EXAM QUESTION SEQUENCING EFFECTS ON MARKETING AND MANAGEMENT SCIENCES STUDENT PERFORMANCE

Michael Russell, St. Bonaventure University
Michael J. Fischer, St. Bonaventure University
Carol M. Fischer, St. Bonaventure University
Kathleen Premo, St. Bonaventure University

ABSTRACT

Marketing and management sciences instructors may utilize multiple versions of a single examination as a deterrent to student cheating, particularly for exams comprised of multiple-choice questions. Viewed from this perspective, understanding the impact that multiple exam versions has on student performance is important, because such an understanding may serve as an aid to instructors in designing equitable examinations.

The results of previous research examining the effect of multiple-choice question sequencing have been inconsistent. This paper reports on a series of investigations into the impact of exam question sequencing on marketing and management sciences student performance. The research involved manipulating the order of presentation of multiple-choice questions on the exams in four different courses taught by two instructors over one semester The results of the investigation reported in this paper suggest that a scrambled (random-ordered) format provides the most consistent responses and can be viewed as the most desirable sequencing method. This paper describes the research design and discusses the implications of the findings and directions for future research.

INTRODUCTION

Marketing and management sciences instructors may utilize multiple versions of a single examination as a deterrent to student cheating, particularly for multiple-choice or other objective questions. Understanding the impact of multiple exam versions on student performance is therefore important because such an understanding aids instructors in designing "equitable" or "unbiased" examinations (i.e., exams that do not unfairly advantage or disadvantage particular students simply because of the version of the exam that they are given). Additionally, identify-ing situations in which student performance is affected by the order in which questions are presented may have broader implications for marketing and management sciences education by providing insights into the organization of student knowledge structures and the manner in which students access their knowledge when confronted by marketing and management sciences problems.

There is very little research that has examined evidence of exam question sequencing impacts on marketing and management sciences student performance. Further, the results of the limited

body of research in this area are mixed. While some researchers have found relatively strong evidence of exam question sequencing impacts, others have failed to find any impact of exam question sequencing on student performance, and still others have found evidence of impact only in some courses, or for certain subsets of students within the studied courses.

The purpose of the research presented in this paper is to extend our knowledge of exam question sequencing impacts on marketing and management sciences student performance. Specifically, this paper reports on a series of investigations that examined multiple-choice question sequencing impacts across four courses taught by two instructors. Additionally, the paper discusses the implications of the results of this stream of research for marketing and management sciences educators and provides direction for future research.

The remainder of this paper is organized into four sections. The first provides a summary of previous studies that have examined test question sequencing impacts. The second section describes the design and execution of our research and the third section presents our results. Finally, the fourth section discusses the implications and limitations of this research.

REVIEW OF THE EXISTING LITERATURE

Early research in the area of test-item order examined the systematic alteration of items using question difficulty as the basis for ordering. Results of these studies reported higher performance on exams when items were arranged from easy to difficult. Differences in motivational level and/or anxiety were determined to be key determinants for the results obtained (Sax and Cromack 1966; Flaugher 1968). Recent research has investigated the relationship between the order of topical presentations in class and the order those topics are presented in an examination. Some researchers have described the variations

as forward-sequential and reverse-sequential. In the forward-sequential method, the test questions are arranged in a sequence coinciding with the topical presentations in the course. In the reverse-sequential approach, the test questions are arranged in reverse order. A third approach is to scramble the items using a randomized approach (Pettit et al. 1986). Researchers examining the effect of multiple-choice question sequencing have reported mixed results. For example, Balch (1989) found that students scored slightly higher when test items were grouped sequentially (relating to text and lectures) than on tests when test items were grouped by text chapter but ordered randomly, or when test items were ordered randomly. Araujo and Semb (1979) completed two experiments using a 2 x 3 x 2 (instructors x test forms x class time) designs to compare three aspects of student performance (mean final examination scores, homogeneity of variance, and retake frequencies) on three test forms. The test items on the three forms were arranged in three ways: (1) forward – items paralleled the presentation of material in class; (2) backward – items were in an opposite order; and (3) random order. The analysis indicated that item order did not significantly influence student performance in either experiment.

Sander (1988) examined the effects of presenting test items in random order or in a sequence parallel to the order of presentation by testing 92 undergraduates in an introductory psychology course at Emporia State University in Kansas. Students were assigned to receive four versions of examinations in various orders. There was no instance in which a sequential test resulted in a higher score than a random test. Carlson and Ostrosky (1992) conducted an order sequencing experiment using a large section (400) of a microeconomics principles class. They concluded that students taking a scrambled form of an examination may be at a disadvantage and students taking the ordered form of an exam may benefit in the form of marginally higher exam scores. While examining the predictors of student performance in the introductory marketing

class, Borde (1998) found that question sequencing on multiple choice exams led to inconclusive results.

Chidomere (1989) examined students enrolled in principles of marketing. Four examinations were administered in a three-week interval. All of the questions used were obtained from a test-bank. Similar to other studies, he altered the format of questions using ordered, reverse-ordered and scrambled formats. His study found no significant differences among the various test scores in test forms analyzed in his study.

In summary, the empirical evidence concerning the impact of question sequencing is limited and mixed. Chidomere (1989) states, "despite the popularity of item rearrangement in objective examinations among marketing and management science educators, little effort has been made to explore its effect on students' performance. Further research is necessary to explore the issue." Table 1 below summarizes the research completed with regard to question sequencing.

RESEARCH DESIGN AND EXECUTION

This study was conducted at a small, private university in the northeast United States. Multiple-choice question sequencing was manipulated in all exams administered in two sections of Principles of Management, one section of Sales Management, one section of Social Roles in Organizations and two sections of Advertising. The Principles of Management class is a required course for all business majors, while Advertising is a required class for marketing majors. The Sales Management class is an elective for marketing majors and the Social Roles in Organizations is an elective for management sciences majors. Two instructors, one from the marketing department and one from the management sciences department taught the classes. In each case, multiple-choice questions were positioned at the beginning of the exam, followed by other types of questions (primarily problems and short essays). The advertising classes used multiple-choice questions exclusively. Information concerning the number (and percentage value) of multiple-choice questions on each exam and the number of students in each condition is summarized in Table 2. The first instructor taught the two Advertising classes while the second instructor taught the remaining classes. There was only one student registered for more than one of the classes simultaneously; thus the results are not significantly influenced by common subjects across sections.

The same research design was used across all courses. There were three versions of each exam, all of which contained the same multiple-choice questions. In the first condition ("ordered"), questions appeared in the same sequence as the related material was presented in class; in the second condition ("reverse ordered"), the sequence of questions was reversed; and in the final condition ("scrambled"), the questions were randomly ordered. Exams were randomly distributed to students in the first exam. Over the next two exams, students were given the exam versions to which they had not been previously exposed. Therefore, each student was exposed to each type of exam (ordered, reverse ordered and scrambled) throughout the semester. When exams were graded, the exam version was recorded to permit subsequent analysis.

This design was intended to allow for the examination of the impact of multiple-choice question sequencing. In an effort to provide more reliable evidence concerning sequencing effects, we replicated the design across several different courses and used it for each exam administered in each of the courses. However, we were primarily interested in determining whether there was a systematic difference in students' performance on the multiple-choice questions based on the version of the exam taken.

ANALYSIS OF RESULTS

The average performance on multiple-choice questions by class and format is presented in Table 3. In addition, the ranking of each exam format in each class is provided. Although no

TABLE 1
SUMMARY OF MAJOR RESEARCH STUDIES

Author	Course Examined	Findings	Additional Comments
Araujo and Semb 1979	Psychology	Order did not significantly influence student performance	3 aspects of student performance were compared
Balck 1989	Psychology	Students had slightly higher scores with sequential grouping	
Borde 1998	Marketing	Inconclusive results	
Carlson and Ostrosky 1992	Microeconomics	Exam scores may be influenced by the order of exam items; students with scrambled version had lower scores	
Chidomere 1989	Marketing	No significant difference was found	Used test-bank
Flaugher et al. 1968	Education	Differences in results of verbal vs. math scores	5000 SAT tests
Pettit et al. 1986	Marketing	Mixed results	
Sander 1988	Psychology	No difference in random vs. sequential test scores	4 test versions; various orders
Sax and Cromack 1966	Education	Differences in motivational level are the and/or anxiety key determinants of test scores	Higher test when questions are arranged from easy to difficult

definitive conclusions can be established, the results presented here are very consistent. In every course, students generally performed best on the ordered exam versions; further, performance was generally poorest on the scrambled version.

A statistical analysis of the data was conducted to determine whether these differences in performance were significant. ANOVA, using a randomized block design, was chosen to analyze the results of this experiment to control for the variation in student ability. Since each student was exposed to each treatment (format of examination) once and only once, blocking was used to remove the variability due to the blocking variable (student ability). This allowed us to focus on differences among the treatment condi-

TABLE 2
SUMMARY OF EXPERIMENTAL CONDITIONS

Course	# of Exams	Format: # of MC questions/ (% of total exam)		Versions	# of Students
Advertising (2 Sections)	3	100	(100%)	Ordered	51
		85	(100%)	Reverse ordered	51
		200	(100%)	Scrambled	51
Social Roles in Organizations	3	27	(54%)	Ordered	45
		21	(35%)	Reverse ordered	45
		26	(33%)	Scrambled	45
Sales Management	3	26	(43%)	Ordered	12
		32	(45%)	Reverse ordered	12
		30	(40%)	Scrambled	12
Principles of Management (2 Sections)	3	57	(71%)	Ordered	64
		61	(85%)	Reverse ordered	64
		47	(62%)	Scrambled	64

TABLE 3
MULTIPLE CHOICE PERFORMANCE ON EXAMINATIONS BY CLASS AND FORMAT OF EXAMINATION

	Ordered	Reverse Ordered	Scrambled
Principles of Management	75.98% (1)	73.75% (2)	72.25% (3)
Social Roles in Organizations	82.70% (1)	78.70% (3)	79.52% (2)
Advertising	70.31% (1)	70.16% (2)	68.76% (3)
Sales Management	75.84% (1)	74.15% (2)	71.76% (3)
All Courses Combined	76.48% (1)	74.65% (2)	73.13% (3)

tions. The randomized block design also provides a more efficient analysis by reducing the experimental error and increasing the level of precision (Berenson 1989).

In each ANOVA model, the dependent variable is the student's score on the multiple choice portion of the exam (performance). The independent variables are student ability (proxied by the student's mean score on the multiple choice questions across all three exams) and examination format (an indicator variable for "ordered," "reverse ordered," or "scrambled" version). We expect, of course, to find that student ability (the

blocking variable) will be significantly related to performance in every case. However, if the examination format variable is also statistically significant, this indicates that the exam format had a significant impact on student performance after controlling for student ability. This is the focus of our study.

We first analyzed the model including all classes in the experiment. As shown in Table 4, the type of examination used is significantly related to performance $(p < .01)$, suggesting that question sequencing impacts are significant. To determine if the results are driven by a particular course, we extended this analysis on a course-by-course basis.

The individual models for each of the four classes participating in the study are presented below in Tables 5 through 8.

As shown in Table 5, there is a significant effect $(p < .05)$ for the type of examination used in the two sections of Principles of Management. The initial analysis presented in Table 3 indicates that student scores were lowest on the scrambled test versions.

As shown in Table 6, there is also a significant difference $(p < .05)$ for the type of examination used in the Social Roles in Organizations class. The initial analysis presented in Table 3 indicates that student scores in this class were lowest on the reverse-ordered format.

As shown in Table 7, there is no significant effect $(p = 0.41)$ for the type of examination used in the Sales Management Class. The lack of significance in this case could be due, in part, to the limited number of students in the class, resulting in a less powerful test. However, the initial

TABLE 4
ANOVA: TWO-FACTOR WITHOUT REPLICATION
ALL CLASSES

Source of Variation	SS	df	MS	F	P-value
Student Ability	44548.95	172	259.05	3.78	0.00
Examination Format	975.06	2	487.53	7.12	0.00
Error	23544.94	344	68.45		

TABLE 5
ANOVA: TWO-FACTOR WITHOUT REPLICATION
PRINCIPLES OF MANAGEMENT SCIENCES CLASSES

Source of Variation	SS	df	MS	F	P-value
Student Ability	12609.66	63	200.15	2.84	0.00
Examination Format	452.01	2	226.01	3.21	0.04
Error	8877.32	126	70.45		

TABLE 6
ANOVA: TWO-FACTOR WITHOUT REPLICATION
SOCIAL ROLES IN ORGANIZATIONS CLASS

Source of Variation	SS	df	MS	F	P-value
Student Ability	5104.59	44	116.01	2.00	0.00
Examination Format	508.45	2	254.22	4.39	0.01
Error	5094.87	88	57.89		

TABLE 7
ANOVA: TWO-FACTOR WITHOUT REPLICATION
SALES MANAGEMENT CLASS

Source of Variation	SS	df	MS	F	P-value
Student Ability	1574.76	13	131.23	2.21	0.04
Examination Format	109.07	2	54.53	0.91	0.41
Error	1424.92	24	59.37		

TABLE 8
ANOVA: TWO-FACTOR WITHOUT REPLICATION
ADVERTISING CLASSES

Source of Variation	SS	df	MS	F	P-value
Student Ability	14693.72	50	293.87	3.68	1.48E-08
Examination Format	74.15	2	37.07	0.46	0.68
Error	7979.17	100	79.79		

analysis presented in Table 3 indicates that student scores were lowest on the scrambled format.

Finally, as shown in Table 8, there is no significant effect (p = 0.68) for the type of examination used in the Advertising classes. It is more difficult to explain the lack of statistical significance in this case, since there were a larger number of students in this course. However, the mixed results of our study are consistent with previous research. Importantly, however, the initial analysis presented in Table 3 indicates that, consistent with other courses examined, student scores were lowest on the scrambled format for this course.

SUMMARY, IMPLICATIONS, AND LIMITATIONS

The results of research examining the impact of exam test-item sequencing on student performance – both our research and the work of our predecessors – could be described as "mixed." However, some researchers have found relatively strong sequencing impacts at least for student performance in certain courses. In this study, our analysis of a single model including data for four different management and marketing courses found that multiple choice question sequencing has a statistically significant effect on student performance. We also found a statistically significant sequencing effect for both management courses. While the effect of question sequencing was not statistically significant for the two marketing courses, the results of our examination consistently suggest that student performance is highest when questions are presented in sequential order. Further, for 3 of the 4 courses examined, performance was lowest for the scrambled version of the exams.

Thus, after reviewing the results of this experiment, we conclude that the best method of testing is the scrambled format. Although it is the most challenging format for students, it allows for distribution of several scrambled formats with identical questions. Scrambled exams are also easier to construct, less likely to allow cheating and provide a truer measure of student learning since there are no cues tied to the material immediately preceding or following the test questions on the exam. This approach can therefore provide exams that are economical, fair and consistent.

We attempted to increase the generalizability of our findings by extending our investigation over four different courses, representing six sections and taught by two instructors over the course of one semester. However, all of these courses were taught at a single university. Additionally, while multiple-choice questions were presented in the order indicated, we do not know for certain that students completed the test questions in the order presented. Thus, our study is best characterized as quasi-experimental in nature. The inability to control the order in which students complete a paper-and-pencil test introduces additional noise into the data, making it more difficult to detect a statistically significant relationship between exam format and student performance. This may explain the mixed findings in our study as well as other research.

Future research could address some of these limitations and enhance our understanding of the relationship between examination design and student performance. The study could be replicated in additional courses at additional universities in an effort to enhance generalizability. Future researchers could control for the order in which questions are answered by administering computerized exams that do not allow students to skip questions or change their answers. Following the lead of earlier researchers, we intend to extend our analysis to explore the impacts of exam question sequencing on sub-groups of students partitioned based on their cumulative GPAs. Despite the mixed results of research that has examined the impacts of question sequencing on student sub-groups to date, we believe that this may still be a fruitful direction for future inquiries. Partitioning students on the basis of learning styles might also provide interesting insights into the impact of test-item sequencing on student performance.

The issue of test-item sequencing impacts has been discussed in the existing literature primarily as relating to testing "equity," as well as motivated by a desire to produce "unbiased" examinations. This reflects the belief that students' examination performance should not be affected by extraneous factors such as the version of the test that is administered to them. Viewed from this perspective, readers of this body of research may wish to use the results of these studies to help them design examinations such that the results are not affected by the version of the exam administered. Our results suggest that scrambling the order of test items is the best way to achieve this result. Alternatively,

an instructor might keep the ordering of the questions the same but vary the order of the response choices offered. However, it does not appear that the hypothesis that this approach will produce "equitable" results has been subjected to empirical testing in marketing and management sciences contexts. Empirical testing in other contexts, however, suggests that response ordering does not influence exam performance (Aiken 1990).

Despite the growing body of research in the area of marketing and management sciences education, we still clearly do not have a well-developed understanding of the factors affecting student exam performance. However, the results of the research into the impact of exam question sequencing on marketing and management sciences student performance, taken as a whole, seem to indicate that students may organize their marketing and management sciences knowledge in highly specific ways (e.g., in terms of sequen-tial textbook chapters). Students' ability to access and utilize this knowledge may be significantly impaired when they are confronted with questions that do not conform to these sequential structures. However, if marketing and management sciences student knowledge structures – perhaps being developed and enhanced through the marketing and management sciences educational process – are as specialized and constrained as the body of research on intertopical ordering effects suggests, perhaps students are not being well prepared for the professional world that they will be encountering after graduation. We believe that this may offer a very promising avenue for future research.

The limitations of our work notwithstanding, we believe that the research presented here meaningfully extends our understanding in this important area of marketing and management sciences education and provides useful direction for future inquiries.

REFERENCES

Araujo, J. and G. Semb (1979), "Effects of Item Order on Student Performance," paper presented at the 8th Annual Conference of the American Psychological Association, New York.

Berenson, Mark L. (1989), *Basic Business Statistics.* Prentice-Hall, 468–74.

Balch, William R. (1989), "Item Order Affects Performance on Multiple-Choice Exams," *Teaching of Psychology,* 16 (2), (April), 75–77.

Borde, Stephen F. (1998), "Predictors of Student Academic Performance in the Introductory Marketing Course," *Journal of Education for Business,* (May).

Carlson, J. Lon and Anthony Ostrosky (1992), "Item Sequence and Student Performance on Multiple-Choice Exams: Further Evidence," *Journal of Economic Education,* (Summer).

Chidomere, Roland C. (1989), "Test Item Arrangement and Student Performance in Principles of Marketing Examination: A Replication Study," *Journal of Marketing Education,* (Fall), 36–40.

Flaugher, Melton, and Meyers (1968), "Item Arrangement Under Typical Test Conditions," *Educational and Psychological Measurement,* 28, 813–24.

Kleinke, D.J. (1980), "Item Order, Response Location, and Examinee Sex and Handedness on Performance on Multiple-Choice Tests," *Journal of Educational Research,* 74, 225–28.

Pettit, Kathy L., Kenneth G. Baker, and Laura D. Davis (1986), "Unconscious Biasing of Student Examination Scores: A Case of Sequential Versus Random Information Retrieval," *Journal of Marketing Education,* (Fall), 20–24.

Sander, Angelle (1988), "Sequential Order – Scrambled Brains: The Effects of Pretesting Test Items in Versus Random Order," Paper

Presented at the Annual Meeting of the Southwestern Psychological Association, Tulsa, Oklahoma, (April), 21–23.

Sax, G. and T.R. Cromack (1966), "The Effects of Various Forms of Item Arrangement on Test Performance," *Journal of Educational Measurement,* (3), 309–11.