ON EVALUATING ELECTRONIC TESTING ENVIRONMENT

Hassan Younies
College of Business and Economics
United Arab Emirate University
Al-Ain, U.A.E.
P.O. Box 17555
Fax: 071 3 7624384

Belal Barhem
College of Business
Abu Dhabi University
Al-Ain, U.A.E.

ABSTRACT

In this article we study the security conditions of exams conducted in online environment. A statistical package is used to detect the security of online exams and the level of students cheating compared to the classical hardcopy type of exams. Two female sections of a management core course were employed in this study. Six different tests were given to the students. Section enrollment ranged between 28 and 35 students. The electronic testing deployment system under study is the one used at United Arab Emirate University (UAEU), namely the Blackboard learning system (Release 6.0). As far as cheating level and security issues, the results indicate very little difference between the traditional type of testing and the electronic online type.

1- Introduction:

Student assessment, or evaluation, is one of the most important aspects of the educational process. It involves many issues such as: 1) assessments’ diversification, 2) assessments’ tools, 3) matching learning outcomes with assessment tools and 4) plagiarism and cheating. The new electronic era, being twofold, has created new tools for students and instructors in data availability, distance learning and assessments’ deployment. It also created new detriments to the instructors and to the evaluation process on the forms of plagiarism and cheating. In particular, instructors should be vigilant and pay attention for the prevention and detection of such issues as plagiarism on research papers and projects. Different tools are in place to handle plagiarism, such as online tools that can help in detecting plagiarism. Most of the research regarding cheating in online exams focused on the issues of identity checking of test takers and copying from material in non-proctored online exams. For instance the paper by McLaffert and Foust (2004) handled the online plagiarism issues in detail.

To the best knowledge of the authors the issues of the effect of online environment on students cheating behavior in a proctored test setting have not been addressed in the literature.

With the advent and accessibility of computers and networking technology, electronic deployment of exams is getting more popular than the traditional exams. In addition, the readily
available electronic versions of test banks, from textbook publishers, ready to use in popular e-deployment software spur the use of online exams. These test banks are already compatible with the most commonly used electronic examination deployment software such as Blackboard (Bb) and WebCT.

One of the most used assessment technique is the multiple choice and true/false type of exam. In this paper we investigate one issue related to electronic deployment of assessments, mainly student cheating in objective tests i.e., multiple choice and true false questions. Cheating detection in multiple choice exams has been under thorough investigation by many researchers in different disciplines, such as medicine (Ercole et al. 2002), Business (Thompson 1994) and Engineering (Kvam 1996). One way to detect cheating on objective tests is by using statistical tools. Section 2 will briefly discuss the basis of these statistical methods. This research study will use a statistical tool developed by Wesolowsky (2000), which is designed to detect student cheating on exams.

Electronic deployment can take place on both non-proctored and proctored environment. The non-proctored environment is the case where students attempt to answer online questions at their own location (home, office etc.). This kind of exam deployment has its own security issues such as: 1) plagiarism, 2) difficulties to identify the identity of individual/s taking the exam and, 3) make use of textbooks or other reading material not allowed by the course instructor, Rovai (2000).

The latter form of electronic deployment of exams is the proctored test-taking environment. Proctored electronically deployed exams have been customary for many years, using customized deployment software, in exams and certifications such as (Test of English as a foreign language (TOEFL), Graduate Record Examination (GRE), Microsoft certificates exams, etc. Security issues in proctored exams are different from the non-proctored environment. The issues of test taker identity and usage of external material can be resolved easily. However, based on the test type, other issues such as cheating from other students or from hidden material is always a concern. Different measures and the proctor vigilance can reduce the latter type of cheating method, i.e., the hidden material cheating type.

Collusion between students by looking at someone-else’s exam paper is hard to control unless excessive spacing between students and different versions of exams are in place. Both student spacing and deployment of different versions of exams - electronically- are not an easy issue under online environment, we will discuss these issues in Section 3. United Arab Emirate University (UAEU), adopted Bb as its electronic learning environment for student resources and exam deployments. The university has a policy of active learning and laptop computerized environment, with instructors and students being equipped with wireless laptops. Classrooms are provided with the necessary technology for this wireless environment. In addition, the UAEU has two separate campuses, one for male and the other for female students. This study was conducted only on the female campus. In Section 4, we will point out the major options available for instructors in deploying their exams under Bb.

The study setting and findings, conclusions and recommendations will be discussed in Section 5.
2- Statistical methods for detection of cheating in objective tests:

Academic dishonesty is a reality in teaching which jeopardize the teaching outcomes and integrity. Therefore, detection of student cheating and security of exams setting is of a great interest for academicians in all disciplines. In fact, some studies show that more than 20% of students are engaged in different forms of cheating, threatening examination quality and educational outcomes, Thompson (1994). To avert cheating in objective tests educators use a variety of exam arrangements such as: seating plans, more spacing between students, randomizing questions and more vigilant inspection and monitoring. These costly and time-consuming arrangements, while necessary, can help only in reduction of cheating from adjacent students but would not guarantee eliminating all forms of cheating.

Therefore there is a need to examine the security of the different exam arrangements and to support any accusations of cheating. Among the soundest tools for detecting possible academic dishonesty is the statistical detection of cheating on objective exams, i.e. multiple choice and true-false type of questions. Statistical detection of cheating on objective tests is, mostly, based on the idea of comparing the number of matching answers between exams. Probabilities of wrong and correct answers can be estimated from the overall class answers. For example Thompson (1994) used a spreadsheet program to compare the results of all pairs of adjacently seated students. McManus et al. (2005) used "regression of similarity correct answers in all possible pairs of candidates or the overall proportion of correct answers". The program by Harpp and Hogan (1993), flags students with "number of errors in common to number of exact errors in common exceeds a certain cutoff value". Identified students are checked with the seating plan to check if they were seated adjacent to each other. These tools were tested on actual exams and classes and proved to be workable methods of detection.

Wesolowsky (2000), proposed a new model based on an algorithm that examines all possible pairings of students and flag student pair with similarity above threshold level. The algorithm is suitable to estimate the level of copying in the class and it can be used for true-false and multiple choice questions. The program is also suitable for testing the security environment of administered exams.

3- Exams issues on proctored online environment:

Cheating on objective exams is common and hard to detect. Prevention techniques vary from various test versions to prearranged seating and more spacing between students. However, some of these measures may “consume time and resources and may contribute to an atmosphere of mistrust leading to unnecessary anxiety among students” Thompson (1994). In online tests, there are other issues specific to this media of examination. First, instructors are limited to the testing software flexibility and available testing options. Second certain types of exam problems are not possible to be deployed, unless combined with other tools, Naslund (2005), and they may not be convenient for students due to various editing issues. An example is formulae writing problems, where editing would be time consuming, with a grading bias toward computer proficiency and not course comprehension.

Equipped classrooms and ample spacing between students are a main issue in online environment. Computers input devices and power supplies require more spacing than normal paper exams. Lack of large classrooms, equipped to handle online exams, would force instructors to forgo spacing between students thus making it easier for students to cheat.
The most common way of cheating in objective exams, glancing into adjacent exam paper, is different in online exams. In laptop environment collusion with other students can be made easier through technological means. If collusion exists, students can make use of technology to their advantage such as adjusting font size on their web browsers, screen brightness and color in order to make collusion possible. Finally, many universities, including UAEU, are adapting online environment where students own their laptops, with little or no control or jurisdiction from the school and the instructors on students’ laptops.

4- Description of the Electronic deployment tool used for this study:

Blackboard (Bb) is turnkey Electronic Learning Environment software. It comes with many built in features, enhancing the teaching and learning experience for instructors and students. One main advantage of Bb is the ability to deploy tests easily and promptly. In this section we point out the main testing features available in Bb that are applicable to objective tests, i.e. multiple choice and true/false type questions.

The Bb testing features can be divided into two main categories: a) test preparing and setting feature and b) test deployment features.

4.a Test preparing and setting in Blackboard (Bb V6.0)

Test preparing features are the options available in preparing questions and setting up exams. Users have the options to write their own questions or to choose from a test bank. If a test bank is used, users can select specific questions or select a random block of questions from the test bank. From the authors’ experience, choosing a random block of questions is not the best way to go. First, same topics may be repeated in two different questions in the same exam. This would require instructors to make sure that questions on similar topics are removed from test banks ahead of selecting questions from the test bank. Second, random blocks would not ensure equity in test level among students. A study by Taub and Bell (1975) shows that students with randomly arranged questions would score lower than students with questions appearing in order of textbook or lecture material. This was supported by Carlson and Ostrosky (1992). Therefore, in this study we did not use random blocks in our exams, instead we selected questions from test banks with different deployment settings, including randomizing questions. Figure 1 shows the test setting features available in Bb (see appendix I). It must be noted that, deployment of different versions of exams on Bb is possible in newer versions of Bb. Students are divided into groups and different versions are made available to different groups. In addition to that, new Bb version has features allowing for randomizing answers for each question. Nevertheless, this feature is not available in some test banks. At the time of this study, these features were not available in the Bb version used at UAEU. These extra features would enhance testing security and reduce any possible collusion. It goes without saying that testing exam security without these features would give an indication about the worst case scenario available on e-deployment method and about exam security in general. We believe that these new features do help to strengthen online testing security conditions.

4.b- Test deployment features available in Blackboard

Test deployment features are the options available in Bb for deploying questions. These features include the following:

1- Randomizing the order the questions appear for students with the following options:
   a- Show all questions at once on students’ browser.
b- Show one question at a time with the backtracking capabilities (ability to revisit previous question).

c- Show one question at a time without backtracking capabilities.

2- Show questions without randomization with the same sub-features as in one above.

These test deployment features can be summarized as shown in Table (1). Instructors may use any one of the six different options available. Figure (2) in appendices gives a screen snapshot of these features.

In Table 1, one question appears on the student browser at a time if instructor selects any of the options A, B, C and D. Students are allowed to get back to the question and change the answers in options A and D. While options B and D won’t allow students to backtrack to previously viewed questions. In addition, questions appearance on students’ browsers can be randomized or kept in order selected by the instructor. In the last two options (E and F) questions appear all at once on the students’ browsers, questions can be randomized as in option E, or it can be displayed in the same order of the test setting.

Generally speaking, having exams where students have no backtracking is not popular with students and instructors, because students would like to review their answers and. On the other hand, it is expected that exams with randomized question displayed on students screens and without backtracking to have the strongest security setting. An example of this feature is option B in Table 1, where questions are randomized and students can see, and must, answer one question at a time without backtracking.

<table>
<thead>
<tr>
<th>Option reference</th>
<th>Test</th>
<th>Randomize Questions</th>
<th>Prohibit Backtracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>One at a Time</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>B</td>
<td>One at a Time</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>C</td>
<td>One at a Time</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>D</td>
<td>One at a Time</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>E</td>
<td>All at Once</td>
<td>Y</td>
<td>NA</td>
</tr>
<tr>
<td>F</td>
<td>All at Once</td>
<td>N</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 1: A Summary of Test Deployment Features Available in Bb.

5- Study Results, Conclusions, and Recommendations

The study used statistical cheating detection software which was developed by Dr. Wesolowsky. Tests were designed according to testing options of Table 1. During all exams, standard proctoring and student monitoring procedures were enforced thus assuring proctoring vigilance during the exams. In this study, students’ names and numbers and seating positions were recorded in each exam. Students were not aware of this study.

The course used for this study was a business core course, Fundamentals of Management, which was offered every semester and in multiple sections. The textbook test bank was used in all the experiments. During the study we conducted several exams, with each deployment setting
available on Bb, for at least two female sections. Sections enrollment during the study varied between 28-33 students. Exams were downloaded from Bb and entered into proper format suitable for the statistical detection software. The software results for various exam options are shown in Table 2.

The study shows that a single suspected collusion was registered for option F for students sitting in adjacent seats. The algorithm did not detect any other collusion for other testing options.

Testing environment security would still be an important issue where institutions and instructors should deal with. Using online testing software for objective testing is not less secure than traditional testing methods. Our study shows that the amount of possible collusion in online environment is within the 5% reported in traditional testing environment. It must be mentioned, however, that there may be other factors that could influence the results such as culture and gender of test takers.

<table>
<thead>
<tr>
<th>Options</th>
<th>Total Number of Students</th>
<th>No. of Pairs of Suspected Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>63</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Number of Cheating Detected for Different Bb Deployment Features.

With the new features of Bb deployment systems we can expect enhancements of test security conditions. Nevertheless, technology is twofold with pros and cons. Institutions and instructors should be aware of the threats of electronic cheating methods. The authors believe that, using a wireless connection and having full jurisdiction of the computers where tests are conducted would eliminate or at least minimize cheating on electronic/online testing environment. The authors recommend the following: a) to enhance security of online testing, instructors must be vigilant and must use the latest electronic techniques in test design and deployment, b) the instructor computer competency should be in the highest level in order to increase his ability to detect the latest software and hardware technology, and c) the testing software must have restrictive access to exams by students and no allowance for copying or transmitting exams and answers to test questions to unauthorized parties. This to be established shortly before, during and at least two days after the exam was administered.

Acknowledgement:
The authors are grateful to Dr. Wesolowsky for providing a copy of his program and for the anonymous referees for their valuable comments.
APPENDIX

Fig 1.a
Add the question text, then set the point value and the number of answers. Add the answer choices, indicate the correct answer, and provide feedback for user responses.

**Multiple Choice Question**

**Question Text**  
Test Settings

**Point Value**  
1

**Answers**  
Select the number of answer choices, fill in the fields with possible answers, and select the correct answer.

**Number of Answers**  
4

**Fig 1.b**

Figure 1: Test setting options available on Bb
Figure 2.a
Reference:


Additional References Available from the Authors.