An Exploratory Case Study of the Use of Video Digitizing Technology to Detect Answer-Copying on a Paper-and-Pencil Multiple-Choice Test

Carlos Zerpa and Christina van Barneveld
Lakehead University
czerpa@lakeheadu.ca

Abstract — In educational testing, answer-copying is considered a behaviour that poses threats to the validity of test scores interpretations, which is a concern when interpreting the test results for the purpose of making changes to curriculum and educational policies. Answer-copying involves at least two examinees, one being the source and the other the copier. While different methods have been developed to detect answer-copying using statistical indices, researchers have not yet examined the use of video digitizing technology via a kinematics analysis of the data to detect answer-copying during test-taking situations. The purpose of this case study was to explore the use of video digitizing technology to detect answer-copying by measuring examinees’ response time, displacement and velocity from item to item on a test. A sample of two university students volunteered to demonstrate the benefits and challenges of using video digitizing technology to detect answer-copying. While this is a small scale demonstration, the outcome of this study may shed light on whether or not the use of video digitizing technology provide evidence of feasibility and some preliminary reliability in the detection of answer-copying. The lesson learned from this study can inform the direction of a future program of research.

Keywords:
Answer-Copying, Validity, Behaviour

INTRODUCTION

Answer-copying involves at least two examinees, one being the source and the other one the copier (Ark, Emos & Sijtsma, 2008). Answer-copying is a problem in the educational system because it can inflate or deflate examinees’ scores on a test. Answer-copying behaviour is more likely to be observed when the stakes are high for the test takers. That is, the examination entails admission, licensing, or certification testing (van der Linden, 2009; Wollack, 2006). Answer-copying behaviour is considered a threat to the validity of examinee test score interpretations in the educational system and society (Ark, Emos & Sijtsma, 2008; Meijer & Sijtsma, 1995; 2001).

Over the past 25 years, researchers in the field of educational measurements have proposed and used statistical indices to detect answer-copying and minimize threats to the validity of test score interpretations (Landis & Koch, 1997; Meijer, 1996; Sotaridona & Meijer, 2003; Sotaridona, van der Linden & Meijer, 2006; van der Linden & Sotaridona, 2004, 2006; Wollack, 2006). For example, Sotaridona and Meijer (2003) proposed two indices to detect answer-copying on a multiple choice test. These two indices are called the S1 and S2. The S1 index is used to detect the number of wrong answer matches between two examinees, the source and the copier. One of the limitations of the S1 index, however, is that it fails to detect answer-copying when the examinees’ responses are based on correct answer matches. In order to overcome this limitation, Sotaridona and Meijer (2003) proposed the S2 index. The S2 index is an improvement in answer-copying detection rate over the S1 because it incorporates information from correct and incorrect responses between two examinees, the source and the copier.

The problem with the S1 and S2 statistical indices (Sotaridona & Meijer, 2003) is that they are population size dependent. For example, if a pair of examinees is suspected of answer-copying, it may be possible that these examinees may get a different value for a copying statistics if their pair of response vectors gets included in a different population. In addition, when the sample is small parameters cannot be estimated reliably.

A better technique to detect answer-copying and overcome the population size dependency problem of the S1 and S2 indices was proposed by Sotaridona, van der Linden, and Meijer (2006). These researchers investigated the use of Kappa statistics by Cohen (1960) as a copying statistic with a null distribution that was population size independent. For instance, if a pair of examinees was suspected of answer-copying, their
distribution would be independent of any other test takers behaviour distributions. In addition, Sotaridona, van der Linden, and Meijer (2006) stated that the advantage of using Kappa statistics is not only that it is sample size independent but also that it can be used for a single pair of examinees such as the source and the copier, without requiring information from any other examinees. For their study, Sotaridona, van der Linden, and Meijer (2006) generated 1000 response patterns of simulated data. The score patterns were grouped into six score categories, which produced 2000 response vectors between two examinees, the source and the copier. The researchers found that Kappa statistics resulted to be a powerful function in detecting answer-copying and minimizing type I error rates when interpreting test results. The Kappa statistics used in their study was given by the following equation:

\[ k = \frac{P_o - P_e}{1 - P_e} \]

Where

- \( P_o \): was the probability of an observed agreement between two examinees.
- \( P_e \): was the probability agreement due to chance if the examinees operated independently.

While Kappa statistics is a more robust method to detect answer-copying between two examinees because it is population size independent (Sotaridona et al., 2006), the method cannot identify who is the source and who is the copier. Based on this concern, there is a need to explore and examine other avenues and techniques to detect answer-copying (Wollack, 2006). One technique that researchers have not yet examined is the role of video digitizing technology in combination with a kinematics analysis of the data and Kappa statistics as an avenue to detect answer-copying and identify the source and the copier. Kinematics is the study of movement of an object without taking in consideration the cause of the movement but rather where and when the movement happens (Wang, Hu & Tan, 2003). Through a sequence of video digitizing images, it is possible to track human movement kinematics and establish a coherent relationship between video frames and human movement related to position, speed and time (Wang, Hu & Tan, 2003). For example, it is possible to track the movement of an examinee hand holding a pencil and measure the examinee displacement, velocity and response time from item to item on a test.

This case study examined the role of video digitizing technology as an avenue to possibly improve the sensitivity of Kappa statistics to not only detect answer-copying but also better identify the source and copier. The question that guided this case study was: To what extent is the use of video digitizing technology via a kinematics analysis of the data able to detect answer-copying and identify the source and the copier when combined with Kappa statistics?

**METHOD**

**Participants**

Two fourth year kinesiology students from Lakehead University volunteered to participate in this case study. One student was asked to be the source and the other to be the copier. Students were given a second year biomechanics exam and were asked to answer 30 questions from the exam to the best of their abilities. During the examination, only the copier was allowed to answer copy from the source.

**Instruments**

For this case study, two testing stations were set up side by side. That is, one for the source and the other one for the copier. Each station contained a video camera set 90 degrees with respect to the desk. The video cameras were used to track examinees’ responses on a scantron sheet for each item on the test as depicted in Fig. 1. The examinees’ video cameras were interfaced to a video analysis computer software (Peak Motus) to capture the data. The captured video data were digitized frame by frame per examinee by using the Peak Motus digitizing module to facilitate the computation of linear velocities for each item on the test.

**Procedure**

Examinees were assigned a station and instructed on how to use the instrumentation prior to the test. Each examinee answered 30 questions on a scantron sheet and the captured data from the video technology were digitized for each examinee to compute linear velocities.
from displacement and time measures as the pencil move from item to item on the test.

**Data Analysis**

Descriptive statistics was used to compute and compare the mean velocities between the two examinees (source and copier) for correct and incorrect responses to the test items. Test scores were computed per examinee and Kappa statistics was used to examine the agreement between the two examinees in terms of correct and incorrect responses.

**RESULTS**

The Kappa statistics analysis revealed that the level of agreement between the source and the copier in terms of correct and incorrect responses to the test items was statistically significant, $Kappa = 0.405, p < 0.05$. This outcome, however, was only marginally convincing. As a rule of thumb, values of Kappa from 0.40 to 0.59 are considered moderate, 0.60 to 0.79 substantial, and 0.80 outstanding (Landis & Koch, 1977).

Descriptive statistics as depicted in Fig. 2 revealed that when both examinees agreed in terms of correct responses to the test items, the mean velocity seemed to be higher for the copier ($M=0.155, SD=0.107$) than the source ($M=0.0694, SD=0.038$). How significant this difference was, could not be revealed with this small sample. In Fig. 2, it can be observed, however, that the velocity was higher for the copier than the source at the instant when the copier gave the correct response to the test item throughout the entire exam.

![Figure 2. Mean Velocity for Correct Responses when both Examinees Agreed](image)

Descriptive statistics as depicted in Fig. 3 also revealed that when both examinees agreed in terms of incorrect responses to the test items, the mean velocity was higher for the copier ($M=0.145, SD=0.116$) than the source ($M=0.102, SD=0.072$). Figure 3 also suggests that the velocity was higher for the copier that the source at the instant when the copier gave the incorrect response to a test item.

![Figure 3. Mean Velocity for Incorrect Responses in when both Examinees Agreed](image)

The data as depicted in Fig. 4 suggest that when both examinees disagreed in terms of correct and incorrect responses, the mean velocities were very similar for the copier ($M=0.0812, SD=0.069$) and the source ($M=0.073, SD=0.041$). In addition, it can be observed that the velocities were very similar for the copier and the source at the instant when the copier gave either the correct or incorrect response to the test item.

![Figure 4. Mean Velocity for Correct Responses](image)
DISCUSSION

This case study examined the role of video digitizing technology as an avenue to possibly improve the sensitivity of Kappa statistics to not only detect answer-copying but also better identify the source and copier. The outcome of this study seems to indicate that by using Kappa statistics in combination with video digitizing technology via a kinematics analysis of the data, it is possible to detect answer-copying and better identify the source and copier.

Although Kappa statistics is a very robust algorithm used to detect answer-copying between two examinees because it is population size independent (Sotaridona, van der Linden & Meijer, 2006), the Kappa statistics index lacks sensitivity in identifying who is the source and who is the copier. The video digitizing technology via a kinematics analysis of the data, however, appeared to be promising in detecting answer-copying when both examinees correct and incorrect responses to test items match. Furthermore, the video digitizing technology appeared to improve the sensitive of Kappa statistics because the copier seemed to exert higher velocities than the source when both examinees agreed in terms of correct and incorrect responses to the test items. Similar velocities were obtained, however, when both examinees disagreed in terms of correct and incorrect responses to the test items.

CONCLUSION

The question that guided this case study was: To what extent is the use of video digitizing technology via a kinematics analysis of the data able to detect answer-copying and identify the source and the copier when combined with Kappa statistics? The outcome of this case study seems to provide an avenue to identify the source and copier via Kappa statistics and video digitizing technology to minimize the effect of answer-copying so that more valid interpretations of test results can be made when assessing examinee performance (Ark, Emos & Sijtsma, 2008; Wollack, 2006). As the technology involves, it may become more portable and feasible to be implemented in the classroom and for large scale examinations. The implementation of this technology, however, may have implications for teachers and test organizations to help them make better decisions about educational policies and curriculum changes.

Acknowledgements

The authors would like thank the Social Science and Humanities Research Council (SSHRC) for funding part of this research work.

References


