

Статья поступила в редакцию 13.08.2022 г.

DOI: 10.24412/2687-0053-2022-4-120-123 EDN: LNKNTNP

**Информация для цитирования:**

Бушмакин А.Д., Луговой К.А., Помешкин Е.В., Тришкин А.Г., Курганова Л.В., Зуева Г.П., Лесников А.И., Бочкарникова А.Г., Елгина С.И., Мозес В.Г., Рудаева Е.В., Мозес К.В. ВЕДЕНИЕ ПАЦИЕНТА С СЕКРЕТОРНОЙ ФОРМОЙ АЗОСПЕРМИИ // Медицина в Кузбассе. 2022. №4. С. 120-123.

**Dr. Srinivasa J,**Zydus Medical College and Hospital,  
Dahod, Gujarat, India

## ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОБНАРУЖЕНИЯ СГОВОРОВ В ОНЛАЙН-ЭКЗАМЕНАХ С МНОЖЕСТВЕННЫМ ВЫБОРОМ – ОБЗОР

Компьютеры использовались в высших учебных заведениях для оценки студентов на протяжении десятилетий. Программное обеспечение, позволяющее проводить оценивание учащихся (часто называемое компьютерным оцениванием или САА), стало широко коммерчески доступным в 1990-х годах, и многие учебные заведения начали экспериментировать с этими пакетами. Растущая доступность сетевых компьютеров к середине 1990-х годов позволила предоставлять оценки и другие образовательные услуги в Интернете с помощью веб-браузеров, но проблема этого онлайн-теста заключается в сговоре и мошенничестве. По мере того, как онлайн-экзамены становятся все более популярными, возможно, с возможностью того, что учащимся будет разрешено сдавать их удаленно, вероятность списывания может быть намного выше, чем на традиционном экзамене. Экзамен, который проводится дистанционно, нуждается в некоторой форме обеспечения качества, чтобы предотвратить или обнаружить мошенничество и сговор между кандидатами. Экзамены с множественным выбором (MCQ) представляют особый интерес для оценивания.

Следовательно, эта статья посвящена обзору различных программ, таких как LERTAP, индекс Harpp-Hogan, Scrutiny!, Integrity и SCheck, которые могут решить проблему сговора при онлайн-экзамене MCQ.

**Ключевые слова:** MCQ; онлайн-экзамен; программное обеспечение; сговор

**Dr. Srinivasa J,**Zydus Medical College and Hospital,  
Dahod, Gujarat, India

### COLLUSION DETECTION SOFTWARE IN ONLINE MULTIPLE CHOICE EXAMINATIONS – A REVIEW

Computers have been used in higher education to assess students for decades. Software to allow assessments to be delivered to students (often referred to as computer-assisted assessment, or CAA), became widely commercially available in the 1990s, and many institutions began to experiment with these packages. The increasing availability of networked computers by the mid-1990s allowed assessments and other educational services to be delivered online with web browsers but the problem of this online test is collusion and cheating. As on-line examinations become more popular, perhaps with the possibility of students being allowed to sit them remotely, the opportunity for cheating could be much higher than the traditional examination. An examination that is sat at a distance need some form of quality assurance to prevent or detect cheating and collusion between candidates. Multiple choice question (MCQ) examinations are of particular interest in assessment.

Hence this article focuses on overview of various software's like LERTAP, The Harpp-Hogan index, Scrutiny!, Integrity and SCheck that can tackle the problem of collusion in online MCQ examination.

**Key words:** MCQ; online examination; software; collusion

Almost overnight, the COVID-19 pandemic forced all classes online, and along with the classes, all exams. This meant that human proctoring was no longer possible, not for formerly face-to face classes, nor for distance-ed classes, which often depended on employers, libraries, or testing centers to supply proctors for individual students. This raised a great deal of concern that cheating would undermine the integrity of exams, and thus subvert the validity of grades.

The emergence of a range of learning and teaching environments beyond the traditional face-to face environments, the increased complexity of academic

work, and heightened expectations of students has imposed a new set of challenges on teaching staff. Multiple-choice question examinations continue to be popular for both formative and summative assessments. The recent advances in computer, web and network technology make online administering of MCQ exams an increasingly attractive and feasible option. Online MCQ examinations have become one of the assessment methods in several universities. The principal motivation for these weekly/monthly quizzes is to enhance the students' learning but collusion has been identified as a major problem, that need to either cancel the online MCQ test and decide to opt for

a traditional approach. In fact, there would be cheating observed with the traditional approach. As we know that the main challenge of online testing is to solve the problem of plagiarism and collusion [1, 2]. This article focuses briefly on different software that can tackle the problem of collusion in online MCQ examination. The implementation of these methods would make difficult for students to cheat and also discouraging them from being dishonest.

There are different ways the students can cheat in the online quiz:

- When large number of students is taking test at the same time the students can typically see the screen of the person sitting next to them easily.

- Students who have taken the test can pass the answers to others.

- There are various cheating devices (technological and otherwise), e.g., it would be possible to communicate answers via a small undetectable mechanism like a vibrating mobile phone to each others.

There are several studies of sociological factors that motivate rational cheating behavior including specific work on behavioral modeling and gender differences [3, 4].

There are a few statistical methods to detect collusion during examination. An earliest statistical method that documented in literature was bird index proposed by bird (1927, 1929) [5, 6]. For pairs of examinees, Bird suggested three approaches which are based on the inspection of observed distributions of the number of identical wrong responses. Subsequently, Crawford devised his index in 1930 similar to Bird's procedure which is also based on the percentage of pair's wrong answers that were similar. Crawford computes the index using a test of the difference between proportions [7].

Cody RP (1985) has used correlated errors as a simple measure of possible collusion in medical MCQ examinations [8]. In another study by Angove WH (1974), observed that the regression analyses between several possible indices to observe correlations in errors and runs of correlated results between pairs of candidates [9]. Kvam PH has recommended that a class be given two subtly different examination papers, so that the copier writes down what is, in effect, the wrong answer. In this way, and by employing a maximum likelihood calculation, it was possible simultaneously to penalize and detect cheats, but this approach is only appropriate in an invigilated environment [10].

There are several disadvantages of using correlation analysis for detection of collusion such as it is necessary to take into account the ability of the students. When comparing two very able students, it would not be surprising to observe a significant number of correlated correct answers, simply because they got so many correct. Similarly, in negatively marked tests, two very risk averse students would be expected to show a relatively high degree of correlation in terms of preferentially choosing a penalty-free 'don't know' option if it is available.

A statistical test for answer copying on multiple-choice tests based on Cohen's kappa was developed by Sotaridona, van der Linden and Meijer (2006) [11]. The test is free of any assumptions on the response processes of the examinees suspected of copying and having served as the source, except for usual assumption that these processes are probabilistic.

Belov and Armstrong (2010) have proposed a bi-stage approach which combines two statistical approaches in successive stages. The first stage uses Kullback-Leibler divergence to identify examinees, called subjects, who have demonstrated inconsistent performance during an exam. For each subject the second stage uses the K-Index to search for a possible source of the responses. Both stages apply a hypothesis test given a significance level. Computational details for Kullback-Leibler divergence index can be found in Belov and Armstrong (2010) [12].

## SOFTWARE PROGRAMS

### 1. Lertap

The Laboratory of Educational Research Test Analysis Package, «LERTAP», is a classical item and test analysis system. Lertap also analyzes surveys and mastery tests. Lertap's original RSA method was based on the "Harpp-Hogan index", also known as the Harpp-Hogan ratio. It uses response similarity analysis (RSA) methods to detect cheating (Larry R. Nelson 2006) [13, 14]. The H-H index is based on two characteristics of the students' item responses: the number of exact errors in common (EEIC) and the number of different responses (D). The H-H index is expressed as a ratio of these two numbers:  $HH = EEIC/D$  [14]. Two students are said to have an "exact error in common" when they both select the similar distractor to an item, that is, when they choose exactly the same incorrect answer to an item. Harpp, Hogan, & Jennings described it to be "a powerful indicator of copyin". They mentioned that analyses of well over 100 examinations during the past six years have shown that when this number is  $\sim 1.0$  or higher, there is a powerful indication of cheating. In virtually all cases to date where the exam has  $\sim 30$  or more questions, has a class average  $< 80\%$  and where the minimum number of EEIC is 6, this parameter has been nearly 100% accurate in finding highly suspicious pairs. But according to a study by Larry R Nelson (2006), have concluded that the H-H index should be used with great caution. RSA is use to see if the responses of any two test takers were "excessively similar", even previous study by Wesolowsky supported this findings [1]. So if RSA is carefully used might check the possible presence of cheating in an examination environment. In Mid-July 2012, Assessment System Corporation released Lertap 5.10 for use with Excel 2010. There were three things particularly addressed in this current edition: providing access to more immediate on-line help for users, getting Lertap to provide more warning flags for items

which may have problems, and to make "packed plots" easier to get.

## 2. Scrutiny!

It is a commercial package which may be obtained from Assessment Systems Corporation. This software helps in the detection of possible cheating and/or test compromise. Scrutiny! uses error similarity analysis to identify examinees whose responses are suspiciously similar and provides precious information to support other indications of possible misconduct. Cizek (2001) stated that Scrutiny! is "easy to use", and "is compatible with many common input file formats". But Cizek also related that Scrutiny! uses "a method which, unfortunately, has not received strong recommendation in the professional literature" [15]. However, Scrutiny! software is no longer available through Assessment Systems Corporation.

## 3. Integrity

This is a wide-ranging system which includes not less than five different methods of cheating detection, among them the "g2" procedure developed by Frary, Tideman, and Watts (1977) [16]. Using Integrity involves an off-line "batch" process somewhat reminiscent of mainframe computing: (1) the two data files required by the program are prepared on the user's computer; (2) the files are uploaded to the Integrity computer via the internet; and, (3) after a period of time, Integrity's results are then downloaded, again using the internet. Users don't have to wait for their job to finish – once a job is submitted in step (2), a user may turn off his/her computer, and re-connect to the Integrity computer at a later time.

## 4. The SCheck program

This is the name of a program written by Professor Wesolowsky of McMaster University, Canada, which is published in the Journal of Applied Statistics in 2000 [1]. The g2 collusion index seen in Integrity stems from what Wesolowsky (2000) has referred to as the "seminal work" of Frary, Tideman, and Watts (1977) [16]. Wesolowsky's paper presented a modification to Frary et al. In researching Wesolowsky's modification, Tideman and Kheirandish (2003) [17] found it had "noticeably better power than the probabilities suggested by Frary et al.". Better power means that Wesolowsky's method has a greater likelihood of rejecting the response-independence hypothesis if the hypothesis is in fact false – it is more capable of detecting possible cheating, less likely to make a Type II error.

Like Integrity, SCheck has clear value for those interested in detecting cheating on multiple-choice exams, even though its output is not extravagantly formatted. Tideman and Kheirandish (2003) gave an edge for SCheck in terms of its methodology [17], and Lertap users will find an in-built interface which eases the process of preparing data for input to SCheck.

Despite the large number of tools, plagiarism-detection software is not used very widely. According to the Wiley survey [18], only 4 % of instructors reported using it, far less than the 16 % who used webcam monitoring, or the 15 % who used lockdown browsers. The reasons for this are many. Detection methods without applying prevention methods could not be effective. As cheating detection and prevention methods are evolved, new cheating types and technologies emerge as well. Consequently, no system can mitigate all kinds of cheating in online exams, and more advanced methods should be employed. It seems the most efficient strategy for cheating handling is to lower cheating motivation [19]. Tools that engages a single statistical test are easy to use, but may not detect all plagiarism. Tools with multiple statistical tests show more, but are harder to interpret. A brief review of methods which are developed and reported in literature is provided in this article. Additionally, methods or indices for detecting collusion are also compared with respect to their effectiveness and the practicality of their application for different groups in terms of their performance. As there are pros and cons of each software used hence there is a need for development a comprehensive method which can tackle the collusion in online MCQ examination and overcome the drawbacks of various methods.

## CONCLUSION

Multiple choice question (MCQ) examinations are of particular interest in assessment because they offer a tractable method of assessing ability across a wide range of topics. It is expected that all of the software systems stated in this article will continue to improve. In conclusion, the appropriate software used should have a good user-friendliness, plagiarism prevention, and above all, software should detect cheating. We feel it is good to use a combination of technologies to provide the support best suited to the assessment, rather than trying to fit the assessment to an off-the-shelf e-learning environment.

## REFERENCES:

1. Wesolowsky GO. Detecting excessive similarity in answers on multiple choice exams. *J Appl Stat.* 2000; 27(7): 909-921.
2. Ercole A, Whittlestone KD, Melvin DG, Rashbass J. Collusion detection in multiple choice examinations. *Medical education.* 2002; 36(2): 166-172.
3. Tibbetts SG. Differences between women and men regarding decisions to commit test cheating. *Res High Educ.* 1999; 40: 323-342.
4. Tibbetts SG. Gender differences in students' rational decisions to cheat. *Deviant Behav.* 1997; 18: 393-414.
5. Bird C. The detection of cheating in objective examinations. *School and society.* 1927; 25: 261-262.
6. Bird C. An improved method of detection cheating in objective examinations. *Journal of Educational Research.* 1929; 25: 261-262.

- 7 Muhammad NK, Zahid M, Naeem AR. Statistical Methods for Answer Copying – A Brief Overview. *British Journal of Arts and Social Sciences*. 2011; 1: 49-61.
- 8 Cody RP. Statistical analysis of examinations to detect cheating. *J Med Educ*. 1985; 60: 136-137.
- 9 Angoff WH. The development of statistical indices for detecting cheaters. *J Am Stat Assoc*. 1974; 69: 44-49.
- 10 Kvam PH. Using exam scores to estimate the prevalence of classroom cheating. *Am Stat*. 1996; 50: 238-242.
- 11 Sotaridona LS, Vander LWJ, Meijer RR. Detecting answer copying using the kappa statistic. *Applied Psychological Measurement*. 2006; 30: 412-431.
- 12 Belov DI, Armstrong RD. Automatic detection of answer copying via kullback-leibler divergence and k-index'. *Applied Psychological Measurement*. 2010; 34: 379-392.
13. Larry R. Nelson (2006). Using selected indices to monitor cheating on multiple choice exams. <http://www.lertap.curtin.edu.au/Documentation/JERM2006.doc>. Date accessed 15 June 2013.
14. Nelson LR. (2006) Using Lertap 5.6 to monitor cheating on multiple choice exams. <http://www.lertap.curtin.edu.au/Documentation/JERM2006mod1.doc>. Date accessed 13 June 2013.
15. Cizek GJ. (2000) An overview of issues concerning cheating on large-scale tests. Paper presented at the Annual Meeting of the National Conference on Measurement in Education, Seattle, Washington.
16. Frary RB, Tideman TN, Watts TM. Indices of cheating on multiple-choice tests. *Journal of Educational Statistics*. 1977; 2: 235-256.
17. Tideman N, Kheirandish R. Structurally consistent probabilities of selecting answers. *Journal of Applied Statistics*. 2003; 30: 803-811.
18. Wiley, Academic Integrity in the Age of Online Learning, <http://read.uberflip.com/i/1272071-academic-integrity-in-the-age-of-online-learning/4?> [accessed March 6, 2021]
19. Noorbehbahani F, Mohammadi A, Aminazadeh M. A systematic review of research on cheating in online exams from 2010 to 2021. *Educ Inf Technol*. 2022; 27(6): 8413-8460. <https://doi.org/10.1007/s10639-022-10927-7>

#### Acknowledgement

This article was a part of GCHE program, Monash University Malaysia . Author wish to thank the Faculty of Higher Education, Monash University for their guidance.

**Корреспонденцию адресовать:** Dr. Srinivasa J, Professor, Department of Physiology, Zydus Medical College and Hospital, Dahod, Gujarat, India, former Faculty of Monash University Malaysia. ORCID: 0000-0001-9473-8011