

The Detection and Minimization of Cheating During Concurrent Online Assessments Using Statistical Methods

John H. Mott
Purdue University

ABSTRACT

As distance learning programs in aviation continue to evolve, a natural step in the progression is the fully online assessment of distance students. In some cases, those assessments must occur in synchronous delivery systems, where distance students view lectures and/or take online examinations concurrently. Among the questions that arise in the mind of the aviation educator in such cases are those of how to determine the magnitude of academic dishonesty occurring during concurrent assessments and how to minimize the opportunities of students being assessed to participate in such dishonesty. The author has developed a statistical algorithm to allow the use of common software to detect instances where students pass answers between themselves through various means in order to improve their assessment scores. A study conducted over a period of two semesters during the 2009 – 2010 academic year in two of the author's distance learning classes indicates the validity of both the algorithm and the methods used to minimize the occurrences of dishonesty.

INTRODUCTION

Many collegiate aviation programs have in recent years added distance education delivery methods; at some such institutions, entire degrees may be earned by distance learners. Distance learning course management systems (CMS) such as Blackboard / WebCT typically include an assessment component that allows the assessment of students entirely online. Benefits of online assessment include increased grading accuracy, reduced grading time, and provision of students with immediate feedback (Eplion & Keefe, 2005). As Eplion and Keefe (2005) suggest, these benefits are also present if online assessment is used with traditional delivery methods, with the additional benefit that such assessment can make available additional class time for other learning activities.

An issue that arises with regard to the use of distance learning CMS is that of the minimization of cheating by students during the process of online assessment. Dick et al. (2003) suggest that managing the issue of cheating consists of three steps: preempting it, detecting it, and responding to it. The research described herein was targeted primarily at the first two of these steps; namely, detection and minimization.

There is no question that cheating on online assessments should be a concern for distance educators. Rogers (2006) conducted a survey of distance learning faculty who regularly use a CMS at a regional university in the southeastern United States. The survey indicated that 33 of 54 respondents also use the assessment feature of the CMS, and that 17 (or 52%) of these had experienced or suspected some form of academic dishonesty during an online exam. A comprehensive literature review by Dick et al. (2003) indicates that additional studies conducted on the frequency of cheating over a period of 37 years from 1964 to 2001 have determined a range of students found to be cheating from as low as 40 percent to as high as 96 percent. Those studies were not specifically focused on distance education, however. Mason and Waight (1998) suggest that online course offerings are typically taken by the “best” students, or those having the highest grade point averages, and that this causes the frequency of cheating on online assessments to be less than that found with traditional assessments, and, indeed, as Crown and Spiller (1998) confirm, a negative correlation exists between grade point average and cheating frequency. Other

characteristics correlated with cheating include age, extracurricular involvement, and major field of study (Burlak, Hernández, Ochoa, & Muñoz, 2006). Rowe (2004), however, suggests that as distance learning becomes more popular (or perhaps a requirement in some programs), distance educators will see an increasingly broader spectrum of students. It is clear, therefore, that preparation for dealing with the academic dishonesty issue needs to occur sooner, rather than later.

With advances in computer and communications technology, there are many innovative methods that may be employed by students who wish to cheat on online assessments. Olt (2002) and Rowe (2004) describe the use of forbidden sources of information during online testing. Such sources during unproctored assessments might range from textbooks and class notes to proxy test takers, who may be easily utilized if some form of identity verification is not employed. It is clear that an instructor is unable to control a student's use of unauthorized sources of information during an assessment where a proctor is not present. Some authors (Abbott, Siskovic, Nogues, & Williams, 2000; Dirks, 1998; Kaczmarczyk, 2001; Kim, Smith, & Maeng, 2008) suggest the use of proctored testing at a predetermined location, but this approach introduces difficulties to the assessment process. The first of these is the problem of testing a large class of distance students in a facility that may not be able to accommodate the entire class (Pain & Le Heron, 2003), and the second is requiring students to travel what may be a substantial distance to reach the testing facility (Cooper, 2000).

Two other means of cheating on distance assessments, and two that are perhaps especially applicable to assessments that are delivered to students concurrently, is that of the passing of e-mail, text messages, or cellphone calls between students taking the exam (Olt, 2002), and the use of spyware placed on other students' machines to allow the cheater to see those students' answers. The research presented in this paper is particularly relevant to combating these two modes of academic dishonesty.

There are a number of authors who suggest best practices for reducing the frequency of cheating on online assessments. Dick et al. (2003) warns educators to recognize that cheating is probably occurring in their classes. Several studies (McMurtry, 2001; Olt, 2002; Rowe, 2004) propose that cheating be clearly defined by the instructor and that an academic dishonesty policy be provided to and discussed with students (and perhaps returned to the instructor with the students' signatures). Dirks (1998), however, found that only 15% of academic syllabi collected during a study actually contained such policies.

According to Olt (2002), distance education assessments should be "oriented toward higher ordered thinking skills, requiring application, evaluation and synthesis rather than mere factual recall." Meyen, Aust, Bui, & Isaacson (2002) suggest the use of collaborative projects, case studies, discussion boards, and peer evaluation in the online environment to promote such skills. However, a study conducted by Kim, Smith, & Maeng (2008) of three different distance education programs at a large university in the Midwest found that a majority of online assessments continue to rely on the same types of assessments found in traditional course delivery methods. The Rogers (2006) survey indicated that 91% of the respondents using the WebCT assessment feature in their distance courses employed a multiple-choice format. This type of question is typically easiest to set up in the CMS and easiest to grade automatically.

Best practices that are specifically applicable to online assessments conducted using a CMS and consisting of multiple-choice questions include randomizing questions selected from a larger pool, limiting the time allowed to take the assessment, and disallowing submission of answers after the time limit has expired (Illinois Online Network, N.D.; California State University, Sacramento, N.D.). Eplion and Keefe (2005) suggest that, in cases where answer patterns appear similar between exams, IP addresses be compared to determine the physical proximity of the machines from which the assessments in question were submitted. While address proximity may be an indicator that unauthorized collaboration has occurred, lack of address proximity does not prove that it has not occurred through electronic means.

Given the fact that we as distance educators do not always follow recommended practices and continue to use the multiple-choice format, it is important that we possess the tools for detecting academic dishonesty on online assessments and the knowledge of the parameters that may be varied within the typical CMS to effectively minimize such dishonesty.

DEVELOPMENT

When comparing two sets of answers to multiple-choice questions in an online assessment in order to determine whether unauthorized collaboration between the students occurred, it is essential to consider the underlying statistics. According to Rowe (2004), because of the fact that little of the test-taker may be observed online, “statistical methods are often the only hope to detect problems. To detect cheating, therefore, we shall consider the commonality of incorrect question responses between the datasets. Rowe (2004) states that the distribution of the number of answers in common between noncheating pairs of students should usually be close to Gaussian, but this observation is incorrect for several reasons; first, because the Gaussian or normal distribution is a continuous probability distribution and the case under consideration is discrete; second, because the Gaussian distribution is real-valued, whereas the number of responses in common between two assessments must clearly belong to the set of nonnegative integers; and third, because a Gaussian distribution implies a large population, which, in most cases, will not be available. The statistic in which we are interested is that of the identical *incorrect* responses between student pairs. The distribution of this random variable should, in fact, possess a Poisson probability mass function,

$$\Pr(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad (1)$$

where k is a non-negative integer, in cases where no academic dishonesty has occurred.

A type of discrete probability distribution known as the negative binomial distribution occurs often in nature and is used by meteorologists to model contagious events, or processes where each event with a favorable result increases the probability of succeeding favorable events. Natural events such as tornado outbreaks are accurately modeled with this sort of distribution. The probability mass function of the negative binomial distribution has the form

$$\Pr(X = k) = \binom{k+r-1}{r-1} (1-p)^r p^k, \quad (2)$$

where r and p are real-valued and k is a non-negative integer.

Because of the fact that we will allow the parameter r to be real-valued, this is more properly known as a Pólya distribution. The rationale for using a Pólya distribution to represent the number of pairwise incorrect identical responses between dishonest students concurrently taking an online assessment is twofold. First, it makes some sense, in that it is easy to see that cheating on an assessment can be considered contagious according to the definition above. Second, it is relatively straightforward to show by using the parameterization

$$p = \frac{\lambda}{r + \lambda}, \quad (3)$$

that the Pólya probability mass function converges to the Poisson probability mass function as the parameter r approaches infinity; hence, an increase in r implies a decrease in cheating on the assessment.

The process, then, of detection of cheating on concurrent online assessments becomes one of calculating pairwise vectors of identical, incorrect answers between all students taking an exam concurrently, ascertaining the number of identical, incorrect responses per pair, and then fitting the aggregate data to a Pólya distribution. A test for goodness of fit will then allow the researcher to determine how close the actual distribution comes to the ideal case, and this closeness presents a picture of the validity of the model. The degree of actual cheating for each assessment will logically follow the skewness, or third central moment, of the distribution. This is easy to see since skewness represents the “tilt” or tail-heaviness of the distribution, and it is straightforward to understand that Pólya distributions that are positively tail-heavy, or that exhibit positive skewness, imply that a handful or more students have a great many incorrect answers in common between them.

The Blackboard CMS has the ability to export results files in Microsoft Excel format for all students participating in an assessment. Assuming that multiple-choice questions are used, the worksheet can then be translated easily into a numerical format suitable for import to a matrix analysis package such as MATLAB, used for the research described herein.

The MATLAB algorithm for calculating the skewness of the observed distributions first allows the input of an $n \times q$ matrix of student answers, where n is the number of submitted assessments and q is the number of multiple-choice questions. From this, a three-dimensional matrix is computed where the third element corresponds to the answer selection. For most multiple-choice responses it is simplest just to allow the position of a “one” to designate the alphabetic response selected by the student, and to fill the remainder of the dimension with “zeroes.” Similarly, a $1 \times q$ matrix of correct answers is entered, and this is converted to a three-dimensional matrix to represent the proper alphabetic response. An “incorrect answer matrix” can then be constructed from the two three-dimensional matrices described above. This resultant is also three-dimensional. The final product, \mathbf{X} , a vector of $\binom{n}{2}$ elements, is calculated by making pairwise comparisons of the columns of the incorrect answer matrix. One may then calculate requisite statistics from \mathbf{X} , including those of the most interest to us, the mean and skewness.

Once statistics for the observed distribution have been calculated, it remains to fit the data to a Polya distribution and use a chi-square test to determine goodness of fit. The value of the parameter chi-square provides a good indication of how well the data fits the ideal Pólya distribution; the better the fit, the more cheating has occurred. The actual determination of the parameters p and r is somewhat difficult, in that it requires the solution of a logarithmic equation, but generally an iterative method can be used, in which an initial value of one parameter is provided, the value of the other is calculated, and these starting values used in a nonlinear programming exercise that, one hopes, converges to the proper solution. In practice, convergence was obtained in all cases in which the Pólya distribution was appropriate; those cases in which the iteration did not converge were, in fact, found to exhibit little cheating and could be fitted to Poisson distributions with ease.

After the distribution has been fitted, it is straightforward to calculate the ideal skewness of the model and compare this with the skewness of the data. The difference in these parameters, along with chi-square, allows the educator to quantify the degree of cheating on the online assessment in question.

RESEARCH METHODOLOGY

Purpose of the Study

The purpose of this study was to determine from a relatively small set of testing parameters common to typical distance learning course management systems those parameters producing the more significant impact relative to the minimization of cheating on concurrently administered online assessments. The study was conducted in two aviation courses taught at the Indianapolis Campus of Purdue University over a two-semester period during the 2009-2010 academic year. Both of the courses used a hybrid delivery method; that is, traditional class lectures integrated with online homework and assessments. It was originally planned to extend the study to a third, purely-asynchronous online course, but the enrollment figures in that course were too low ($N = 5$) to allow any sort of meaningful analysis.

Study Population

Fifty-three unique students who enrolled in the two aviation technology courses over the two-semester period were targeted as the sample population for this study. Forty-nine of those students were enrolled in only one of the courses; four students were enrolled in both.

Methodology

Online assessments administered to all members of the population consisted of two section examinations and one final examination, comprising a total of six exams for the two courses. All of the assessments were of the multiple-choice type, with section exams consisting of 50 questions and final exams consisting of 100 questions. While it was possible to select the questions randomly from a pool, the pool sizes were no larger than the number of questions on the assessments.

Three parameters were varied among the assessments in order to investigate the effects of the changes on the degree of cheating detected. The three parameters, all easily controllable in Blackboard / WebCT, were

1. Time allowed for submission,
2. Order in which the questions were presented, and
3. Whether students were allowed to revisit questions once they had been answered.

Other parameters available for variation in the CMS include the choice of when to present assessment results to students (immediately, or after all assessments have been completed), whether to disallow answer submission after time for the assessment has expired, and the point at which the assessment is available for starting by the student. These additional parameters were all fixed for this study; assessment results were released to the student immediately after completion, and answer submission was disallowed after the submission time had expired. The assessments were made available only during a one-hour window; students were required to begin the assessment at some point during the specified hour. The availability window does not affect the time available for completion of the assessment. Assessments were not accessible at any point outside of the available window.

RESULTS OF THE STUDY

The first section assessment was taken by 34 students ($n = 34$) and consisted of 50 multiple-choice questions ($q = 50$). From combinatorial analysis, we have that the total number of unique pairwise comparisons to be made is $\binom{n}{2}$, or 561. Figure 1 gives the calculated frequency distribution.

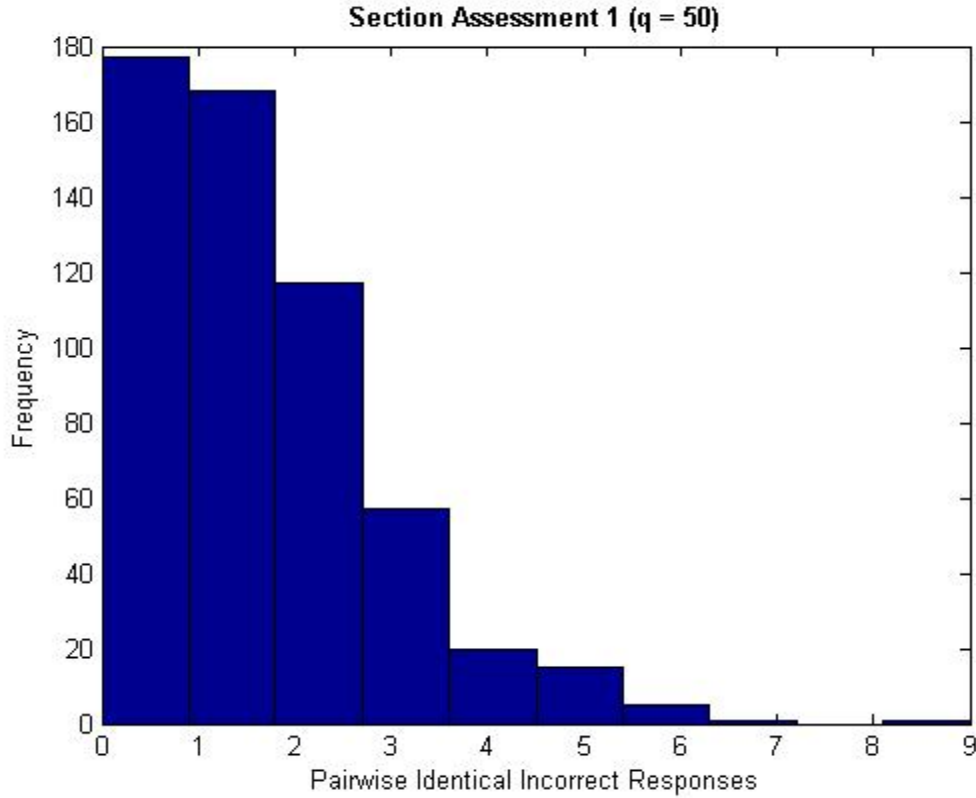


Figure 1. Pairwise identical incorrect responses, Section Assessment 1.

Study parameter values for Section Assessment 1 were an allowed test time, t , of 1.5 hours, a non-random question order (meaning that all students received the same questions in the same order), and allowed revisitation.

The calculated mean, μ , is 1.3797. The skewness, γ , calculated from the dataset is 1.3109, while the skewness in the ideal case with no cheating, that is, the skewness of the Poisson distribution of mean μ , is $\frac{1}{\sqrt{\mu}}$, or 0.8535. The disparity between the two skewness values ($\Delta = 0.4574$) suggests that some degree of cheating occurred, and this is substantiated by the chi-square calculation, which yields a value of $\chi^2 = 5.02$. The 9 degree-of-freedom values, or confidence limits, corresponding to the 97.5% confidence interval are 2.7 and 19.0, and χ^2 falls well within the range, so this demonstrates that the Pólya distribution is a good fit for the data. The values of p and r for the model, calculated using the nonlinear programming method, are $p = 0.7061$ and $r = 3.3155$.

The second section assessment was also taken by 34 students ($n = 34$) and consisted of 50 multiple-choice questions ($q = 50$). The total number of unique pairwise comparisons is 561. The corresponding distribution is shown in Figure 2.

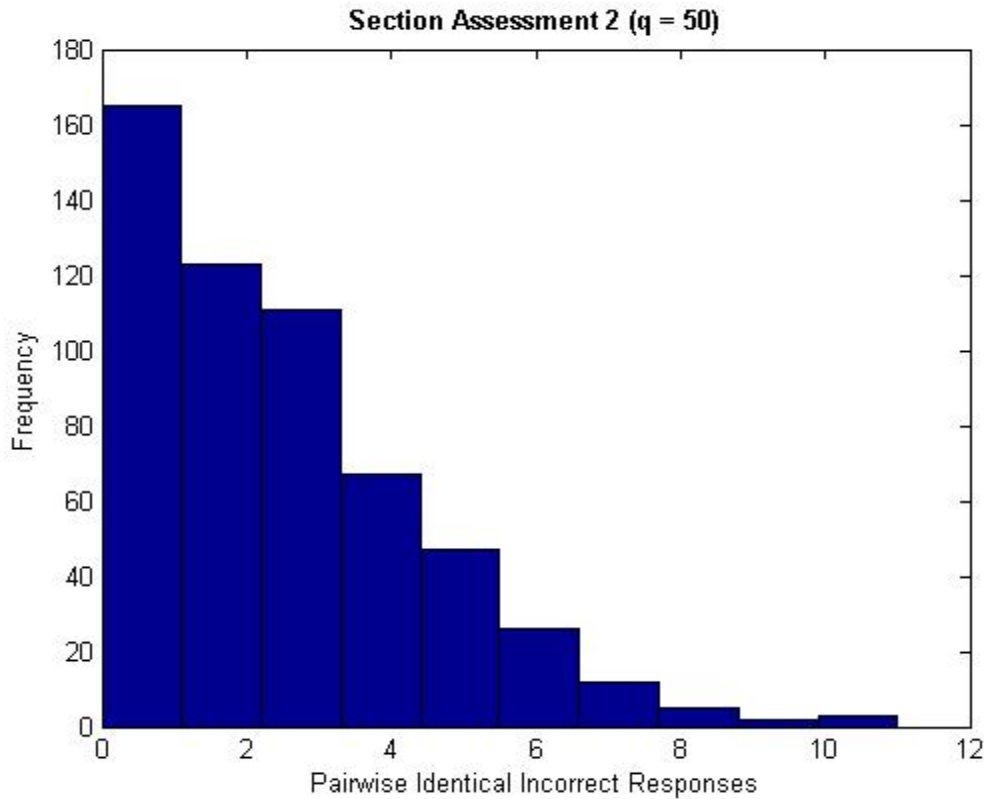


Figure 2. Pairwise identical incorrect responses, Section Assessment 2.

Study parameter values for Section Assessment 2 were $t = 1.5$ hours, a random order of questions presented on the assessment, and allowed revisitation.

The calculated mean, μ , is 2.7291. The calculated value of γ is 0.9254, while the ideal skewness is 0.6053. The fact that the two skewness values are somewhat closer ($\Delta = 0.3201$) suggests that, while some degree of cheating was present, less occurred than in the first case. This is demonstrated by the chi-square calculation, which yields a value of $\chi^2 = 5.675$. The confidence limits corresponding to the 97.5% confidence interval are again 2.7 and 19.0, and χ^2 falls well within the range, so this demonstrates that the Pólya distribution is a good fit for the data. The model parameter values are $p = 0.8564$ and $r = 16.2701$. Note that the higher value of r , as mentioned previously, indicates that the fitted Pólya distribution is closer in nature to the ideal Poisson distribution, and as such implies less cheating.

The first final assessment was taken by 33 students ($n = 33$) and consisted of 100 multiple-choice questions ($q = 100$). The total number of unique pairwise comparisons is 528. Figure 3 shows the corresponding distribution.

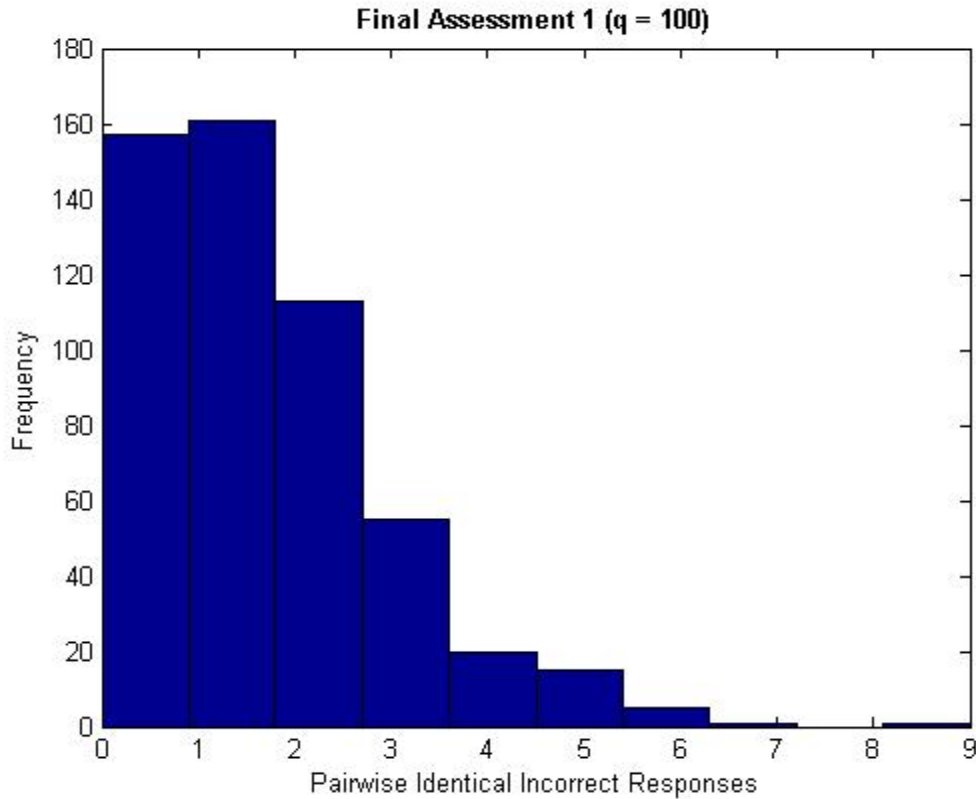


Figure 3. Pairwise identical incorrect responses, Final Assessment 1.

Study parameter values for Final Assessment 1 were $t = 2.0$ hours, a non-random order of questions presented on the assessment, and allowed revisitation.

The calculated mean, μ , is 1.4261. The calculated value of γ is 1.2895, and the ideal skewness is 0.8374. The fact that the difference between the two skewness values ($\Delta = 0.4521$) is greater indicates that some degree of cheating was present, and in fact close to that in the case of the first section assessment. This is demonstrated by the chi-square calculation, which yields a value of $\chi^2 = 4.97$, again close to that of the 5.02 value in the first case. The confidence limits corresponding to the 97.5% confidence interval are again 2.7 and 19.0, and χ^2 again falls well within the range, so the Pólya distribution is a good fit for the data. The model parameter values are $p = 0.7293$ and $r = 3.8424$. Note that the model parameter values are also quite close to those found with Section Assessment 1, and note, as well, that the study parameter values in both cases were a non-random question order and allowed revisitation. The difference between the two cases was the extra time allowed for the final assessment, and it is interesting to see that, because χ^2 is slightly less in the case of the final, the implication is that slightly more cheating occurred on the final assessment than on the first.

The third section assessment (the first assessment of the second course) was taken by 18 students ($n = 18$) and consisted of 50 multiple-choice questions ($q = 50$). The total number of unique pairwise comparisons in this case is 153. Figure 4 shows the corresponding distribution.

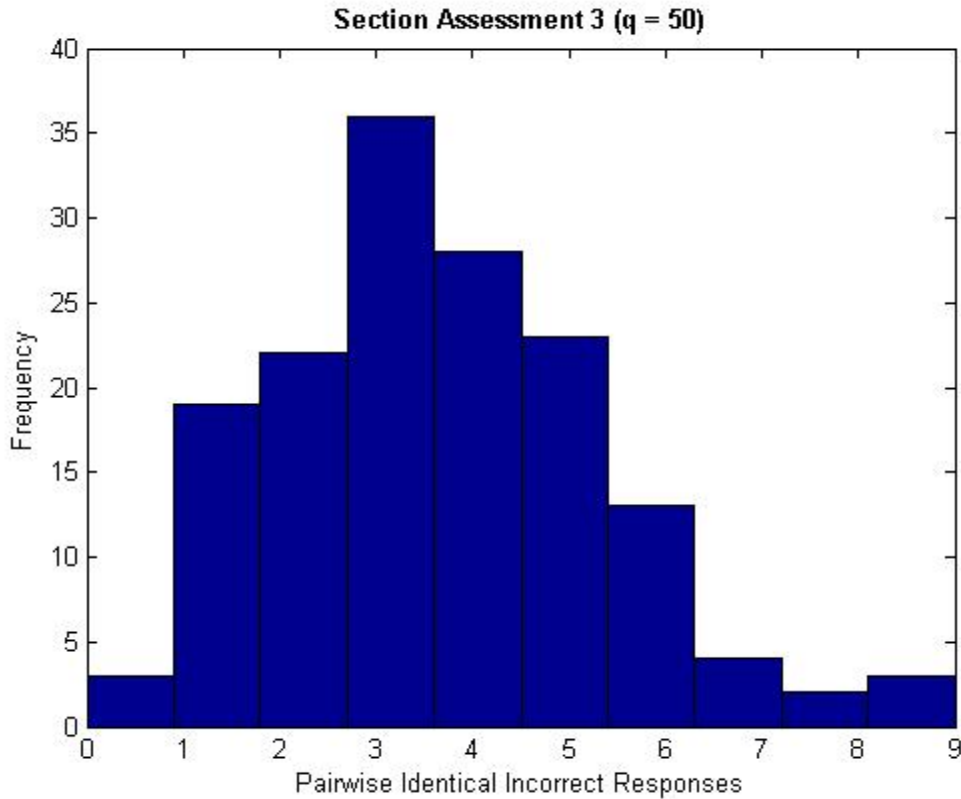


Figure 4. Pairwise identical incorrect responses, Section Assessment 3.

Study parameter values for Section Assessment 3 were $t = 1.0$ hours, a random order of questions presented on the assessment, and allowed revisitation. Note that the value of t was decreased in each of the remaining assessments.

The calculated mean, μ , is 3.5752. The calculated value of γ is 0.5146, and the ideal skewness is 0.5288. The difference between the two skewness values is quite small ($\Delta = -0.142$), and indicates that minimal cheating was present. The initial attempt to fit the data to the Pólya distribution failed due to the lack of convergence of the nonlinear program to calculate the values of p and r , but algorithm convergence was achieved using a nonzero value of k . The chi-square calculation then yielded a value of $\chi^2 = 20.3$. The confidence limits corresponding to the 97.5% confidence interval are again 2.7 and 19.0, but χ^2 falls outside this range, so while the Pólya distribution does not fit the data within these limits, the fit becomes acceptable if we increase the confidence level to the 99%.

The fourth section assessment (the second assessment of the second course) was taken by 17 students ($n = 17$) and consisted of 50 multiple-choice questions ($q = 50$). The total number of unique pairwise comparisons is 136. The corresponding distribution is shown in Figure 5.

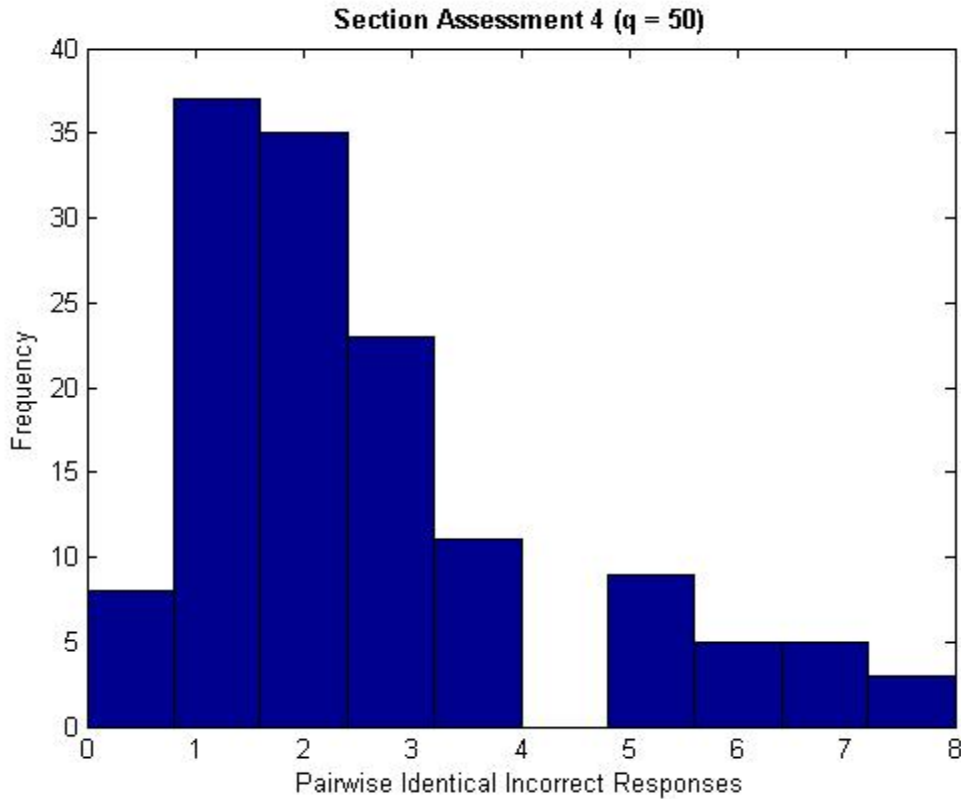


Figure 5. Pairwise identical incorrect responses, Section Assessment 4.

Study parameter values for Section Assessment 4 were $t = 1.0$ hours, a random order of questions presented on the assessment, and disallowance of question revisitation.

The calculated mean, μ , is 2.6029. The calculated value of γ is 1.0483, and the ideal skewness is 0.6198. The difference between the two skewness values ($\Delta = 0.4285$) is greater than that seen in the case of Section Assessment 3, and, as was the case with that assessment, the initial attempt to fit the data to the Pólya distribution failed due to the lack of convergence of the nonlinear program. Algorithm convergence was again achieved using a nonzero value of k . The chi-square calculation, however, yields a value of $\chi^2 = 233.7$, which widely exceeds the 97.5% confidence level, indicating that the Pólya distribution does not fit the data acceptably.

The second final assessment was taken by 17 students ($n = 17$) and consisted of 100 multiple-choice questions ($q = 100$). The total number of unique pairwise comparisons in this last assessment of the study is, again, 136. Figure 6 shows the corresponding distribution.

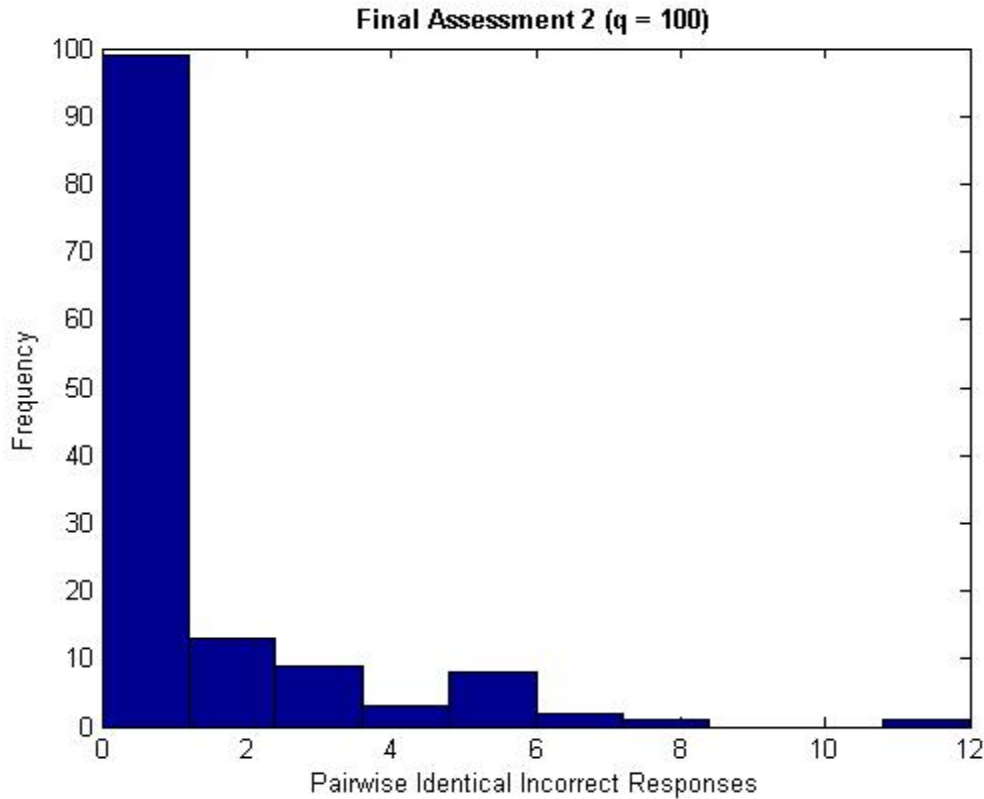


Figure 6. Pairwise identical incorrect responses, Final Assessment 2.

Study parameter values for Final Assessment 2 were $t = 1.5$ hours, a non-random order of questions presented, and disallowance of question revisitation.

The calculated mean, μ , is 1.2868. The calculated value of γ is 2.3856, and the ideal skewness is 0.8815. The fact that the difference between the two skewness values is rather large ($\Delta = 1.5041$) indicates that a significant degree of cheating was present. This is demonstrated by the chi-square calculation, which yields a value of $\chi^2 = 13.67$. The confidence limits corresponding to the 97.5% confidence interval are again 2.7 and 19.0, and χ^2 falls within the range, so the Pólya distribution fits the dataset well. The model parameter values are $p = 0.048$ and $r = 0.1348$. The low value of r substantiates the inference that significant cheating occurred on this exam; as noted earlier, the Pólya probability mass function converges to that of the Poisson distribution as $r \rightarrow \infty$.

The study parameters and χ^2 values are summarized for ease of reference in Table 1 below.

Table 1. *Summary of Study Parameters and Goodness-of-fit Results.*

Assessment	<i>t</i> (hours)	Random	Revisitation	χ^2
S. A. 1	1.5	No	Yes	5.02
S. A. 2	1.5	Yes	Yes	5.68
F. A. 1	2.0	No	Yes	4.97
S. A. 3	1.0	Yes	Yes	20.3
S. A. 4	1.0	Yes	No	233.7
F. A. 2	1.5	No	No	13.67

ANALYSIS

Given the structure of the study as it was conducted, it is both possible and instructive to compare the effects of variation of the three parameters of interest by examining the section assessments. We first examine changes in the parameter t ; specifically, a reduction in that parameter with the other parameters held constant. One can readily do this by comparing the results from Section Assessments 2 and 3. The reduction in the allowed test time by half an hour (33.3%) led in this case to an increase in χ^2 by a factor of 3.6, indicating that the degree of cheating decreased, since a higher value of χ^2 implies a poorer fit of the data to an ideal Pólya distribution. From a more empirical viewpoint, in general, the reduction of allowed test times between the first course (Section Assessments 1 and 2 and Final Assessment 1) and the second (Section Assessments 3 and 4 and Final Assessment 2) led to considerable increases in each relevant value of χ^2 . This makes good sense, since unauthorized collaboration between students requires time, and is less possible when less time per question is allowed for the assessment.

Next, we examine a change from a non-random order of question presentation to a random order. It should be noted again that the question pool sizes for the random-ordered assessments were the same as the number of questions in the assessments; in other words, all available questions were presented to each student. Section Assessments 1 and 2 may be compared to provide an insight to the value of randomization. It can be seen from this comparison that use of randomization, all other parameters being constant, leads to an increase in χ^2 , and thus a decrease in cheating, as expected.

Lastly, we examine the disallowance of question revisitation by comparing Section Assessments 3 and 4. Note that a substantial increase in χ^2 occurs (by a factor of 11.5) when revisitation is disallowed. Note also, though, that the skewness actually increases as a result of the disallowance, and when this is taken in the context that the programming methods for fitting each of these two assessments originally failed to converge to solutions for the Pólya distribution parameters, it is evident that this comparison is not particularly useful. A more valid comparison can be made between Final Assessments 1 and 2, in which

the disallowance of revisitation led to an increase in χ^2 as calculated for Final Assessment 2. This occurred, however, with a corresponding decrease in the time allowed for the assessment, so while this provides an indication that the disallowance of revisitation leads to a decrease in cheating, the results are not entirely conclusive.

CONCLUSION AND FURTHER INVESTIGATION

The study presented here demonstrates that a strong positive correlation exists between time allowed for the online assessment and frequency of cheating, and that randomization of question order is also moderately to strongly correlated with reduced cheating. Also, while indications are that disallowing question revisitation reduces cheating, especially when coupled with reduced testing times, numerical results related to this parameter are inconclusive.

Two potential items for further research are an additional study of the revisitation parameter so that it may be decoupled from the test time parameter, and the inclusion of additional parameters in the study; of particular interest are the effects of the immediate release of the results to the student (and the determination of whether this affects the mean of the actual assessment scores), and the effects of the broadening and narrowing of the assessment availability window.

Additional research in this area might include a study to see whether a correlation exists between exam scores and the degree of cheating present on subsequent exams; i.e., whether lower exam scores lead to increased cheating later in the course due to a sense of desperation on the part of some students. Finally, further study in a somewhat different direction might address the question of what we as instructors can actually do when cheating is detected by the methods suggested herein; how the problem students should be handled, and how the perpetrators might be separated from the victims.

While the benefits of using online assessments in fully-online, hybrid, and traditional methods of course delivery are many, it is important that the distance educator become aware of the opportunities for academic dishonesty that exist in the online delivery environment, understand how to minimize those opportunities, and avail himself or herself of the technology to detect dishonesty when it does in fact occur. Only then will the integrity of the environment reach the level that it must in order for the promise of distance learning to realize its fullest potential.

REFERENCES

- Abbott, L., Siskovic, H., Nogues, V., & Williams, J. (2000). Learner assessment in multimedia instruction: considerations for the instructional designer. Paper presented at the Society for Information Technology & Teacher Education International Conference, San Diego, CA. Retrieved from <http://eric.ed.gov>. (ED444516)
- Burlak, G., Hernández, J., Ochoa, A., & Muñoz, J. (2006). The use of data mining to determine cheating in online student assessment. In *Proceedings of the 2006 Electronics, Robotics and Automotive Mechanics Conference*, 1, 161-166. Washington, DC: IEEE Computer Society.
- California State University, Sacramento. (n.d.). Online exams: strategies to minimize cheating. Sacramento: California State University System. Retrieved from <http://www.csus.edu/webct/faculty/online-exams-best-practices.pdf>
- Cooper, L. (2000). Online courses: tips for making them work. *THE Journal* [Online], 27(8), 86-92. Retrieved from <http://thejournal.com/Articles/2000/03/01/Online-Courses--Tips-for-Making-Them-Work.aspx>
- Crown, D., & Spiller, M. (1998). Learning from the literature on college cheating: a review of empirical research. *Journal of Business Ethics*, 17, 683-700.
- Dick, M., Sheard, J., Bareiss, C., Carter, J., Joyce, D., Harding, T., & Laxer, C. (2003). Addressing student cheating: Definitions and solutions. *ACM SIGCSE Bulletin*, 35(2), 172-184.
- Dirks, M. (1998). How is assessment being done in distance learning? Paper presented at the Northern Arizona University Web '98 Conference, Flagstaff, AZ. Retrieved from <http://eric.ed.gov>. (ED423273)
- Eplion, D., & Keefe, T. (2005, April). On-line exams: strategies to detect cheating and minimize its impact. Paper presented at the Tenth Annual Instructional Technology Conference, Murfreesboro, TN. Retrieved from <http://frank.mtsu.edu/~itconf/presentations/eplion.html>
- Illinois Online Network. (n.d.). Strategies to minimize cheating online. Champaign-Urbana: University of Illinois System. Retrieved from <http://www.ion.uillinois.edu/resources/tutorials/assessment/cheating.asp>
- Kaczmarczyk, L. (2001). Accreditation and student assessment in distance education: why we all need to pay attention. In *Proceedings of the 6th Annual Conference on Innovation and Technology in Computer Science Education*, 113-116. New York, NY: Association for Computing Machinery.
- Kim, N., Smith, M., & Maeng, K. (2008). Assessment in online distance education: a comparison of three online programs at a university. *Online Journal of Distance Learning Administration* [Online], 11(1). Retrieved from <http://www.westga.edu/~distance/ojdl/spring111/kim111.html>.
- Mason, D., & Voit, D. (1998). Integrating technology into computer science examinations. In *Proceedings of the Twenty-Ninth SIGCSE Technical Symposium on Computer Science Education*, 140-144. New York, NY: Association for Computing Machinery.

- Meyen, E.L., Aust, R. J., Bui, Y. N., & Isaacson, R. (2002). Assessing and monitoring student progress in an e-learning personnel preparation environment. *Teacher Education and Special Education*, 25(2), 187-198.
- McMurtry, K. (2001). E-cheating: combating a 21st century challenge. *T.H.E. Journal* [Online], 29(4), 36-38, 40-41. Retrieved from <http://eric.ed.gov>. (EJ641503)
- Olt, M. (2002). Ethics and distance education: strategies for minimizing academic dishonesty in online assessment. *Online Journal of Distance Learning Administration* [Online], 5(3). Retrieved from <http://www.westga.edu/~distance/ojdla/fall53/olt53.html>.
- Pain, D., & Le Heron, J. (2003). WebCT and online assessment: the best thing since SOAP? *Educational Technology & Society*, 6(2), 62-71.
- Rogers, C. (2006). Faculty perceptions about e-cheating during online testing. *Journal of Computing Sciences in Colleges*, 22(2), 206-212.
- Rowe, N. (2004). Cheating in online student assessment: beyond plagiarism. *Online Journal of Distance Learning Administration* [Online], 7(2). Retrieved from <http://www.westga.edu/~distance/ojdla/summer72/rowe72.html>.