have on students' decision of when to take the test? In other words, do a large percentage of students select Saturday?

3. How do exam scores on Saturday compare to scores on other test days?
4. Do better students, in terms of GPA or ACT score, tend to take the exam earlier in the exam period?
5. How do the results compare to similar multi-day exams in other courses, such as introductory managerial accounting?

References

Reed, S.A, & Holley, J.M. (1989). The Effect of Final Examination Scheduling on Student Performance. Issues in Accounting Education, 4(2), 327-344.