Does Embedding Test Instruction in Classroom Presentations Improve Test Performance?

Robin G. Yaure, Ph.D.
The Pennsylvania State University, Mont Alto

Jacqueline E. Schwab, Ph.D.
The Pennsylvania State University, Mont Alto

ABSTRACT. This study was conducted to determine whether embedding instruction about exam questions in class leads to higher exam performance. Students ($N = 64$) in an introductory level Human Development and Family Studies (HDFS) course were tested on questions discussed in class, questions that were in online materials, and new questions. Students' scores on the Study Behavior Inventory (SBI; Bliss & Mueller, 1986; 1993) were correlated with exam performance. On Exam 1, students performed better on the questions that had been presented in class than on the other two types of questions. Students' test scores also correlated significantly with their SBI results. On the second exam, students' performance on the questions presented in class and those available online differed significantly from the new questions, although exam performance was not significantly correlated with the SBI results. These results suggest that students benefit from embedded test instruction and alter their study habits following test performance feedback.

Introduction

The literature on student performance has three primary foci: student characteristics, test characteristics, and faculty characteristics. Student characteristics include motivation, learning style or orientation (Boyle, Duffy & Dunleavy, 2003; Busato, Prins, Elshout & Haramaker, 1998; Entwistle, Meyer & Tait, 1991; Tait & Entwistle, 1996; Vermunt, 1996), use of test-taking strategies (Dodeen, 2008), ability to self assess (Hacker, Bol, Horgan, & Rakow, 2000; Peverly, Brobst, Graham & Shaw, 2003; Rachal, Daigle & Rachal, 2007), text anxiety (Davis, DiStefano & Schutz, 2008; VanZile-Tamsen & Livingston,1999), motivation (Maurer, 2006; Sankaran, & Bui, 2001), and academic confidence (Bliss & Mueller, 1986; Hacker, et al., 2000).

Test characteristics include lecture- or text-based questions (Zamboanga, Padilla-Walker, Hardy, Thompson, & Wang, 2007), amount of practice provided (Maurer, 2006), time permitted for testing (Herman, 1997), and questions requiring primarily inference or memory (Johnson & Mayer, 2009; Peverly, et al., 2003).

Please direct correspondence to Dr. Robin G. Yaure at r2y@psu.edu.

Faculty characteristics include personality characteristics such as approachability and warmth (Alutu, 2006; Keeley & Shemberg, 1995), expertise (Alutu, 2006) and use of different techniques for providing information (McClanahan & McClanahan, 2002). The latter includes whether the faculty member uses primarily lecture or other techniques such as active and collaborative learning as well as how overtly test preparation is organized into the course.

Although a student's performance on an exam involves more than one of these components, only some of these can be controlled by the instructor. The purpose of the current study was to examine a method which could be controlled by the instructor: whether embedding direct instruction about understanding and preparing for test questions can assist students in their preparation for an exam.

The current study began in response to student comments about their performance on exams. The feedback on the test performance, routinely requested during each semester of the course, included an appraisal of the exam, a self-assessment of performance on the test, and an estimate of their study time and effort. On multiple occasions in past semesters, a number of students reported that the test questions were 'tricky.' Many had also reported that they had not studied a great deal for the exam. Thus, the current study was designed with two points in mind: first, to help the students understand how the questions were constructed and to provide suggestions on how to answer them correctly, and second, to embed test-taking instruction into the class itself. The latter situation would provide more opportunity for the students to study for the test.

Some research suggests that students who report that the questions on a test are 'tricky' may be more likely to have an external attribution style (Hacker, et al., 2000). In other words, they are more likely to blame outside factors for their failure rather than consider their own actions as the problem. Students who are more familiar with the course material are less likely to suggest that the questions are to blame or are confusing. Although attribution style and lack of preparation may be part of the reason for why students have found tests difficult in the course under investigation, it may also be that many of the students in the introductory class are new to college and have little or no experience in taking college-level multiple-choice tests. Thus, providing additional assistance to the students in understanding how questions are framed and examples of how analysis of the questions may be conducted may be helpful.

Davis et al. (2008) identified five different types of student patterns of appraising tests: tests are out of the control of the student, students are well-prepared-for challenges, students are hopeless, tests should be kept in perspective, and students require bracing for the worst. By providing direct support for students with different patterns of test appraisal, it is possible that some of the effects of test anxiety and a sense of lack of control may be reduced. Peverly et al. (2003) found that students who used more effort when processing course material achieved higher test performance than those who used less effort. Having direct instruction for test-taking strategies may also provide students with higher levels of effort and intention in processing course materials. Rachal et al. (2007) suggested that 'meta-curricular' assistance to students, which involves explicit instruction as an embedded part of a course, may be an effective way to enhance student learning. In their study examining the effects of self-reported learning problems of students, Rachal et al. found that first-year students reported fewer problems, which may have been more of a function of their lack of understanding about their learning problems rather than their actually having fewer problems. Volet, McGill, and Pears (1995) found that students whose instructors provided more explicit support of their learning performed better at problem solving, were more satisfied with their learning, and had greater levels of motivation in upper-division courses.

The current study examined whether providing direct instruction on how to understand and interpret exam questions may assist students in preparation for their exams. By embedding actual test questions into the class presentations and providing students with the opportunity to discuss the correct answers, ways that those answers may be obtained, and ways to identify incorrect answers, it was hypothesized that students would be able to perform better on subsequent exams.

Methods

Participants

The participants included students from two sections of an introductory Human Development and Family Studies (HDFS) class taught in Spring 2010 by the first author. After Institutional Review Board approval, students were informed of the purpose of the study and provided with a copy of the informed consent form for the study. Those who were under 18 were unable to participate in the study. A total of 64 students out of 80 enrolled in the two sections agreed to participate in the study. The students were exclusively first- and second-year students at a small branch campus of a large northeastern university. The course is generally taken by students majoring in Human Development and Family Studies and also by those who need a general education course in social and behavioral science. Seventy-three percent were female, 81% white, 19% African-American, and 73% were traditional-aged (age 23 or older) students. Each student was ascribed an identification number to provide confidential handling of his or her data.

Procedures

The students were presented with multiple-choice questions from previous exams, available in an online course management system (ANGEL), and online quizzes to be completed prior to the course material being covered in class. Students were able to access both the sample exams and online quizzes throughout the semester and were able to take them multiple times. Feedback was given to show which answers were correct or incorrect, but the letter of the correct answer on the multiple-choice question was not provided. A subset of these questions was embedded in a Power Point presentation for each chapter covered during the semester. Each question was shown directly after the corresponding material was presented. Students were shown the question and then given a short period to think about the answer. A student was called upon to provide the answer, and then discussion ensued regarding the correct answer and why the incorrect options were not correct.

After every two chapters, the students received a quiz with the questions from the presentations. The students were told the quiz grade would not be recorded since the purpose was to help them assess their understanding of the material, unlike a summative assignment. The corrected quiz was returned to the students prior to the in-class exam, or the answer key was provided in the case when not enough time was available to return the quizzes graded.

Twice during the semester, an in-class multiple-choice exam was given to the students that covered four chapters of the text. Each exam was made up of 50 questions. Eight questions were on the sample exam and had been reviewed in class, eight were on the sample exam but were not seen in class, eight were on the online quiz and seen in class, eight were on the online quiz and not seen in class, and 16 were new questions, based on course material, but not questions previously seen on a sample exam, online quiz, or in class.

Students were also asked to complete the Study Behavior Inventory (SBI), an instrument designed to measure three components of student study behaviors, including academic

confidence and short- and long-term study behaviors (Bliss & Mueller, 1986, 1993). Short-term behaviors include preparing for class on a daily basis, and long-term behaviors include preparing for exams and writing papers. One of the goals of having students complete this inventory was to help them become more aware of their study strategies and to gain confidence in their academic abilities. Forty-seven students completed the survey. They received nominal course credit for completing the survey. The survey was completed early in the semester, and then a workshop was presented in class by a learning specialist from the campus academic support center. The learning specialist and the instructor discussed how to interpret the results and how the students could prepare for class and exams. Periodically throughout the semester, the instructor referred back to the inventory, making suggestions of how the students may incorporate different suggested strategies in their own repertoires. The results for the inventory were recorded and used in later analyses.

There were two hypotheses for the study. First, it was expected that those students with higher SBI scores would perform better on the exams than those with lower SBI scores. Given that the SBI identifies students' study skills, it was expected that it would predict how well students would perform on the exam, regardless of how the instruction in the class influenced their performance. Second, it was expected that students would perform better on questions they had seen in class and had embedded into instruction than those they had only seen in online quizzes or sample exam.

Results

Hypothesis 1: To test the hypothesis that those students with higher SBI scores would perform better on the exams than those with lower SBI scores, a Pearson product moment correlation was computed with the SBI total percent score and the total number of correct answers on Exam 1. The results indicated a statistically significant relationship between SBI score and score on Exam 1, $r = 0.440$, $n = 46$, p < .001. Thus, students with higher SBI scores also tended to have higher scores on Exam 1. There are three factors measured by the SBI: academic confidence (Factor 1), short-term study habits (Factor 2), and long-term study habits (Factor 3). Further analysis examined the relationships between these three factors and performance on the exams. Pearson product moment correlations were computed, comparing the total correct for Exams 1 and 2 with each of the three SBI factors. Factor 1 (Academic Confidence) was significantly correlated with total correct on Exam 1, $r = 0.384$, $n = 46$, $p = 004$. Factor 2 (Short-term Study Habits) was significantly correlated with the total correct on Exam 1, $r = .290$, $n = 46$, $p = .025$. Factor 3 (Long-term Study Habits) was also significantly correlated with total correct on Exam 1 ($r = 0.370$, $n = 46$, $p = 006$). A significant relationship between the SBI and the total number of correct answers on Exam 2 was not found, $r = 0.075$, $n = 43$, $p = .316$. There were no significant correlations between any of the three factors and total correct on Exam 2. Thus, the SBI predicted performance on Exam 1 more accurately than for Exam 2.

Hypothesis 2: To test the hypothesis that being exposed to questions both in class and online would lead to better performance on the exams, a one-way repeated measures ANOVA was conducted with the within-subjects factor being type of exam question—question embedded in class (Seen), question on online quiz or exam (Not Seen), or a new question (New) as the grouping variable) and the dependent variable being the total correct answers. Given that there are unequal numbers of questions for the Seen, Not Seen and New questions, percentage scores were calculated for each type of question. An ANOVA-RM indicated a significant effect for question type, Wilks' $\lambda = .296$, $F(2,60) = 71.426$, $p < .000$ for Exam 1. Students performed significantly better on the questions that had been seen in class ($M = 83.59$, $SD = 16.946$)

compared with both those only seen in the online quizzes and sample exam ($M = 68.37$, $SD = 15.552$) and those that were newly constructed for the exam ($M = 61.65$, $SD = 14.465$). Post hoc analysis showed a significant difference between the questions that students had seen and those that were new, as well as those they had seen and those that were on the sample exam. Thus, the hypothesis that the embedded instruction would lead to improved performance on the exam was supported. Further analysis showed that students scored better on questions that had been on the sample exam and online quizzes compared with the new questions, Wilks' $\lambda=.426$, F (2, 60) = 40.365, $p = .000$. Post hoc comparisons showed that performance on all three types of questions differed significantly with the best performance on the questions from the sample exam ($M = 78.47$, SD = 16.759), next on the questions from the online quizzes ($M = 73.79$, $SD = 16.254$), and last on the new questions ($M = 61.65$, $SD = 14.465$) for Exam 1.

For Exam 2, Mauchley's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 6.362$, $p = .042$. Thus, degrees of freedom were corrected using the Huynh-Feldt correction ($\varepsilon=.903$). The results show that there was a significant difference between the different types of questions, F (1.862, 102.905) = 104.215, $p = .000$. Post hoc analyses demonstrated that performance on questions that were embedded in the class presentations ($M = 85.12$, $SD = 18.01$) did not differ significantly from those on the online quizzes or sample exam ($M = 84.18$, $SD = 84.18$), but that performance on both of those types of questions were significantly better than on the new questions ($M = 61.02$, $SD = 14.919$). An analysis also compared the performance on questions that had been on the online quizzes and sample exam as well as the new questions, as was done for the first exam. This time there were no significant differences between the scores on questions that had been in the online quizzes ($M = 85.80$, $SD = 15.695$) and sample exam ($M = 83.53$, $SD = 83.53$), but performance on both of these types of questions were significantly better than on the new questions ($M = 61.02$, $SD = 14.919$), Wilks' $\lambda=.206$, F (2, 56) = 107.834, $p = .000$.

Discussion

The purpose of this study was to determine whether embedding instruction of test-taking strategies and providing sample questions would help students perform better on their exams. This hypothesis was supported in the results from the first exam; students' performance was better on the questions they saw in class compared with those that were only on the online quizzes or sample exam and the questions that were new to them. The effect of embedding the instruction on specific questions may not have been present for the second exam since it is typical for students to focus more of their study time for the second exam on studying the online quiz and exam questions. They often realize that doing so is useful since some of the exam questions are available online. This finding is helpful for those designing ways to assist students in preparing for exams, although there is a limitation that this may not lead to transfer of knowledge in new situations, which is a typical pedagogical goal. Future studies might examine ways in which students are assisted in learning how to apply the concepts being assessed in one question to questions that are different from those they had seen previously.

The analyses regarding the Study Behavior Inventory, completed by the students prior to the first exam, partially supported the hypothesis that the SBI would predict the exam performance. This was true for the first exam but not for the second. This may be a result of student changes in their study habits and their confidence in their skills between the two exams. This is a likely situation for at least some students who gained a better understanding of how the exams were constructed and the effort they would need to perform well on the exams. It is not uncommon for students to report that they did not study much for the first exam, thinking that they knew the

material and were prepared adequately. After they get the first test's results, students often realize their efforts were not sufficient, and then they change their study behavior for the next exam. This outcome fits with the literature that indicates that direct instruction leads to increases in effortful processing by students (e.g., Rachal, et al., 2007). Effortful processing involves more direct focus by students on their learning. This notion is further supported by the analysis showing that performance on the embedded questions was not significantly better than those that were on the online quizzes or sample exam for the second exam. Students realized that using these resources was helpful for studying for the exam, so they used these more for the second exam. This was corroborated by the students' self reports that they gave after receiving their grade for the second exam.

While this study has limitations in demonstrating clearly how embedded instruction leads to improved exam performance, it has shown that students' exam performance improved over the semester. What may be important about embedding material into the course is not whether students learn about the specific questions for the test, but getting a sense that their own behavior may affect their performance on the exam. Thus, by emphasizing studying for the exams and concentrating on how to study, the embedded material may help the students realize they need to put more effort into their studying. By providing more opportunities for students to understand their abilities are not immutable and that they have control over their own performance, instructors can help students perform better in classes and beyond.

References

Alutu, A.N.G. (2006). The guidance role of the instructor in the teaching and learning process. *Journal of Instructional Psychology,* 33 (1), 44-49.

Bliss, L.B. & Mueller, R.J. (1986). An instrument for the assessment of study behaviors of college students. Paper presented at the Annual Meeting of the American Educational Research Association : San Francisco, CA, April 16-20, 1986. Retrieved from: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2f/30/aa.pdf

Bliss, L.B. & Mueller, R.J. (1993). An instrument for the assessment of study behaviors of college students. *Reading Research and Instruction*, 32 (4), 46-52.

Boyle, E.A., Duffy, T., Dunleavy, K. (2003). Learning styles and academic outcome: The validity and utility of Vermunt's Inventory of Learning Styles in a British higher education setting. *British Journal of Educational Psychology,* 73, 267-290.

Busato, V.V.,Prins, F.J., Elshout, J.J., Hamaker, C. (1998). Learning styles: A cross-sectional and longitudinal study in higher education *British Journal of Educational Psychology,* 68, 427-441.

Davis, H.A., DiStefano, C., & Schutz, P.A. (2008). Identifying patterns of appraising tests in first-year college students: Implications for anxiety and emotion regulation during test taking. *Journal of Educational Psychology,* 100 (4), 942-960.

Dodeen, H. (2008). Assessing test-taking strategies of university students: Developing a scale and estimating its psychometric indices. *Assessment & Evaluation in Higher Education,* 33 (4),409-419. DOI: 10.1080/02602930701562874

Entwistle, N.J., Meyer, J.H.F., & Tait, H. (1991). Student failure: Disintegrated patterns of study strategies and perceptions of the learning environment. *Higher Education,* 21 (2), 249-261.

Hacker, D.J., Bol, L, Horgan, D.D. & Rakow, E.A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology,* 92 (1), 160-170: DOI: 10.1037//0022-0663.92.1.160

Herman, W. (1997). The relationship between time to completion and achievement on multiple-choice exams. *Journal of Research and Development in Education*, 30 (2), 113-117.

Johnson, C. & Mayer, R. E. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology*, 101 (3), 621-629. Retrieved from: http://psycnet.apa.org/journals/edu/101/3/

Keeley, S.M. & Shemberg, K.M. (1995). Coping with student resistance to critical thinking. *College Teaching*, 43 (4), 140-146.

Maurer, T. W. (2006). Daily online extra credit quizzes and exam performance. *Journal of Teaching in Marriage and Family*, 6, 227-238. Retrieved from: http://familyscienceassociation.org/jtmf62006.php

McClanahan, E.B. & McClanahan, L.L. (2002). Active learning in a non-majors biology class: Lessons learned. *College Teaching, 50 (3),* 92-96.

Peverly, S.T., Brobst, K.E., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking. *Journal of Educational Psychology,* 95 (2), 335-346. DOI: 10.1037/0022-0663.95.2.335

Rachal, K.C., Daigle, S., & Rachal, W.W. (2007). Learning Problems Reported by College Students: Are They Using Learning Strategies? *Journal of Instructional Psychology,* 34 (4),191-199.

Sankaran, S. & Bui, T. (2001). Impact of learning strategies and motivation on performance: A study in web-based instruction. *Journal of Instructional Psychology,* 28 (3), 191-198.

Tait, H. & Entwistle, N. (1996). Identifying students at risk through ineffective study strategies. *Higher Education*, 31 (1), 97-116. Retrieved from: http://www.jstor.org/stable/3447710

VanZile-Tamsen, C. & Livingston, J.A. (1999). The differential impact of motivation on the self-regulated strategy use of high- and low-achieving college students. *Journal of College Student Development*, 40 (1), 54-60.

Vermunt, J.D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: a phenomenographic analysis. *Higher Education,* 31 (1), 25-50. Retrieved from: http://www.jstor.org/stable/3447707

Volet, S., McGill, T. & Pears, H. (1995). Implementing process-based instruction in regular university teaching: Conceptual, methodological and practical issues. *European Journal of Psychology of Education,* 10 (4), 385-400.

Zamboanga, B.L., Padilla-Walker, L.M., Hardy, S., Thompson, R.A., & Wang, S.C. (2007). Academic background and course involvement as predictors of exam performance. *Teaching of Psychology,* 34 (3), 158-162.