

EFFECTS OF PRACTICE OPPORTUNITIES ON EXAM PERFORMANCE IN  
AN UNDERGRADUATE BEHAVIOR MODIFICATION COURSE

BY

Copyright 2007  
Wesley H. Dotson

Submitted to the graduate degree program in Applied Behavioral Science  
and the Faculty of the Graduate School of the University of Kansas  
In partial fulfillment of the requirements for the degree of  
Master's of Arts

---

James A. Sherman, Co-chair

---

Jan B. Sheldon, Co-chair

---

Daniel J. Bernstein, Committee Member

Date Defended: \_\_\_\_\_

The Thesis Committee for Wesley H. Dotson certifies  
That this is the approved Version of the following thesis:

EFFECTS OF PRACTICE OPPORTUNITIES ON EXAM PERFORMANCE IN  
AN UNDERGRADUATE BEHAVIOR MODIFICATION COURSE

Committee:

\_\_\_\_\_  
James A. Sherman, Co-chair

\_\_\_\_\_  
Jan B. Sheldon, Co-chair

\_\_\_\_\_  
Daniel J. Bernstein, Committee Member

Date Approved: \_\_\_\_\_

## ABSTRACT

All students enrolled in an undergraduate behavior modification course during two semesters participated. During each of five course sections, students could participate in two forms of optional review activities: 1) answering practice questions for feedback and extra credit, and 2) participating in a mock exam where students took a practice exam under “test-like” conditions and evaluated and discussed their answers. Exams were given at the end of each section. On average, section exam performance improved as students completed more practice questions accurately. Students also performed better on section exams if they took the mock exam. The effect of taking the mock exam on exam performance was the most pronounced for students who did not answer the practice questions or who did not answer them correctly. A survey of student opinion indicated that students perceived both practice questions and the mock exam as helpful to them in doing better on examinations.

## EFFECTS OF PRACTICE OPPORTUNITIES ON EXAM PERFORMANCE IN AN UNDERGRADUATE BEHAVIOR MODIFICATION CLASS

### Introduction

Undergraduate education has been a concern of behavior analysts from the earliest days of the field. Interventions to improve student learning have ranged from environmental manipulations within traditional lecture courses (Semb, 1974; Semb, Hopkins, & Hursh, 1973) to the development of entirely new instructional techniques and systems (Boyce & Himeline, 2002; Keller, 1968; Skinner, 1968).

Interventions within the teaching of psychology have aimed to increase the quality of several types of student behavior including note-taking (Baker & Lombardi, 1985; Fisher & Harris, 1973), attendance (Clump, Bauer, & Whiteleather, 2003; Gunn, 1993; Hancock, 1994; Lamdin, 1996; Wilder, Flood, & Stromsnes, 2001), homework completion (Dufresne, Mestre, Hart, & Rath, 2002; Mayfield & Chase, 2002; Miller & Malott, 1997; Ryan & Hemmes, 2005), study behavior (Balch, 2001; Dickson, Devoley, & Miller, 2006; Fleming, 2002; Gurung, 2005; Jenkins & Neisworth, 1973; Kouyoumdjian, 2004; Mawhinney, Bostow, Laws, Blumenfeld, & Hopkins, 1971), and exam performance.

Efforts to improve student's performance on exams have included increasing student engagement with the course materials outside of class periods through the use of study guides (Balch, 2001; Fleming, 2002; Jenkins & Neisworth, 1973), review activities (Dickson et al., 2006; Flora & Logan, 1996; Miles, Kibler, & Pettigrew, 1967; Oliver & Williams, 2005; Semb et al., 1973; Simon, 2005), and

practice tests (Balch, 1998; Bol & Hacker, 2001; Peat & Franklin, 2002). This review will include only those studies which dealt with undergraduate or graduate level students and which reported effects of interventions specifically on exam performance. A potentially relevant literature that is not reviewed here involves the effect of homework assignments on exam performance. Some studies (Dufresne et al., 2002; Ryan & Hemmes, 2005) have demonstrated that homework activities that potentially resemble exam materials may have a similar impact on exam performance as the review activities discussed below. The type of intervention attempted, the format of the exam given, and the effect size (as indicated by the percentage points of difference out of 100) will be discussed for each article.

### *Study Guides*

For this review, study guides were defined as any document or instruction that informed the students about the test format, material to be covered on the exam, performance expectations for the test, or tips for studying. Interventions which included example or practice questions are classified as review activities and will be discussed in the next section.

Two studies (Balch, 2001; Fleming, 2002) have reported on the effects of providing general study tips to students in introductory psychology classes. Balch handed out a sheet containing six study tips on the first day of class. The tips were tailored to help students succeed in his course. Tips advised students to partake in review activities such as group study and self-help quizzes and also recommended tracking their performance in the course using a provided chart. Balch also spent

fifteen minutes reviewing the tips with the students during the first day. Students reported that the tips were helpful, but the effect of the tips on exam performance was unclear. Balch reported a weak correlation between students' self-report of using more of the tips with a very small increase in exam performance, but he did not report the exam scores nor provide any information regarding the format of the exams he gave. Fleming (2002) reported a similar attempt at improving performance by teaching general study skills. At the beginning of the semester, she taught the skills to one of two sections of her Introductory Psychology course. She spent a few minutes during each of the first six days of class introducing skills such as how to review class materials, reading tips, and ideas for studying for course exams. Students were also asked to formulate study goals and to report on how accurately they had met their goals using an activity sheet. The goal and activity sheets were presented to the students at the beginning of each section of the course. She examined the effect of teaching general study skills and associated activities by comparing the exam scores of those students who completed those activities to students in another section who had not participated in these activities. Students within each section were divided into two groups (first year and upper class students). On the multiple choice exams given in the course, the average grades were uniformly high (above 80%) for all groups except for the first year control students on the first two exams (who averaged around 70% correct). By the third and fourth exams, there were no significant differences across the four groups of students. The results imply that the study guide and explanations were helpful for

first year students who may not have known how to study for classes in general. The results also suggest, however, that experience in the course itself appeared sufficient to teach those skills since the first year control students were matching the other groups' exam performances by the third exam.

Knowing general expectations may help students prepare for exams, but knowing specific expectations for what students may see on a single exam may be more helpful. One study (Jenkins & Neisworth, 1973) presented objectives for a reading test given to students in an introductory educational psychology course in an effort to determine the effect of providing specific and accurate learning objectives for an exam on the performance of students taking that exam. A group of 55 students with experience using study guides was randomly divided into two groups. On one Thursday class period, each group was given ten learning objectives for a set of two readings. An example of a learning objective was, "Given five attributes, identify one which is not, according to the article, one of the four basic attributes of criterion measurement." For each group, five of the expectations they were given corresponded with questions that would be asked on the exam and five did not. On the following Tuesday the students were given an exam over the material. There were ten multiple-choice questions on the exam, and each corresponded to an accurate objective presented to half of the students. One group of students received accurate learning objectives for questions 1, 3, 5, 7, and 9 while the other group received accurate objectives for questions 2, 4, 6, 8, and 10. Students scored higher on the questions for which they had received accurate learning objectives than for

questions which they had not received accurate learning objectives. The difference in accuracy ranged from 2-55 percentage points of difference and the differences were statistically significant for 7 of the 10 questions. The results demonstrate that when students have experience with receiving learning objectives that are predictive of questions that will be asked on the exam, they are more likely to be accurate on subsequent questions for which those objectives have been provided. The results also suggest that providing specific learning objectives for each section of a course has the potential to produce more benefit to students than the general study tips described above (Balch, 2001; Fleming, 2002) as evidenced by larger improvements in exam performance.

#### *Review Activities*

Review activities are any activities that provided students with opportunities outside of class periods or with the assistance of notes to answer questions or complete tasks that were identical to or similar to the questions that would be on the exam.

In one study (Dickson, Miller, & Devoley, 2005), 236 students in two sections of an introductory psychology course took four exams during the course of a semester. In both sections, the instructor made available a comprehensive textbook study guide that contained review activities in a number of formats ranging from true/false to essay questions. The study guide covered some of the material that would be on the exams (specifically, the information from the textbook). Approximately 22% of the questions on each exam came directly from the study



guide. The other 78% of the questions on each exam tested for knowledge of material covered in classroom lectures and activities. In one of the course sections, students were required to complete the study guide to earn 8% of their course grade. In the other section, students had the option to complete the study guide, but it was not required. While the students who were required to complete the study guide performed better on average for each of the four multiple-choice exams (77.96 versus 75.16), there was only 2.8 percentage points of difference on average, and the largest difference on a single exam was only 5%. For the section in which the study guide was required, there was no difference in performance for students who completed 25% versus at least 75% of the study guide. The small effect size may be a result of the structure of the review activities. The authors report that a large number of the study-guide questions were either irrelevant to the material that would be on the exam or were presented in a different format (true/false, essay) than the questions in the exams. Thus, students may have learned that the study guide was not a good predictor of the exam material, especially since only 22% of the exam questions came directly from the study guide, whereas 78% of the questions on the exam came from material presented in class. Thus, students may not have spent a great deal of effort completing the study guide questions when preparing for the exam since other review activities such as re-reading lecture notes may have been more helpful in preparing for the actual exam. Another reason for the small effect size may be the relatively small percentage of the exam questions for which the study guide prepared students. Since the study guide only covered 22% of the

questions presented on the exam, any differences in exam performance as a result of completing the study guide would only be seen in the answers students provided to that small number of questions. The authors do not report a separate analysis of student performance on only those questions from the study guide.

Another important factor in the effectiveness of review activities may be the accuracy with which they are completed. A study by Oliver and Williams (2005) found that when students completing a multiple-choice set of practice questions earned credit based on their accuracy in answering questions their performance was 4-6 percentage points better on the actual multiple-choice exam than when they got credit simply for completing the practice questions. Since there was no comparison group who did not complete any practice questions, it is difficult to determine the overall effectiveness of the practice questions in that course.

Flora and Logan (1996) used a computerized study guide that allowed students to complete review questions online and print out summary “report cards” to document their activity. The study guide contained questions in a variety of formats such as true/false, fill-in-the-blank, and multiple-choice. Students were required to complete a study guide with 85% accuracy prior to two of a total of four multiple-choice examinations in a general psychology class. Students were not required to complete a study guide before the other two exams, although the computerized study guide questions were available before all four exams for all four sections of the course. The exams that were preceded by required completion of the study guide were counterbalanced across four sections of the same course.

Questions for the exam were selected from a test bank provided by the textbook publisher, and the same test bank was also used to select questions for the computerized reviews. It is unclear whether all of the questions on the exam were also among those in the study questions, if some of the questions were used in both the study questions and the exam, or if none of the study questions appeared on the exam. Nevertheless, the authors report there was a statistically significant difference between exam scores when accurate completion of the study questions was required prior to the exam versus when it was not. The difference, however, was small, roughly 2 percentage points of difference between groups. As noted above, the study guide was available to the students prior to all of the exams. This may have diminished the size of the effect of the intervention. Requiring students to complete the review accurately, then, did not seem to lead to meaningfully improved performance on the exam. It is also possible that students did not benefit as much from a study guide that was not highly similar in format to what they would see on the actual exam. The study guide contained questions in a variety of formats such as true/false, completion, and multiple-choice while the exams were purely multiple-choice. It is also possible that mastery of material at the 85% correct criteria using, for example, true/false questions may not have been as easily generalized to demonstration of the same mastery when the questions were asked in the multiple-choice format.

Miles, Kibler, and Pettigrew (1967) taught a beginning educational psychology course that was divided into six units. There were 320 students enrolled

in the course, and on the first day of class the students were randomly assigned to one of six groups. Each group was given a pre-test. Each pre-test was identical to one of the six unit exams, so that at least fifty students had taken each unit exam as a pre-test. The pre-test scores for those students were then compared to their scores on the actual unit test following instruction. The unit tests were the post-tests. Unit exams contained fifty multiple-choice questions each. Prior to the second, fourth, and sixth unit exams, the instructor provided students a study guide that contained eighty multiple-choice practice questions. For those three units, fifty of the study questions were identical to the fifty questions used on the multiple-choice unit exam and on the pre-test. For the first, third, and fifth units, students were simply provided a list of topics to be covered in that unit, a reading list, and the date of the unit exam. The average improvement between pre-test and post-test for each of the six groups of students was compared. For the three groups of students who completed pre-tests and unit tests when the review questions were not provided prior to the unit tests, the average improvement was roughly 18 percentage points. For the three groups of students who completed pre-tests and unit tests when the review questions were available prior to the unit test, the average performance improvement was over 30 percentage points. Thus, there was a larger average improvement in exam performance (12 percentage points of difference) for those groups who were given study questions prior to the unit exam for which they had taken the pre-test than for those groups for which the study questions were not available. A potential confound in the results is that the differences in exam performance may have been a result of

differing difficulty levels of material presented in the six units of the course. Nevertheless, the results of this study do suggest that access to specific study questions that are directly relevant to what students can expect to see on the exam may have a larger effect on exam performance than more varied or less relevant review materials such as the study guides discussed above (Dickson et al., 2005; Flora & Logan, 1996) that contained a large number of review activities in multiple formats that did not accurately represent the majority of questions students would see on the exam.

Simon (2005) found that students performed better on exam questions that had been practiced during two types of review activities than on exam questions that had only been discussed in class. In the first of two experiments, students were given the opportunity to complete two review activities: a daily task in class involving answering a single fill-in-the-blank question from the day's lecture, and a study guide at the end of the unit which contained questions very similar to the daily tasks that asked for the same information but used slightly different phrasing. When compared to accuracy on control questions that were only discussed in class, students were 20% more accurate on questions that had been reviewed. In addition, students continued to answer daily questions and complete study guides even in the absence of contingencies rewarding doing so. The second experiment in the study explored whether use of the daily practice questions together with completing the study guide was any more effective than completion of the study guide alone. The results indicated that there were no differences in accuracy on tests preceded by answering

the daily questions and completing the study guide than tests preceded only by completion of the study guide. In both experiments the study guide questions and the daily questions were very similar in content and identical in form (fill-in-the-blank) to the questions students saw on the actual exams. These results, like those of Miles, Kibler, and Pettigrew (1967), also showed a larger effect on exam performance for a study guide that more closely resembled the actual exam than that seen for study guides (Dickson et al., 2005; Flora & Logan, 1996) that less closely resembled the actual exam.

A study by Semb, Hopkins, and Hursh (1973) suggests that review activities that closely resemble the actual exam, and to which students have multiple exposures, lead to more improved performance on the exam than review activities with which students have less experience. The authors provided fourteen multiple-choice study questions to all forty-six students at the beginning of each of twelve units of an introductory child development course. In addition, a twelve-question, multiple-choice quiz was given following completion of each unit. Four, hour-long exams were given throughout the semester. The hour-long exams each contained forty-eight multiple-choice questions. Some of the questions on the hour-long exams had appeared in both the study questions handed out at the beginning of each unit and on the quizzes given out after the unit was completed. Other questions on the hour-long exams had only appeared in either the study questions or the quizzes, but not in both. Finally, some of the questions had never appeared in either the study questions or the end-of-unit quizzes. On hour-long exams, the questions that had

been on both the study questions and on an end-of-unit quiz were the ones answered correctly most often. Students were also more accurate on questions that had been asked before in either the study questions or on quizzes than on questions that they had seen for the first time on the exam. Questions seen twice before were answered at least 25% more accurately on each exam than those questions appearing for the first time on the exam, and between 4-20% more accurately than questions seen only once before.

### *Practice Tests*

Practice tests were defined as any activity in which the students answered questions identical to or similar to the questions that would be on the exam in a classroom setting and without the assistance of their notes. The effects of practice exams on actual exam performance have been mixed.

In an article describing several review strategies, Peat and Franklin (2002) reported the use of a mock exam to assist students in preparing for exams. They administered the mock exam to students during class time, and then gave them access to grading materials online to evaluate their answers. Only 43% of students chose to take the mock exam, and of those students, only 37% graded their answers online afterward. Nevertheless, students ranked it as helpful in surveys handed out subsequently. The authors report that there was a positive correlation between taking the mock exam and actual exam performance, but no data are presented. With no data, it is impossible to determine the effect the mock exam may have had on student

exam performance, but the student rankings suggest students found the option helpful.

Bol and Hacker (2001) compared the exam performance of students who took a practice test to the exam performance of students who took part in a more traditional review. Students in one section of a graduate-level introductory research methods course took a practice test during the class periods before both the course midterm and final exams. Students in the other section engaged in a traditional review during the class periods before the midterm and final. The two practice tests each contained 25 multiple-choice questions and five short-answer questions. The practice test questions were selected from the same test bank as the actual exams and were matched according to difficulty with questions from the actual exams, but the items were not identical. Students in the practice test section of the course took the practice exam during the class period before each exam. Students were allowed an hour to complete the exam, and then 45 minutes to grade and discuss their answers with the professor. Students scored their answers by viewing the correct answers on an overhead projector. On the day before an exam in the traditional review section, students were shown an outline of the section materials, listened to a lecture that reviewed the section, and were able to ask questions about the material or about the upcoming exam. The actual midterm and final exams each contained twenty-five multiple-choice questions and five short-answer questions. The questions were selected from the same test bank as the practice test questions, but they were not identical. The only significant differences between the two groups in terms of exam



performance were that the practice test group scored slightly lower on the multiple-choice items on the midterm exam than the traditional review group. No other differences were observed between the groups on the short answer portion of the midterm or on either portion of the final. Even though there were no large differences between the students receiving the practice exams and the students receiving the more traditional reviews, the students in the practice test section ranked the practice test as the most helpful review activity in which they participated.

Balch (1998) compared the performance of students who were exposed to study questions with that of students who were exposed the same study questions and who also had to answer those questions without the assistance of their notes during a practice test. A group of volunteers from the same two sections of an introductory psychology course were randomly assigned to one of two groups and matched according to their all-but-final point totals in the course so far. Both groups were given practice questions on the last class period of the semester, one week before the final exam in the course. Both the study questions and the final exam were multiple-choice and questions were selected from the same test bank but were not identical between the two. The two groups met at different times on the same day to receive the study questions and to complete some review activities. During the class period that the study questions were handed out, one group was exposed to the study questions, and in order to encourage them to read all of the questions, was asked to complete a questionnaire ranking the likelihood of the questions appearing on the coming exam. The other group was asked to answer the study questions without aid

of notes. Following that, the groups were shown the answers to the practice questions on an overhead projector in order to review the material. The group of students who completed the practice tests scored higher on the last exam than the students who were simply exposed to the questions regardless of their achievement in the class up to that point, although the largest difference between students was for those with the lowest achievement entering the exam. The improvements in exam performance between those reading the questions and those taking the practice test were 4.4 percentage points difference for low-achieving students, 3.9 percentage points for medium-achieving students, and 3.5 percentage points for high-achieving students. The small differences may be a result of the availability to both groups of the study questions and the time spent providing both groups of students with the correct answers. Given the results obtained by Miles, Kibler, and Pettigrew (1967) and Simon (2005) that students provided with study guides closely resembling the format and content of the actual exams perform better on those exams, it may be that both groups benefited from the exposure to the study questions, especially given that students were able to take their study questions with them from the review to use in preparing for the exam. Also, a larger difference might have been seen between the two groups if the final exam would have been given immediately following the review day rather than one week later.

The results of the studies described above when students are exposed to study guides indicates that when the information provided in the study guide is very similar to what the students will see on the exam (Jenkins & Neisworth, 1973), the apparent

effect on exam performance is greater than that seen when the study guide contains less exam-specific material (Balch, 2001; Fleming, 2002). The results of studies involving review activities also indicates a potentially greater effect for activities that more closely resemble the actual exam in format (Miles et al., 1967; Semb et al., 1973) or content (Simon, 2005) than for those that are less relevant (Dickson et al., 2005) or in different formats (Flora & Logan, 1996). The results of studies reviewed in the section on practice exams are more ambiguous. Having students complete practice exams that contain questions that are very similar in form and content to the actual exam appears to be associated with only small improvements in exam performance. The smaller magnitude of the effect in the practice exam studies as opposed to several of the studies involving study guides and review activities may be the result of two things. One possibility is that in the studies in which there was the largest effect of the interventions on exam performance (Jenkins & Neisworth, 1973; Miles et al., 1967; Semb et al., 1973), the questions contained in the study guide or review activity were exactly the same as those that appeared on the actual exam, whereas in the studies using practice tests (Balch, 1998; Bol & Hacker, 2001) the questions were similar but not identical. A second possibility is that simple exposure to the questions on the practice exams was sufficient to improve exam performance in both of the studies reviewed above (Balch, 1998; Bol & Hacker, 2001). In both studies, all of the exams for which data are presented were preceded by students being given the practice exams, but for only some of the exams were students required to complete them.

Many of the studies reviewed above used only multiple-choice exams. Little work has been done to examine the effects of study guides, review activities, and practice tests on exam performance for exams given in other formats such as short-answer or essay. The current study adds to the literature by exploring the effectiveness of practice opportunities at improving test performance on essay exams. In addition, it seeks to provide additional empirical evidence regarding the possible effectiveness of practice tests in improving actual exam performance. The intervention involved offering students multiple opportunities to engage with and apply course material with activities and settings designed to be similar to what they would see when they took the actual section exams.

## Method

### *Participants*

The participants were 182 students who enrolled in an upper-division undergraduate course in behavior modification (Principles and Procedures of Behavior Modification and Therapy) during the spring and fall semesters of 2006 and who took at least one exam during the course of the semester. Ninety students during the spring semester and ninety-two during the fall semester participated.

### *Structure of the Course*

The course was divided into five sections, each covering some aspect of behavior modification and therapy. The content became progressively more complex as the semester went on. The first section dealt with the definition and measurement of behavior, the second section with teaching new behaviors, the third

with reducing problem behaviors, the fourth with behavioral contracting and other more complex procedures such as systematic desensitization, and the fifth with token economies and the legal and ethical considerations of using behavioral techniques. Each section required eight to nine 50-minute class periods to complete. Students were also required to read several chapters from the textbook (Martin & Pear, 2007) throughout each section. At the beginning of each section, an outline of the section and a copy of a practice exam were made available to students online. At the end of each section, students took an exam over the material covered in the preceding section.

Section exams were worth thirty points apiece, and consisted of essay and short-answer questions. The majority of questions (80-90%) required students to apply the principles of behavior and the techniques discussed in lectures to address an applied problem. For example, during the second section, students were given a scenario describing a young man with disabilities who struggled with independence and they were then asked to describe in detail how they would teach him to engage in some type of self-care behavior. Short answer questions from the book and from lecture materials accounted for 10-20% of the credit available for each exam.

In addition to the lectures and class discussions, students had opportunities to participate in two forms of review activity during each of the course sections. One involved answering practice questions over the course material and submitting those answers for feedback, and the other was attending and participating in a mock exam where they took the practice exam under “test-like” conditions.

For each section of the course, there was a practice exam. The practice exam consisted of a description of a clinical situation and some questions. Some practice exams had more than one clinical situation. Each question required students to use the information presented in the lectures, chapters, and discussions to develop one or more behavioral solutions to the issues described in the situations. The practice exam for each section was posted electronically online on the first day each section of the course began. Individual questions about the practice exam were posted online two to five days before the answers were to be submitted.

The primary difference between the practice exams and the actual exams given in class for each section was the description of the situations to which the principles of behavior and the specific techniques discussed in the section needed to be applied. The questions asked on the practice exam and on the actual exam were essentially the same.

For example, a question on both the practice and actual exam for the third section was, "Describe how you would use extinction to reduce the client's problem behavior." The client described in the practice exam might be a typically developing young man who threw tantrums at home whenever his parents did not pay attention to him or give him what he wanted. On the actual exam, the client might be a young man with disabilities who became aggressive and destructive at school whenever he was asked to complete tasks or to leave a preferred activity. In order to receive full credit for their answers, students had to not only know what extinction was –not allowing the consequences maintaining behavior to follow that behavior- but how

that would be accomplished in each case. For the first case, extinction would involve not paying attention to the child or giving him what he wanted when he threw a tantrum. For the next case, extinction would involve continuing to present requests to complete the task and to transition away from the preferred activity even if the young man became aggressive or destroyed materials. Thus, a correct answer for one of the applied situations would not necessarily be a correct answer for the other, even though the question was the same.

Questions from the practice exam were assigned between four and six times during each section. Students could turn in answers online for grading and feedback from the course TA. Answers were usually due in the morning just prior to the beginning of the next class period following the posting of the assignment. Students who submitted answers received a copy of the grading key for those answers and also individual feedback in the form of the amount of credit they earned on each portion of their answer. Students who turned in these answers on time could earn up to three extra-credit points for each section of the course. These extra credit points were added to the points earned on the section exams and counted towards the final course grade. The amount of extra credit students earned for completing practice questions depended on the correctness and thoroughness of their answers. Thus, a student who turned in all of the assignments, but was only 50% correct, received 1.5 extra credit points for the section, while a student who turned in complete and correct answers to all practice exam assignments earned 3 extra credit points for the section. Thus, it was possible for students to earn up to 15 extra credit

points during the semester. Since there were 150 points possible in the course (30 points for each of 5 section exams), extra credit could add up to 10% to a student's final grade in the course.

Students also had the option to participate in a mock exam. The course TA (and lead author) led each session. Each mock exam lasted between two and three hours and was held in a classroom on campus two days before the section exam. The mock exam for the fifth section during the spring semester of 2006 was held nine days before the section exam due to an existing exam schedule during finals week.

Mock exams were divided into three parts: an introduction, the administration of the mock exam itself, and the grading and discussion of answers. First, during the 5-10 minute introduction, students were told they would be taking the mock exam under "test-like" conditions (without notes, working independently, and under time limits) and were given a general description of the mock exam. Second, students were given a paper copy of the mock exam (which included the same situations and questions as the ones posted online in the practice exam for that section), and they were asked to write their answers within forty-five to fifty minutes. Following this, the students were given a short break. Third, the TA handed out a grading key for the mock exam and discussed the grading criteria question by question. Students were asked to grade their answers using the grading key and were also encouraged to volunteer answers for discussion and analysis by the group. After discussion of answers for each question, the TA displayed a correct sample answer on an overhead projector, or via a PowerPoint slideshow, to help illustrate the grading criteria for



each question. The third portion of the mock exam lasted between sixty and ninety minutes. At the end of the discussion, the students were allowed to take their answers and their grading criteria home for further study.

Because the fourth and fifth sections of the course covered a larger and more diverse amount of material, the practice exams for those sections were longer. For example, in the fourth section students learned about several different types of behavioral contracts and were required to prepare to write each type of contract for the exam. While the actual exam asked students to describe how to create only one type of contract in reference to an applied situation, the students were not told in advance which contract they would have to describe in full. So that they had the opportunity to prepare for all the possible types of behavioral contracts, the practice exam posted online asked students to create each type of contract for different situations.

It would not have been possible for students to write out answers for each type of contract within the 45 minutes allotted during the mock exam. Therefore, there were three different versions of the mock exam, each requiring students to describe a different type of behavioral contract for a particular situation. All three versions of the mock exam required students to answer all of the short answer questions from the online practice exam. Students were randomly given one of the three versions. During the discussion period of the mock exam, all of the possible contracts that might appear on the exam were discussed. For the fifth section there were also three versions of the mock exam, and again, some students completed each

version of the mock exam and all three versions of the mock exam were covered during the discussion period of the mock exam.

### *Dependent variables*

The primary measure in this analysis was student performance on the actual exams taken in class. Student attendance at mock exams and completion of practice questions were also recorded in order to examine whether they were correlated with student exam performance.

### *Reliability*

Ten percent of the total exams given during both semesters were graded by one of the course professors (the second and third authors) in order to determine the reliability of grading. Grading tables were used by both the Graduate Teaching Assistant and the professors to evaluate student answers. Reliability was calculated by determining the total number of points on which the two graders agreed for each question on the exam. The total points of agreement were then added together and divided by thirty to determine the overall percentage of agreement for each individual exam. Reliability percentages were calculated for 10% of the exams for all five exams in both semesters. Overall reliability was 92.80% for both semesters (range 90.08-95.94%).

## Results

### *Exam Performance*

The average scores on the actual exams taken by students during the spring and fall semesters of 2006 are shown in figures 1-5. These scores do not include any

extra credit students earned by turning in answers to practice questions. The test data from both semesters were combined because the effects from both semesters were highly similar. Table 1 shows the total number of exams in each of the comparison conditions.

Figure 1 shows the exam performance of students who did or did not take the mock exam for each section. Across all five exams, students who took the mock exam scored higher on the actual exam. There was a 13 percentage point difference in exam performance on average between students who did and did not take the mock exam.

In order to try to factor out the possible effect of completing practice questions on exam performance, only the exam scores of students who did not complete any practice questions for a section were then analyzed. Figure 2 summarizes those results. Again, students who did not complete any practice questions, but did take the mock exam, scored higher on the exams for each section (a 15 percentage point difference), although their scores were on the average lower than students who both took the mock exam and completed some practice questions.

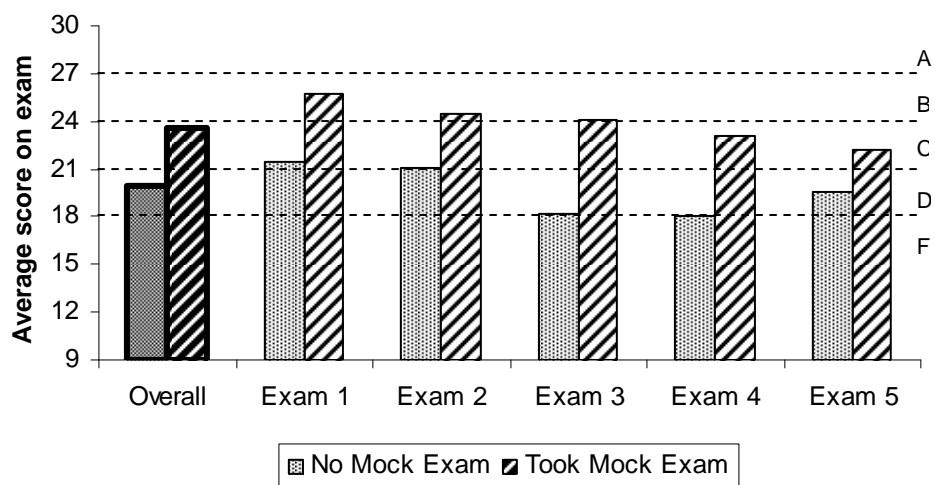
Figure 3 shows the effect of completing practice questions on exam performance. Students earned extra credit based on the accuracy of their answers, and, as the amount of extra credit earned increased, so did their average exam score.

To try to factor out the effect of the mock exam on exam performance, only the scores of those students who did not complete the mock exam were analyzed.

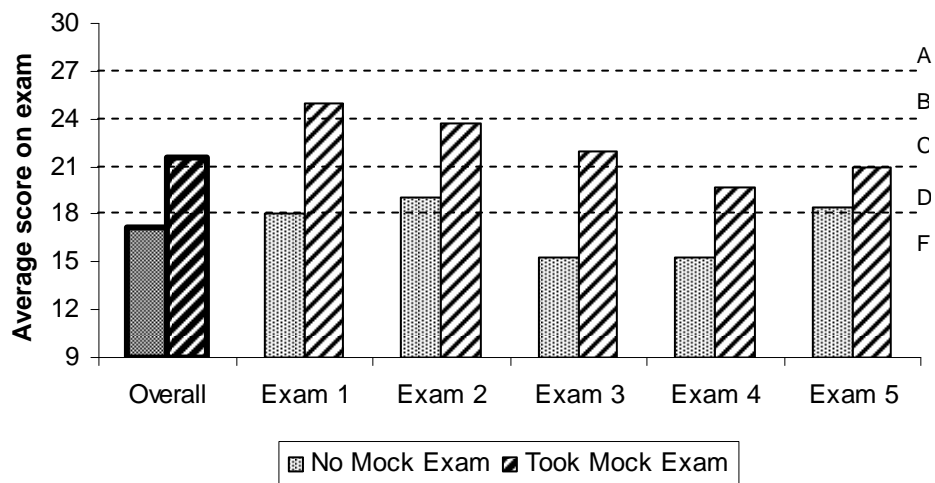
Table 1

*Number of Exams in Each Comparison Condition*

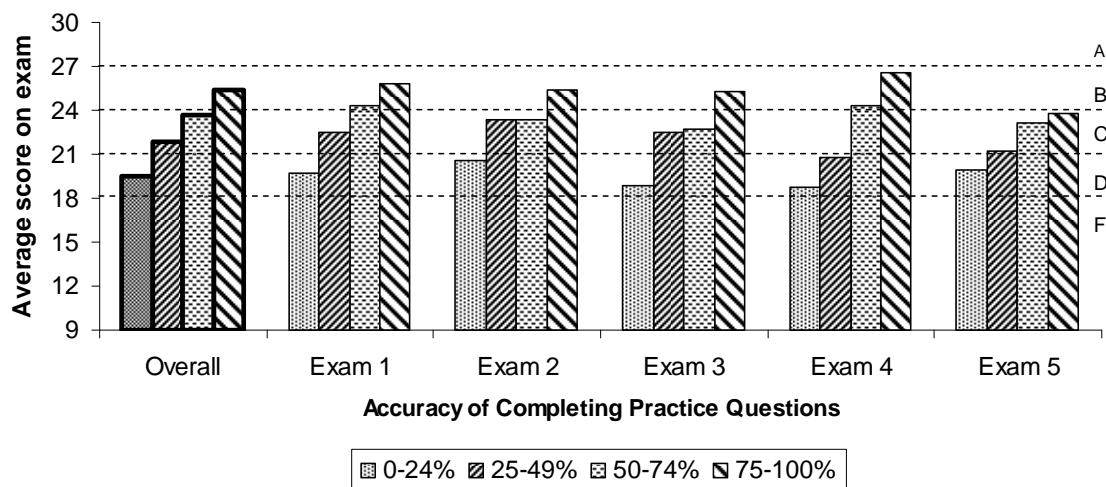
<u>Amount of Extra Credit and Mock Exam Attendance</u>	<u>Number of Exams</u>
Student earned 75-100% of extra credit:	
Attended mock exam	55
Did not attend mock exam	71
Student earned 50-74% of extra credit:	
Attended mock exam	73
Did not attend mock exam	79
Student earned 25-49% of extra credit:	
Attended mock exam	61
Did not attend mock exam	88
Student earned 0-24% of extra credit:	
Attended mock exam	188
Did not attend mock exam	259



**Figure 1:** Average performance on section exams of students who did and did not take the mock exam for that section



**Figure 2:** Average performance on section exams of students who did and did not take the mock exam for that section and who earned no extra credit



**Figure 3:** Average performance on section exams of students by the amount of extra credit they earned for that section (Extra credit based on accurate completion of practice questions)

Figure 4 shows those results. Students performed better on section exams when they earned more extra credit during that section, but their scores were on average lower than those students who both completed the mock exam and earned an equivalent amount of extra credit.

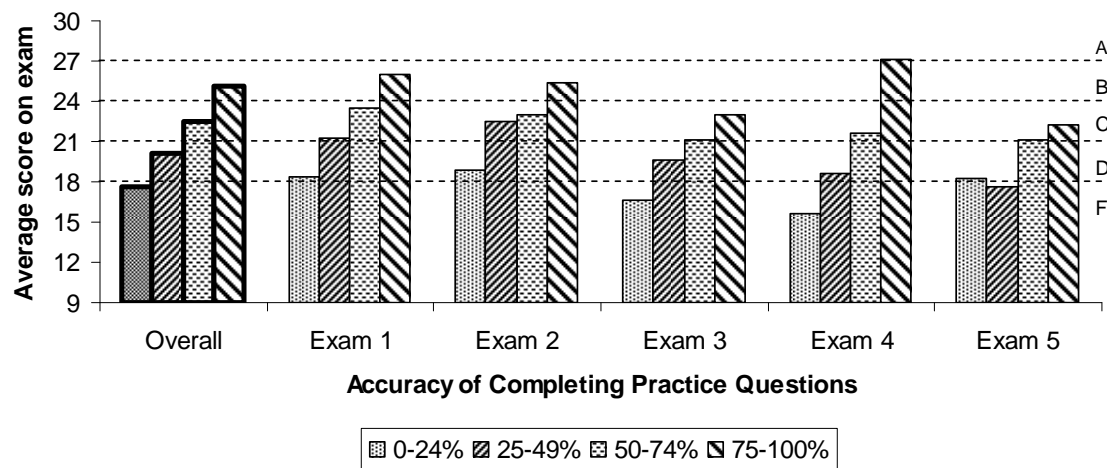
Figure 5 shows the interaction between taking the mock exam and answering practice questions accurately. When matched for the amount of extra credit earned, students performed better on the actual exam if they took the mock exam, however the effect of taking the mock exam on test performance was the most pronounced (15 percentage points difference) for those students who earned the least amount of extra credit on the practice questions.

### *Social Validity*

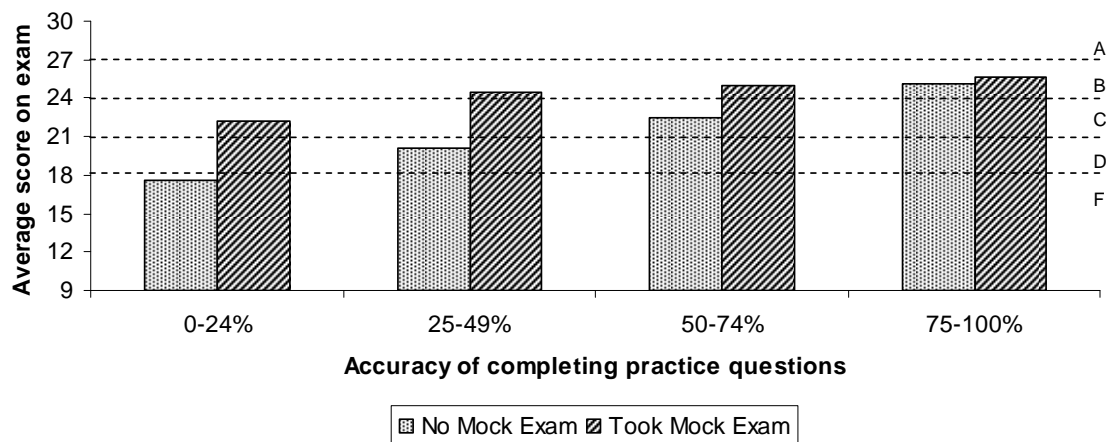
Two surveys were handed out at the end of the spring 2006 semester to assess student opinions about the review options available to them. Both were anonymous and students were not required to complete them. The first survey was handed out to all of the students in attendance on the day teacher evaluations were completed. The survey contained several short questions about the practice and review options available to students throughout the semester. The second survey was given out to students attending the fifth mock exam and only contained questions relevant to the mock exam.

The survey handed out in class contained questions asking if the students had utilized any of the options for submitting answers to practice questions and to take





**Figure 4:** Average performance on section exams of students who did not take the mock exam by the amount of extra credit they earned for that section (Extra credit based on accurate completion of practice questions)



**Figure 5:** Average exam score of students earning different amounts of extra credit (by completing practice questions) and also divided by participation in the mock exam.

mock exams and then asked the students to rate the helpfulness of each of those options on a scale of 1-5 (1= waste of time, 3=helpful, 5=vital). A summary of the survey results is shown in Table 2. Students rated both the practice questions and the mock exams as being helpful, with the mock exam receiving the higher rating.

The survey handed out at the fifth mock exam consisted of four questions about the mock exam, and had an area for student comments. Table 3 summarizes the survey results. Students again rated the mock exam as very helpful.

The in-class survey was handed out again at the end of the Fall 2006 semester. Table 4 summarizes the results. As in the spring semester, students ranked both the practice questions and the mock exams as helpful, with the mock exam receiving the higher rating.

Student attendance at the mock exams provided another indication of the helpfulness of the mock exam. Attendance during both semesters increased over the course of the semester. Figure 6 shows the increasing percentage of students who participated in the mock exams across both semesters. Since there was no extra contingency on attendance such as extra credit points, it is assumed that students came because the sessions were helpful.

Figure 7 shows the percentage of students earning extra credit on each of the five section exams across both semesters of the course. There was not an overall increasing trend in the number of students completing the practice questions accurately, but at least 50% of the students who took each exam earned at least some of the available extra credit for that section.

Table 2

*Student classroom survey results –Spring 06*

Survey Question Ranking	Student
Question 1: Did you complete any of the optional practice questions?	Y: 54 N: 2
Question 2: How helpful did you find the practice questions? (1= waste of time, 3=helpful, 5=vital)	Mean: 3.83 Mode: 4 Median: 4
Question 3: Did you complete any of the mock exams?	Y: 34 N: 22
Question 4: How helpful did you find the mock exams? (1= waste of time, 3=helpful, 5=vital)	Mean: 4.62 Mode: 5 Median: 5

Table 3

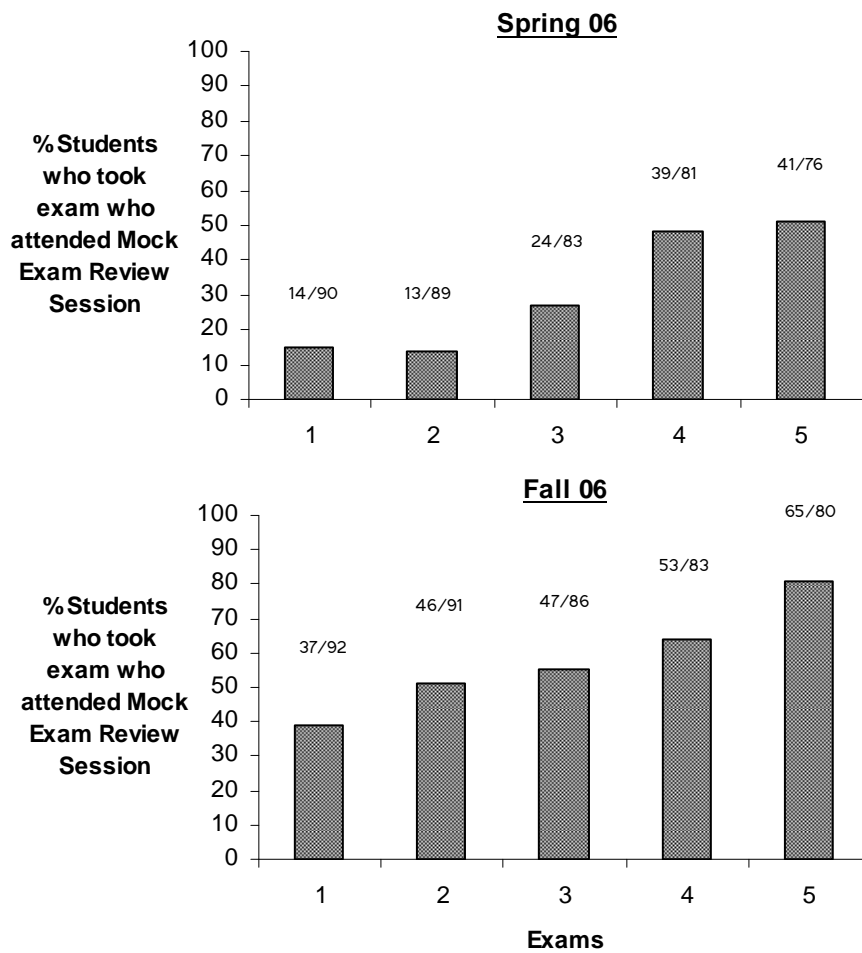
*Mock exam survey results –Spring 06*

<u>Survey Question</u>	<u>Student Ranking</u>
Question 1: Overall, How helpful did you find the mock exams? (1= waste of time, 3=helpful, 5=vital)	Mean: 4.59 Mode: 5 Median: 5
Question 2: How helpful did you find taking the practice exam under “test-like” conditions? (1= waste of time, 3=helpful, 5=vital)	Mean: 3.78 Mode: 4 Median: 4
Question 3: How helpful did you find reviewing the grading criteria? (1= waste of time, 3=helpful, 5=vital)	Mean: 4.58 Mode: 5 Median: 5
Question 4: How helpful did you find reviewing sample answers? (1= waste of time, 3=helpful, 5=vital)	Mean: 4.66 Mode: 5 Median: 5

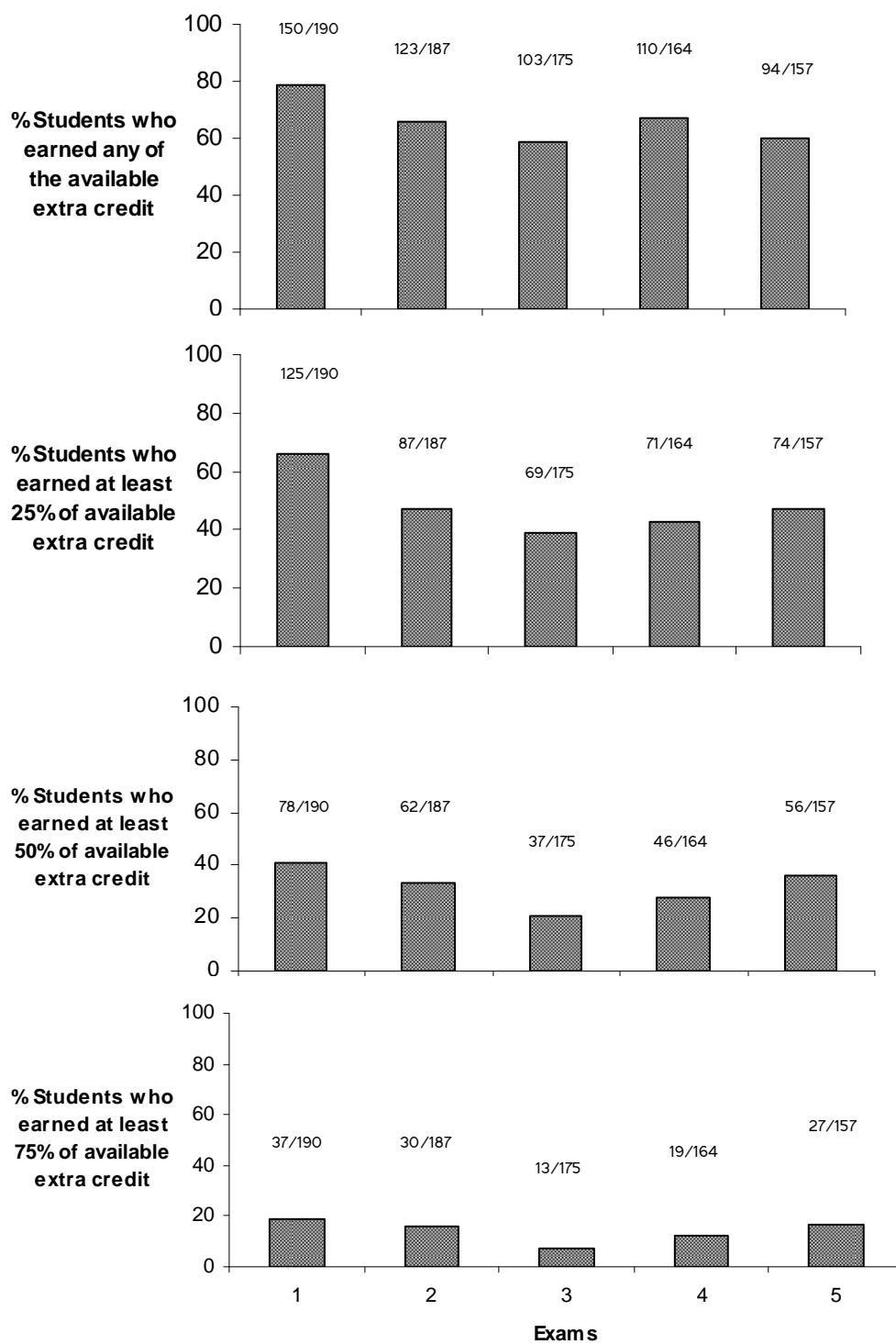
Table 4

*Student classroom survey results –Fall 06*

<u>Survey Question</u>	<u>Student Ranking</u>
Question 1: Did you complete any of the optional practice questions?	Y: 51 N: 2
Question 2: How helpful did you find the practice questions? (1= waste of time, 3=helpful, 5=vital)	Mean: 3.88 Mode: 4 Median: 4
Question 3: Did you complete any of the mock exams?	Y: 43 N: 10
Question 4: How helpful did you find the mock exams? (1= waste of time, 3=helpful, 5=vital)	Mean: 4.76 Mode: 5 Median: 5



**Figure 6:** Percentage of students who took actual section exam who took the mock exam for that section



**Figure 7:** Percentage of students earning extra credit (by completing practice questions accurately) on each of the five section exams across both semesters.



## Discussion

The results from this study indicate that both completion of practice questions and participation in a mock exam were associated with positive effects on average student performance on section exams. Students who answered more practice questions correctly scored higher than students who answered fewer or none, replicating the effects seen by Oliver and Williams (2005). Students who participated in the mock exam for a section scored higher on the section exam than those who did not. The students who performed the best on the exams were those who earned at least 75% of the extra credit points and also attended the mock exam for the section. The students who performed worst were the ones who completed neither practice questions nor the mock exam.

The size of the effect (26.76 percentage points difference between students completing both the mock exam and 75% of practice questions and those completing neither) may be a result of improved generalization of student responding. For responding to generalize, it must occur. The students who correctly answered practice questions and attended the mock exam had answered the questions they would see on the actual exam at least twice before entering the examination room. Students who had practiced more often scored better than students who had practiced once (by only completing one of the review activities) and students who had not practiced scored worse than either. This is very similar to Semb's (1973) finding that the questions students had seen twice before were answered more accurately than those only seen once or not at all. The results of Miles, et al. (1967) also

suggest a role for familiarity –and thus increased likelihood of generalization- in producing larger improvements in exam performance. The review activities they gave students contained the exact questions that students would see on the exam and also were presented in the same format. Students performed 12 percentage points better on exams on which they had seen the questions before than they did on exams for which they had not been given access to questions. The weaker effects seen in other studies (Balch, 1998; Bol & Hacker, 2001; Dickson et al., 2005; Flora & Logan, 1996) may be partially a result of the review activities not so closely resembling the actual exam experience.

In the current study the questions, general issues, and grading criteria were the same for the practice, mock, and actual exams, but the different situations presented in the actual exams required somewhat different answers than those required in the review activities. In the earlier studies (Miles et al., 1967; Semb et al., 1973) the questions and answers on the exams were identical to those seen in the review activities. Students in the current study had to learn not only a single answer to a single question, but also a method to analyze a new situation and apply the previously learned principles to that situation to produce a new answer in the same form and general content as the answer they had practiced writing for earlier review activities. This new answer had to contain the same elements as the earlier answers while incorporating details appropriate to the novel situation. It is likely that mastering the ability to analyze a new situation and identify important components and then further to select and apply a previously learned procedure or principle to

generate a new answer required a higher level of intellectual engagement from students than simply memorizing answers to multiple-choice questions. The results of the current study suggest that students were able to generalize these higher-order skills, and further, that the students who most practiced these higher-order skills (as indicated by both completing 75% of the practice questions accurately and attending the mock exam) performed, on average, better on the actual exams than those students who practiced less.

The results of the current study suggest that both completing the practice questions accurately and attending the mock exams contributed to student success on real exams, but it is difficult to evaluate the separate contribution of each of the components. In order to better understand what components of the practice questions and mock exam were most helpful, it may be necessary to evaluate them separately.

For example, future research might explore the role each individual component of the mock exam session plays in helping students perform better on real exams. In particular, if there are easier or less time-consuming ways to conduct the session to avoid such a large time-commitment both from the students and the instructor it may be more readily adopted.

The results of the Simon (2005) and Semb (1963) studies, along with the exam performance of a small number of students during the fall semester of 2006 in the present course suggest that having the students actually complete answers to the practice exam under test-like conditions may not be necessary as long as the students

do complete the practice exam at some point before the answers are discussed. Due to scheduling conflicts with the mock exam sessions, several students during some sections of the Fall 2006 semester participated in an abbreviated version of the mock exam session. These students did not complete the practice exam in the presence of the TA, but instead brought the finished practice exam with them to a make-up session. The TA then inspected each practice exam to ensure the answers were complete and hand-written, and the sessions began with the distribution of the grading key, and proceeded in the same fashion as the second half of the normal mock exam sessions. The sessions lasted roughly ninety minutes, and at the end students could take their practice exams and grading keys home for study. The average exam score of the students participating in the abbreviated mock exam sessions was 23.86/30 (37 exams total). The average exam score of the students participating in the regular mock exam sessions was 23.54/30 (202 exams total). The exam scores of all of the students participating in both the abbreviated and regular mock exam sessions were included in the analysis for the effect of the mock exam presented in the results section of this paper.

These results, along with the results demonstrating improved performance with improved completion of practice questions, suggest that completing the practice exam may be an important variable in improved exam performance regardless of whether it is done under test-like conditions or under different conditions. If the mock exam session can be shortened by an hour and the effect on exam performance

remains the same, then that could be a practical advantage in the adoption of the method for use in other courses.

One limitation of this study was the lack of measurement of student attendance at lectures. It has been reported that attendance at lectures is correlated with improved exam performance (Borland & Howsen, 1998; Clump, Bauer, & Whiteleather, 2003; Gunn, 1993; Hancock, 1994; Shimoff & Catania, 2001). If those students who most often completed practice questions and attended mock exam sessions were also the students most often attending lectures, then the positive effects of both the mock exam and the practice questions could be confounded by lecture attendance. Additionally, it is possible that those students who did not attend lectures and attended the mock exams in an attempt to “catch up” or learn all of the missed material were responsible for the few failing scores among those students who attended mock exams. A more formal analysis is needed to determine the role of attendance on exam performance in this particular course.

An additional variable that may account for the results is the grade point average (GPA) of the students who did or did not take the mock exam or complete practice questions. It may be that the most successful students (as evidenced by higher overall GPA's prior to the beginning of the course) are the ones who are more likely to attend mock exam sessions and complete practice questions. If so, then perhaps the results simply indicate that students who do well in other courses also do well in the present course. An analysis of the GPAs of students in the course and

their completion of the practice questions and mock exams was not done. This analysis might be very beneficial in future research.

Nevertheless, some of the present results suggest that the positive effects of the mock exam may not simply be due to better students taking the mock exam. As shown in Figure 6, an increasing proportion of enrolled students participated in mock exams as each of the two semesters progressed. If the students with higher entering GPAs took the mock exam early in the semester and the students with lower entering GPAs then began taking the mock exams later in the semester, then we would expect the possible effects of the mock exam to diminish across the semester. This did not appear to be the case, although a much more detailed analysis needs to be done.

It may be the case that the higher achieving students are those most likely to attend mock exams and to complete practice questions. Nevertheless, they may not be the ones most helped by doing so. The larger effect of the mock exam for students who had earned fewer or no extra credit points suggests that it may be the lower achieving students for whom the mock exam is the most beneficial. This is also a replication of the results of Balch (1998), who found that a practice exam was most helpful for students who had been the lowest achieving in the course before that point. The difference in the effect size for low versus high-achieving students who completed practice questions is unknown. Future research may explore whether students with lower GPAs entering the course benefit more from the practice questions and mock exams than students with higher GPAs.

In summary, the current study adds to the literature by demonstrating that both practice questions and a mock exam may improve the performance of students on essay and short-answer based exams. Students who completed more practice questions accurately and who attended mock exams scored higher than students who did neither activity. By completing review activities that closely resembled the actual exam experience, it appeared that students were better able to correctly apply the principles and techniques presented in class to novel clinical situations described on section exams. This improved performance suggests that behavioral interventions to improve generalization can be successful at teaching even advanced academic skills involving higher-level competencies such as analysis and synthesis.

## REFERENCES

- Balch, W.-R. (1998). Practice versus review exams and final exam performance. *Teaching of Psychology, 25*, 181-185.
- Balch, W.-R. (2001). Study tips: How helpful do introductory psychology students find them? *Teaching of Psychology, 28*, 272-274.
- Bol, L., & Hacker, D.-J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *Journal of Experimental Education, 69*, 133-151.
- Borland, M.-V., & Howsen, R.-M. (1998). Effect of student attendance on performance: Comment on Lamdin. *Journal of Educational Research, 91*, 195-197.
- Boyce, T.-E., & Himeline, P.-N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *Behavior Analyst, 25*, 215-225.
- Clump, M.-A., Bauer, H., & Whiteleather, A. (2003). To attend or not to attend: Is that a good question? *Journal of Instructional Psychology, 30*, 220-224.
- Dickson, K. L., Devoley, M.-S., & Miller, M.-D. (2006). Effect of study guide exercises on multiple-choice exam performance in introductory psychology. *Teaching of Psychology, 33*, 40-42.
- Dickson, K. L., Miller, M.-D., & Devoley, M.-S. (2005). Effect of Textbook Study Guides on Student Performance in Introductory Psychology. *Teaching of Psychology, 32*, 34-39.



- Dufresne, R., Mestre, J., Hart, D.-M., & Rath, K.-A. (2002). The effect of web-based homework on test performance in large enrollment introductory physics courses. *Journal of Computers in Mathematics and Science Teaching, 21*, 229-251.
- Fleming, V.-M. (2002). Improving students' exam performance by introducing study strategies and goal setting. *Teaching of Psychology, 29*, 115-119.
- Flora, S.-R., & Logan, R.-E. (1996). Using computerized study guides to increase performance on general psychology examinations: An experimental analysis. *Psychological Reports, 79*, 235-241.
- Gunn, K.-P. (1993). A correlation between attendance and grades in a first-year psychology class. *Canadian Psychology, 34*, 201-202.
- Gurung, R.-A. R. (2005). How Do Students Really Study (and Does It Matter)? *Teaching of Psychology, 32*, 239-241.
- Hancock, T.-M. (1994). Effects of mandatory attendance on student performance. *College Student Journal, 28*, 326-329.
- Jenkins, J.-R., & Neisworth, J.-T. (1973). The facilitative influence of instructional objectives. *Journal of Educational Research Vol, 66*, 254-256.
- Keller, F.-S. (1968). "Good-bye, teacher . . ." *Journal of Applied Behavior Analysis, 1*, 79-89.
- Kouyoumdjian, H. (2004). Influence of Unannounced Quizzes and Cumulative Exam on Attendance and Study Behavior. *Teaching of Psychology, 31*, 110-111.

- Martin, G., & Pear, J. (2007). *Behavior modification: What it is and how to do it (8th ed.)*. Upper Saddle River: Pearson Prentice Hall.
- Mawhinney, V. T., Bostow, D. E., Laws, D. R., Blumenfeld, G. J., & Hopkins, B. L. (1971). A comparison of students studying: Behavior produced by daily, weekly, and three-week testing schedules. *Journal of Applied Behavior Analysis Vol, 4*, 257-264.
- Mayfield, K.-H., & Chase, P.-N. (2002). The effects of cumulative practice on mathematics problem solving. *Journal of Applied Behavior Analysis, 35*, 105-123.
- Miles, D.-T., Kibler, R.-J., & Pettigrew, L. E. (1967). The Effects of Study Questions on College Students' Test Performances. *Psychology in the Schools, 1*, 25-26.
- Miller, M.-L., & Malott, R.-W. (1997). The importance of overt responding in programmed instruction even with added incentives for learning. *Journal of Behavioral Education, 7*, 497-503.
- Oliver, R., & Williams, R.-L. (2005). Direct and Indirect Effects of Completion Versus Accuracy Contingencies on Practice-Exam and Actual-Exam Performance. *Journal of Behavioral Education, 14*, 141-152.
- Peat, M., & Franklin, S. (2002). Supporting student learning: The use of computer-based formative assessment modules. *British Journal of Educational Technology, 33*, 515-523.

- Ryan, C.-S., & Hemmes, N.-S. (2005). Effects of The Contingency For Homework Submission On Homework Submission And Quiz Performance In A College Course. *Journal of Applied Behavior Analysis, 38*, 79-88.
- Semb, G. (1974). The effects of mastery criteria and assignment length on college-student test performance. *Journal of Applied Behavior Analysis, 7*, 61-69.
- Semb, G., Hopkins, B. L., & Hursh, D.-E. (1973). The effects of study questions and grades on student test performance in a college course. *Journal of Applied Behavior Analysis, 6*, 631-642.
- Shimoff, E., & Catania, A. C. (2001). Effects of recording attendance on grades in introductory psychology. *Teaching of Psychology, 28*(3), 192-195.
- Simon, J. L. (2005). Effects if a Study Guide and Written Rehearsal on Undergraduate Test Performance. Unpublished doctoral dissertation, University of Kansas, Lawrence.
- Skinner, B. F. (1968). *The technology of teaching*. New York: John Wiley.