Student Test Scores And Their Relationship To Order Of Test Completion

Ron Deiter and Vern Pierce

Introduction

A number of different instruments are typically used by instructors to assess an undergraduate student's level of understanding of course subject matter and, ultimately, to assign each student a course grade. While homework assignments, quizzes, term papers, group projects, class participation and oral reports are all commonly used, generally, a major part of a student's course grade is based on classroom performance on exams. Many instructors assume or like to believe that the major, if not the only, determinant of student scores on exams is student knowledge and understanding of the material. However, Marotz and Young (1988) suggest that "exam taking ability" measured, in part, by the amount of time spent taking the exam may also be an important factor that affects student grades.

The main purpose of this paper is to provide evidence of the relationship between student test scores and the amount of time spent taking an exam as measured by the student order of test completion. This topic was not directly addressed by Marotz and Young as they specifically examined the relationship between course grades and time spent taking the final exam. Although not reported in their article, forty percent of each student's course grade was based on the final exam. This paper will also provide additional evidence on the general applicability of the Marotz and Young findings which they state in their paper may have been dependent upon the particular exam (a 2-hour final) that was given in their specific course.

Instructors may assume students who did poorly on an exam did so because they did not know the material. Many of those same students often have other explanations of their poor performance, including that they knew the material, but just ran out of time. A negative correlation between student test scores and order of test completion (i.e. students who take longer to complete a test tend to receive lower test scores) would give credence to the student's lack of time lament. This would also suggest that instructors may want to consider reducing exam length or using more nonexamination-type instruments to evaluate students' performance which could especially benefit students who may know the material but who tend to "panic" when they are placed under time pressure such as that in an exam setting. Test scores of these students may not be commensurate with their knowledge of the material.

Psychological research has repeatedly reported a negative correlation between performance and anxiety (Tryon, 1980). Students who take longer to complete an exam are likely to experience more anxiety due to the pressure to complete an exam on time and, hence, might be expected to score lower. Marotz and Young (1988) initially offered two alternative hypotheses as to which students would take less time to complete an exam: 1) those who did not know the material and thus had little to write, or 2) those who knew the material and who could quickly write correct answers to examination questions. Results of their case study supported the first hypothesis in that students who spent more time writing exam answers averaged higher course grades. The findings by Marotz and Young, as well as those reported by Covington and Omelich (1987) suggest that using test anxiety to explain poor test performance may be overemphasized.

Procedure

The results presented in this paper are based on information collected during spring semester 1990 in two agricultural economics courses taught at Iowa State University. Agricultural Cooperatives (Econa 421) had an enrollment of 63 undergraduates, most of whom were seniors (44) and Agricultural Business majors (53). The course, usually an elective or restricted elective for students, emphasizes the identification, management, financing, legal aspects, and taxation of rural cooperative businesses. Agricultural Marketing (Econa 335) had an enrollment of 85 undergraduates with 26 seniors, 37 juniors, and 22 sophomores. The class consisted of 40 Agricultural Business majors and 45 students who were majors in other departments within the College of Agriculture. The course, a requirement for Ag Business majors and generally an elective course for other majors, studies producer marketing alternatives with emphasis on futures contracts, options, contracting, and commodity price analysis.

The order of student test completion was recorded as the students turned in their exams for all 63 students for two separate hourly (noncomprehensive) exams in Econa 421. Similar information was collected in Econa 335 for 68 students for an hourly exam and for the two-hour final exam. Some graduating seniors were excused from taking the Econa 335 final. Only those Econa 335 students who completed both exams are included in this study.

Exam #1 in Econa 421 consisted of 25 multiple choice questions, 15 cooperative identification problems, and 5 short answer/discussion questions. Exam #2 in Econa 421 consisted entirely of 45 multiple choice questions. Each of these Econa 421 exams accounted for 40 percent of a student's course grade. Exam #1 in Econa 335 contained 40

Deiter is an associate professor and Pierce is an instructor in Agricultural Business in the Department of Economics at Iowa State University, Ames, IA 50011-1070.

multiple choice questions of which 10 questions were work problems that required students to do some mathematical calculations. Exam #2 (the course final exam of which about one-third was comprehensive) contained 50 multiple choice questions of which 10 questions required students to perform mathematical calculations. Each of these Econa 335 exams accounted for 20 percent of a student's course grade. Each hourly exam included in this study was completed by approximately three-fourths of the students prior to the expiration of the one-hour time limit. These students could have taken longer to write the exam. The remaining students took all of the allotted time and, in some cases, up to three or four additional minutes. Spending more time on writing the exam was not an option for these students. In contrast, the final exam (Exam #2) in Econa 335 was equivalent in length to an hourly exam, yet students were given two full hours to complete it. All students completed this exam within the two-hour time limit. Thus, this study analyzes the relationship between exam scores and order of exam completion for exams that differ by type of questions asked, length of allowable completion time, and relative weight in a student's overall course grade.

Previous academic performance information on each student was also obtained from the Agricultural Classification Office. In particular, each student's cumulative ISU grade point average through the fall semester 1989 (the semester prior to taking the course) was obtained. This information was considered to be an indicator of whether a student was a "good" student independent of their exam performance.

Results

Table 1 presents a two-way contingency table or distribution of students by finish group and score group for each exam and for all exams combined. Each finish and score group contains approximately one-fourth of the students. In some instances, the numbers of students are not precisely equal across groups because the total number of students does not divide into equal quartiles and also because students with identical exam scores were put into the same score group having been assigned the same rank using the mean rank method of handling ties. Students in finish group #1 and score group #1, for example, were the first one-fourth to complete the exam who also received a score in the top group (approximately one-fourth) of the class. The numbers in a given row in Table 1, show the finish group distribution for students who were in a given score group. Similarly, the numbers in a given column in Table 1, indicate the score group distribution for students who were in a given finish group.

If the findings of Marotz and Young (1988) were to hold for this study, a skewed distribution would be expected. However, a visual inspection of the data in Table 1 does not reveal any apparent score tendency for students who finished in a particular group or finish tendency for students who scored in a particular group. This observation holds for all of the different types of exams included in this study. A chi-square nonparametric distribution test (see Ostle and Mensing, 1975) was used to determine if any of the distributions for a given exam in Table 1 were significantly different from a uniform distribution (i.e. equal distribution of students across cells). Specific χ^2 values are reported in Table 1. None of these values are significant at the 90 percent probability level. Therefore, the null hypothesis of a uniform distribution cannot be rejected. These results do not reinforce the findings of Marotz and Young (1988) which suggested a negative correlation between finish group and score group.

Spearman rank correlation coefficients (see Ostle and Mensing, 1975) were calculated to determine if there were any significant correlations between students' 1) finish rank (FIN) and score rank (SCORE) for each exam, 2) FIN and cumulative grade point average (GPA) for each exam, 3)

Table 1. Number of Students Classified by Finish Group and Score Group.

χ^2	Score Group		Finish Group			
			#1	#2	#3	#4
9.26	421 EX#1	#1	4	3	6	2
		#2	4	6	4	5
		#3	5	3	2	2
		#4	2	4	4	7
9.90	421 EX#2	#1	2	4	5	6
		#2	5	5	3	2
		#3	2	6	4	3
		#4	6	1	4	5
7.29	335 EX#1	#1	4	2	6	6
		#2	5	5	6	3
		#3	3	4	2	3
		#4	5	6	3	5
9.64	335 EX#2	#1	7	4	3	4
		#2	5	3	5	3
		#3	4	6	3	3
		#4	1	4	6	7
11.61	ALL EXAMS	#1	17	13	20	18
		#2	19	19	18	13
		#3	14	19	11	11
		#4	14	15	17	2-

Table 2. Spearman Rank Correlation Coefficients.

Rankings	Coefficient	
	All Students	First 3/4 Students Completed
FIN and SCORE		
421 EX#1	+ .13	- .04
421 EX#2	- .06	- .124
335 EX#1	- .108	- .114
335 EX#2	+ .219(**)	
FIN and GPA		
421 EX#1	- .013	+ .018
421 EX#2	- .105	- .075
335 EX#1	+ .08-	+ .228
335 EX#2	+ .312(*)	
FIN EX#1 and FIN EX#2		
421	+ .639(*)	
335	+ .664(*)	
SCORE and GPA		
421 EX#1	+ .53(*)	
421 EX#2	+ .5-3(*)	
335 EX#1	+ .623(*)	
335 EX#2	+ .705(*)	

* Significant at the 95 percent probability level.
** Significant at the 90 percent probability level.

Table 3. Number of Students Classified by Finish Group and Cumulative GPA Group.

χ^2	Cum GPA Group		Finish Group			
			#1	#2	#3	#4
8.83	421 EX#1	#1	6	1	4	5
		#2	2	5	5	3
		#3	4	3	4	5
		#4	3	7	3	3
7.20	421 EX#2	#1	4	1	6	5
		#2	4	5	2	4
		#3	3	5	3	5
		#4	4	5	5	2
5.88	335 EX#1	#1	6	2	4	5
		#2	5	6	2	4
		#3	3	4	6	4
		#4	3	5	5	4
11.05	335 EX#2	#1	8	4	2	3
		#2	5	4	3	5
		#3	3	5	6	3
		#4	1	4	6	6
14.79*	ALL EXAMS	#1	24	8	15	18
		#2	16	20	12	16
		#3	13	17	19	17
		#4	11	21	19	15

* Significant at the 90 percent probability level.

finish rank on the first exam (FIN EX#1) and finish rank on the second exam (FIN EX#2) for each class, and 4) SCORE and GPA. These results are reported in Table 2.

For each of the exams where there was a one-hour time constraint, there was no significant correlation between students' finish rank and score rank. This lack of correlation is consistent with the χ^2 tests reported in Table 1. This result also held when students in the last group to finish exams were omitted from the analysis. Omitting these students who had completed the exam at approximately the same time having used all of the allotted time allowed for a test of correlation between finish rank and score rank for those students who could have taken more time in writing the exam.

There was a significant positive correlation at the 90 percent probability level between FIN and SCORE rankings for students on exam #2 in Econa 335. This suggests, on this test, students who performed the best tended to finish earlier, a result which contradicts those reported by Marotz and Young (1988). This test was the only 2-hour exam included in this study which was completed by all students within the allotted time period. This suggests that for hourly exams there may not be enough time for finish ranks to effectively distinguish the good students from the poor students.

Students who finished early on exam #1 also tended to finish early on exam #2 as indicated by significant correlation coefficients for both classes. This suggests a certain consistency about the time spent taking tests by students from one exam to another. There also was an expectedly high correlation between a student's previous academic performance as measured by GPA and their individual exam performance (SCORE) in each class.

Another attempt to determine whether or not "good" students tend to use more or less time to finish exams than do "poor" students is presented in Table 3. This table contains a two-way contingency table of students by finish group and cumulative grade point average group. The same chi-square test was used to measure actual distribution deviation away from that of a uniform distribution. For each individual exam, the chi-square test did not allow rejection of the null hypothesis of a uniform cell distribution. There was no discernible correlation between student groups based on cumulative GPA and exam finish rankings. However, for all exams combined, the uniform distribution hypothesis was rejected with the chi-square test at the 90 percent confidence level. This is due to the apparent tendency, as reported above, for the early finishers to score relatively well on exam #2 in Econa 335.

Conclusions

The purpose of this paper was to examine the relationship between student test scores and the amount of time spent taking an exam as measured by the student's order of test completion. The statistical results did not reveal any significant correlation between student relative performance on hourly exams and their order of test completion. In other words, some of the early finishers of exams scored well and some scored poorly. This was also true for the in-the-middle and the late finishers of exams as well. These results were true for hourly exams of different types. For the one exam in this study for which there was no major time constraint and, hence, reduced potential for test anxiety, there was some evidence that the "best" students finished relatively sooner.

These results do not support the Marotz and Young (1988) hypothesis that "poor" students tend to be earlier finishers because they don't know the material and thus have less to write. These results (with the possible exception of those for Exam #2 in Econa 335) also do not support the anxiety-based hypothesis that "poor" students are late finishers who have their exam performance adversely affected by extra anxiety associated with their anticipated running out of time to complete the test.

One possible explanation of our results is that within both the "good" and "poor" student categories there may be different types of exam takers. For example, there can be "good" students who complete exams quickly because they know the material, are quickly able to answer the questions, and are confident in their answers the first time through the exam. Other "good" students may also know the material and write answers quickly, but these students are extra cautious in using most or all of the allotted time to carefully review and double check their answers before turning in their exams. There can also be "poor" students who complete exams quickly because they don't know the material, quickly guess at answers, and see very little value in using more time to review their answers. Other "poor" students may also not know the material but they use most or all of the allotted time writing answers that sound good or hoping that if they wait long enough the right answers will ultimately dawn upon them.

Our results suggest that "exam taking ability" as measured by the amount of time spent taking an exam is not likely to be a major factor that affects student's performance on exams. Therefore, it seems unlikely that examination length

Deiter (continued on next page).

Using SAS As A Gradebook

Donald M. Johnson

Introduction

Of all the college administrative and management tasks performed by classroom agriculture teachers, recording and calculating student grades is one of the most tedious and time consuming chores. This is especially true in large classes with several graded assignments. In such a situation, the traditional method of recording and averaging student grades (e.g. use of gradebook and calculator) is especially cumbersome.

According to Baker and Creel (1985, p. 23) "The use of a computer to assist in storage and calculation of student grades can greatly simplify the process." Several software programs have been suggested for use in managing student grades. Wood (1990) reviewed commercial gradebook programs, while Burton (1983) discussed the use of spreadsheets in academic record keeping. Other authors have reported on user written gradebook programs (Baker and Creel, 1985; Patterson and Reneau, 1985).

An alternative method of managing student academic records is the use of SAS[1] software as a computerized gradebook. According to the *SAS Introductory Guide for Personal Computers, Release 6.03* (SAS Institute, 1988)

> The SAS System is a software system for data analysis and report writing. With base SAS software you can store data values and retrieve them, modify data, compute simple statistics, and create reports all in one SAS session. Other SAS software provides graphics, forecasting, data entry, and sophisticated statistics (p.3).

These capabilities make SAS an excellent alternative to other less powerful grade management software programs.

Purpose

Many college agriculture teachers already use SAS to analyze research data. However, some many not realize that SAS can also be easily used to manage student academic records. Therefore, the purpose of this article is to provide an example of how SAS can be used to record, analyze, and report student grades.

Example

The use of SAS in student grade management will be illustrated by applying it to an introductory agricultural communications course, ACO 3203, taught by the author. In this example, a series of SAS statement will be used to manage course records for five students (Table 1). The output to be produced includes a student grade summary, a measure of the relationship between the number of absences and final course percentage, and an assignment summary.

SAS statements

The SAS statements in Figure 1 were used to record, analyze, and report student grades for ACO 3203. In this example the data (student names, scores, and absences) and SAS statements were both entered into the same file. However, separate data and SAS statement files can be established and then brought together using an "infile" statement. This practice may be desirable with large classes or if a teacher is responsible for a multiple sectioned course.

SAS output

The summary of student scores is presented in Figure 2. This output can be used to determine final course grades. Similar summaries can be produced throughout the semester to provide feedback on course performance.

Figures 3 and 4 present the relationship between student absences and final course percentage. Figure 3 presents this information numerically; Figure 4 presents it graphically.

Table 1. Student Academic Records for ACO 3203.

		Assignment*			
Student	Absences	#1	#2	#3	#4
Andrews	2	90	85	100	180
Baker	0	100	90	91	192
Jones	4	60	80	82	170
Smith	3	84	80	82	164
Young	1	100	90	68	100

*Note. #1=business letter, #2=news article; #3=technical report; and #4=final examination

Johnson is an assistant professor in the Department of Agricultural and Extension Education, P.O. Drawer AV, Mississippi State University, Mississippi State, MS 39762.

[1] SAS is a registered trademark of the SAS Institute Inc., Cary, North Carolina.

Deiter (continued from previous page).

has much of an impact on the relative performance of students on exams. This should be reassuring to instructors who primarily base student course grades on student scores on exams. This also suggests that there are factors possibly more important than test length that instructors should be concerned about in constructing tests, such as clarity and content validity. Our results do not address the question of whether the relative performance of students on written examinations would be consistent with their relative performance on nonexamination type evaluation instruments. We leave this issue to further research.

References

Covington, Martin V. and Carol L. Omelich. (1987). "I Knew It Cold Before the Exam: A Test of the Anxiety-Blockage Hypothesis," *The Journal of Educational Psychology*, 79(4):393-400.

Marotz, Craig and Douglas Young. (1988). "Time Spent Writing Exam Answers as a Factor Related to Final Course Grade," *NACTA Journal*, 32(1):16-18.

Ostle, Bernard and Richard Mensing. (1975). *Statistics in Research*, 3rd edition, Ames, Iowa: Iowa State University Press.

Tryon, G.S. (1980). "The Measurement and Treatment of Test Anxiety," *Review of Educational Research*, 50, 343-372.