

FACILITATING ELECTRONIC EXAM PERFORMANCE BY PROVIDING A HARDCOPY DURING TESTING

Kathy L. Pettit-O'Malley, Department of Business, University of Idaho, P.O. Box 443178, Moscow, ID 83844, 208-885-6819, komalley@uidaho.edu

ABSTRACT

With increasing class sizes, many instructors are now using electronic exams. However, performance on electronic exams is often worse than on similar paper-and-pencil tests. The current study explores a means for achieving test performance comparable to that of hardcopy exams, while taking advantage of the superior grading efficiency of electronically-delivered exams.

INTRODUCTION

Universities facing declining revenues and increasing costs are turning to larger classes. With large classes, and increasing technological support options, instructors are using various electronic-assisted means to facilitate learning and testing. For example, use of “clicker” quizzes resulted in slightly improved subsequent exam scores [1]. Both immediacy of feedback and reduction of grading time have led to greater use of clickers. Many instructors have also tried electronic exams. But, in a study of electronic vs. identical paper-and-pencil tests, the former resulted in significantly lower scores [3].

There are two tasks which differ between electronic and traditional tests. First, a student has to *read* questions from either medium. A study comparing use of either a traditional or e-textbook showed no differences in overall learning [2]. But, testing produces greater anxiety than does reading a text. Any superiority in reading ease from paper could be magnified during a test, especially if the whole question could not be seen simultaneously on a smaller screen. The second task, *responding* to a questions, also differs between testing media.

We conducted a comparative study to determine if disadvantages of electronic exams could be mitigated, while keeping improved grading efficiency. Exams in a senior-level business (marketing) class were exclusively multiple choice and had three parts. For Part 1 about 100 testbank questions per chapter were uploaded to Bb Learn, which selected a unique, random subset of three questions per chapter per student. Part 2 was based on other readings and class discussions. Again, Bb Learn randomly selected ten questions per student from a pool of 35. Part 1 consisted mainly of definitional questions, whereas Part 2 questions involved applications and/or critical thinking.

Part 3 consisted of M-C question sets, each with five questions based on previously-distributed essay questions. The essays were complex, multi-part, and application-oriented, requiring critical thinking. Students could prepare essays together prior to the test. Each of six essays were converted into M-C questions sets. Partial credit was given for some responses. Three instructor-selected sets were available on the test, and a student could answer one, two, or three “essay” sets. With two or more sets, the best two were counted. If only one set was answered, Part 1 and 2 questions were 75% of the exam score, whereas with two or more sets, Part 1 & 2 were 50% of the total score. On average, students answered 2.61 essay sets, as they knew the substance of the questions during test preparation.

The same instructor taught both sections, with no changes in text, lecture material, or provided essays. For one section all parts of the exam were given/taken through Bb Learn (i.e., electronically). *The alternative test was identical to the previous one, except that students were provided with a hardcopy of Part 3 questions, to use during the test.* All questions were answered through Bb Learn. The same three essay-related question sets were included for both sections. Independent samples t-tests were used to test performance differences. Parts 1 and 2, combined, did not differ between sections ($t_{65.211}$, not significant). However, those with the hardcopy of Part 3 performed significantly better on the test than did those reading solely from a computer screen ($t_{67.871} = -2.742$, $p < .01$). The exam mean (raw score) for the totally electronic exam was 58.0, whereas the version accompanied by a hardcopy of Part 3 had a mean of 66.1. Both the instructor and students were less frustrated in the fall semester, when students performed better. Limitations, conclusions, and implications will be discussed.

REFERENCES

- [1] Morling, B., M. McAuliffe, L. Cohen, and T. M. DeLorenzo. Efficacy of Personal Response Systems ("Clickers") in Large, Introductory Psychology Classes. *Teaching of Psychology*, 2008, 35(1), 45-50.
- [2] Rockinson-Szapkiw, A. J., J. Courduff, K. Carter, and D. Bennett. Electronic versus Traditional Print Textbooks: A Comparison Study on the Influence of University Students' Learning. *Computers and Education*, 2013, 63 (April), 259-266.
- [3] Wigenbach, G. J. Agricultural Students' Computer Skills and Electronic Exams. *Journal of Agricultural Education*, 2000, 41(1), 69-78.