How Integrity can help evaluate and address issues of academic integrity

http://integrity.castlerockresearch.com
Castle Rock Research Corp.
November 17, 2005

Rational for conducting collusion detection analyses

The rational for the use of a system to identify students who may have engaged in test collusion (i.e., cheating on tests, answer copying) when taking multiple-choice tests lies fundamentally in the idea of creating a "level playing field" for all students. If some students are cheating on tests and obtaining higher grades than they should, then by comparison, students who do not cheat may be disadvantaged. The pressure for those honest students to cheat on tests then increases. Cheating on tests has major implications for students, the public, businesses, and education centers in that they may be hiring, selecting, or passing students who are not properly qualified.

From a test-validity perspective, cheating on tests impacts the inferences that can be made from the results. Test validity is generally defined as how well the test measures what it is supposed to measure. If students are cheating on tests, the validity of the test results is eroded because a "true" reflection of the students' ability (or aptitude, or achievement, etc.) is not being measured: one is no longer as certain as to what the scores mean. Academic centers that do not monitor and address issues of cheating may receive negative reputations from employers who eventually hire their graduates and find that the performance of those graduates is below acceptable standards.

In the popular book *Cheating on Tests, How to Do It, Detect It, and Prevent It,* Dr. Gregory Cizek[1] states that nearly all research conducted on cheating indicates that "cheating is rampant" and that the number of students who admit to cheating has increased in the last 30 years. This is true at all levels of schooling, but especially at the post-secondary level. There is obviously a need for statistical tools, such as those offered through Integrity, that evaluate the academic integrity of test takers and test administration centers.

Role of statistical collusion detection methods

Professional opinions regarding the use of statistical methods to detect cheaters vary. Some people argue that statistical methods to detect cheaters should only be used if a pair of students has already been identified as being suspicious (e.g., a test proctor believes that he/she saw suspicious behavior occurring between two students during an

[1] Cizek, G. J. (1999). *Cheating on Tests, How to Do It, Detect It, and Prevent It.* Lawrence Erlbaum Associates: Mahwah, New Jersey.

examination). Others argue that using statistical tests to identify cheaters in a more exploratory manner is appropriate but should not be the sole source of information.

Clearly a balanced approach is necessary when addressing such a complex and delicate area as test collusion–cheating. Simply broadcasting the systematic use of Integrity at an institution may reduce the number of examinees who consider engaging in collusion. Most educational institutions have protocols in place to investigate allegations of cheating. Using information provided by Integrity as one piece of evidence in the existing investigative process is the recommended approach. It is not recommended that the collusion detection results of Integrity be used as the only source of evidence in an investigation of cheating.

The main purposes of the collusion detection features of Integrity are to: a) provide information on the integrity of writing centers (i.e., are the test-taking processes secure), b) provide information on the academic integrity of examinees through their responses to multiple-choice type questions.

Collusion reports produced by Integrity

The collusion detection analysis is an optional analysis within Integrity. When submitting test data to Integrity, you will be asked via the step-by-step wizard whether you would like to have the collusion detection analysis run. If you select the collusion detection analysis, you will be asked to choose one of three options for how the analysis should be conducted:

1) All students – this option conducts collusion analysis on all students who took the test. It compares the responses on each question for all possible pairs of students.
2) Specific students on all students – this option compares the responses of students that you specify (e.g., suspect) with the responses of all other students who took the test.
3) Only investigate specific students – this option compares the responses of only the students that you specify (e.g., suspect) with each other.

After you have selected the appropriate collusion detection options and finished submitting a job (test) to Integrity, the collusion detection reports are generated. Integrity uses five statistical collusion detection methods. Those interested in learning more about each method are encouraged to review the descriptions of each method within the application by clicking on the terms or visiting the glossary.

An example of a collusion detection summary report is presented below. The example shows three pairs of examinees that have been identified by at least one of the five collusion detection methods. The writing center information (an optional piece of information that can be contained in the data file) is displayed in order to ensure that the pair of examinees took the test at the same writing location.

For each of the five collusion detection methods, four categorizations of *statistical certainty* are possible: Low, Moderate, High, and N/A (not applicable). *Statistical certainty* indicates the level of statistical confidence that a pair of examinees have engaged in collusion.

An example of a collusion detection summary report produced by Integrity.

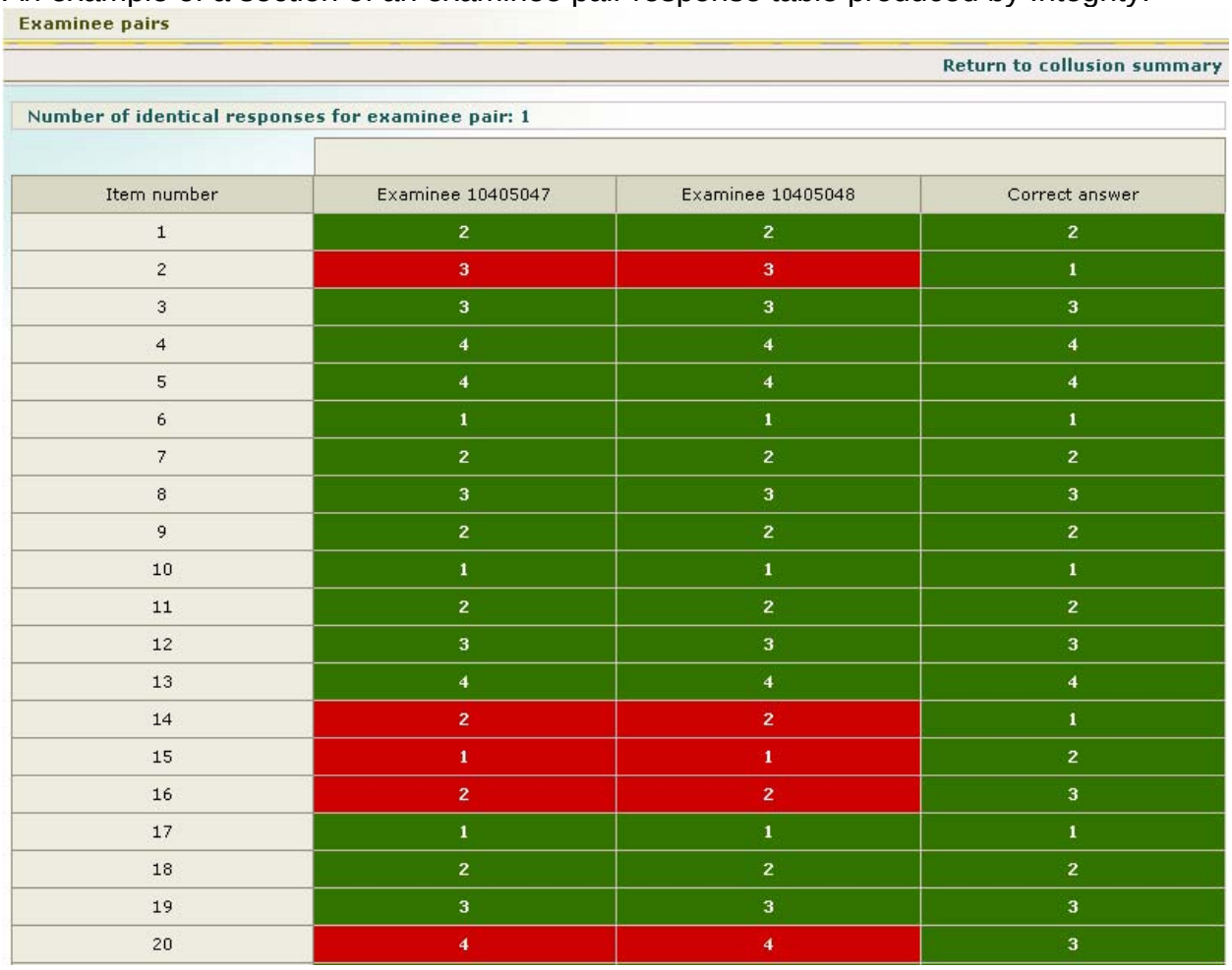
Collusion detection report

	Examinee ID	Writing center	B-Index	PAIR1	PAIR2	MESA	g2
Examinee pair 1	10405047 10405048	1111 1111	High 11.707	High 2900	N/A	Low 1.568E-008	High 8.678 8.678
Examinee pair 2	10505049 10505050	1111 1111	High 10.231	N/A	N/A	Moderate 6.979E-011	High 9.181 9.094
Examinee pair 3	105050149 105150150	3333 3333	Low 7.738	N/A	N/A	N/A	Low 6.063 6.089

The thresholds for these statistical confidence levels were set to minimize false-positives (the flagging of examinees who did not engage in collusion). Even the threshold for “Low” statistical certainty has been set at a very conservative level; therefore, very few false positive cases can be expected at this statistical certainty level. Because each of the five collusion detection methods uses a different approach, if even one identifies a pair of students as potentially engaging in collusion, even at a “Low” statistical certainty, that examinee pair should be scrutinized.

By clicking on the examinee pair (e.g., “Examinee pair 1”), you can access a table that lists the item responses for each examinee. Green cells represent correct responses to items and red cells represent incorrect responses to items. The resemblance of the responses to items between the pair of examinees allows for an intuitive review of how similar, and therefore how rare, the pattern of responses for that pair of examinees is. An example of the examinee pair response table is presented below.

An example of a section of an examinee pair response table produced by Integrity.



Examinee pairs

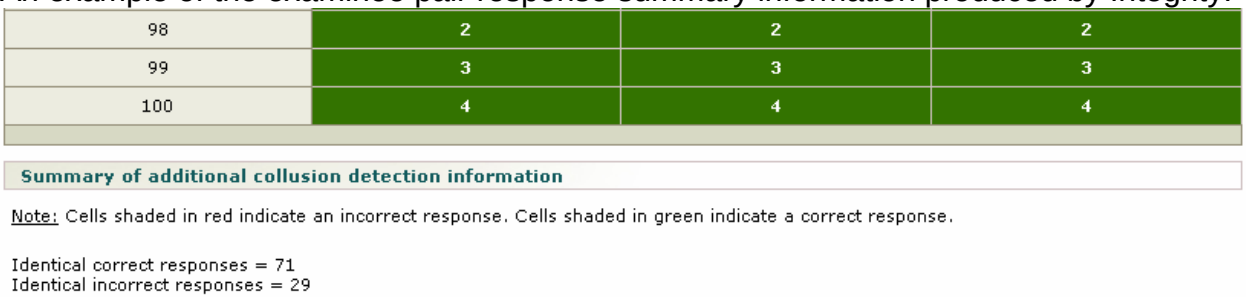
Return to collusion summary

Number of identical responses for examinee pair: 1

Item number	Examinee 10405047	Examinee 10405048	Correct answer
1	2	2	2
2	3	3	1
3	3	3	3
4	4	4	4
5	4	4	4
6	1	1	1
7	2	2	2
8	3	3	3
9	2	2	2
10	1	1	1
11	2	2	2
12	3	3	3
13	4	4	4
14	2	2	1
15	1	1	2
16	2	2	3
17	1	1	1
18	2	2	2
19	3	3	3
20	4	4	3

At the bottom of the examinee pair response table is a summary of the number of identical correct and incorrect responses for the pair of examinees, an example of which is presented below. In this example, we see that the examinee pair selected 71 identical correct responses and 29 identical incorrect responses. In other words, all of their responses to all of the 100 items composing the test were identical.

An example of the examinee pair response summary information produced by Integrity.



98	2	2	2
99	3	3	3
100	4	4	4

Summary of additional collusion detection information

Note: Cells shaded in red indicate an incorrect response. Cells shaded in green indicate a correct response.

Identical correct responses = 71
Identical incorrect responses = 29

Integrity provides very detailed but highly accessible reports to investigate cases of potential violations of academic integrity. The figure below is an example of a graph that

shows the rarity of responses for pairs of examinees. The red arrows represent the cases of examinees that had extremely similar, and therefore extremely rare, patterns of responses to items. Notice that the three potential cheating “events” are far to the right relative to all other pairs of examinees that took the test.

An example of a collusion detection graph produced by Integrity.