A theory of cheating

STEPHEN W. LINK and RICHARD B. DAY
McMaster University, Hamilton, Ontario, Canada

The theory of motivated cheating postulates that test takers may cheat when they do not know an answer. With probability k , an "observer" is unsure of an answer and will copy from a nearby "target" with probability c . The corresponding parameters for the target may be entirely unrelated to those of the observer. Thus, the undesirable feature of bidirectionality of parameters found in correlational techniques is not an inherent feature of this theory of cheating. Predictions are derived, and estimates of k and c are proposed. Statistically large values of c suggest that an observer was copying from a target. High values of c for both the observer and the target suggest collusion. The theory is applied to a 40-item five-choice test taken by students in an introductory psychology section. From the full paired comparison matrix of target $\times$ observer parameter estimates, the method identifies 2 students who were probably in collusion.

A university professor recently charged an undergraduate student with academic dishonesty. Although an undergraduate hearings committee (UHC) and the University Senate agreed that the student copied information from another student during an examination, the student made a final appeal to a legally constituted tribunal. At the appeal, the student was represented by legal counsel. The student's lawyer argued that the UHC found the student guilty because of the correspondence between the student's answers to multiple-choice questions and the answers provided by another student seated next to the appellant. The representatives of the UHC pointed out that the student was observed by an examination invigilator to be looking at the answer sheet of the other student. Furthermore, the correspondence between the answers of the two students was astonishingly similar. Even a crude statistic showed that the probability of the two students making by chance as many independent, but identical, erroneous answers as they did was vanishingly small.

The lawyer made three counterarguments. First, what was a crude statistic doing in a case where the student's future hung in the balance. A crude statistic is not sophisticated enough to meet the demands of a case held before a tribunal. Second, the students appeared to make the same errors because they studied together and misunderstood exactly the same material. Third, the correspondence between tests did not prove which student was cheating. If there was any cheating, it was certainly not by the lawyer's client.

As for the invigilator observing the client looking at the answer sheet of the other student, the amount of looking was apparently not sufficient for the invigilator to stop the offending student from completing the exam by picking up the exam at the moment of the offense as is re-

quired by university regulations. Therefore, the failure of the invigilator to act immediately was evidence of some uncertainty on the part of the invigilator. If the invigilator was not certain of cheating and was a witness to the student's actions, how could the UHC decide that the student was indeed cheating?

The tribunal granted the student's appeal. Professors stopped bringing charges of academic dishonesty before the UHC.

This story brings to light several disturbing issues lurking in the pit we call academic dishonesty. First, when students are represented by counsel but the faculty have no legal representation, the student has every advantage. No faculty member should be refused legal counsel when counsel is available to the student. Second, unless there is an actual measure of cheating, arguments about the correspondence between test answers hinge on the question of whether a correlation can be used as evidence of causation. Because the statistical measure of covariation between test performances is bidirectional, cause and effect cannot be established.

Determining a measure of cheating is a step forward in resolving one of these issues. In the case of multiple-choice examinations, the establishment of a measure of cheating is advanced by defining a relation between the performance of one student, the "observer," conditioned on the performance of another student, the "target." If the observer cheats, then the measure of cheating should satisfy the property that the amount of cheating for the target depends only upon measurement variability, while the amount of cheating for the observer deviates widely from that expected on the basis of no cheating at all. If there is collusion, the method should reveal similar measures for both the observer and the target because each is, at once, both observer and target. The conditional nature of the measure removes the bidirectionality of correlation and substitutes methods of statistical decision making that are meant to establish beyond a reasonable doubt the occurrence of cheating.

Correspondence should be addressed to S. W. Link or R. B. Day, Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1, Canada.

The Theory of Motivated Cheating

Figure 1 shows the key elements of the theory applied to an m -item multiple-choice test with a single correct response for each test item. The theory postulates that the observer cheats when unsure of an answer. The target of the observer's wandering eyes can provide either a correct or an erroneous response. When the target makes a correct response, the observer will produce exactly the same correct response. If the target makes an error, the observer will make exactly the same error. The probability of the observer's not knowing an answer is $(1-k_o)$ and the probability of cheating is c_o . Therefore, the probability that the observer copies the same answer as the target equals $(1-k_o)c_o$.

Yet there are other ways in which identical answers may occur. Correct responses may occur because the target and the observer both know the answer. The same error can occur when the target does not know an answer and

neither does the observer, but the observer guesses the same erroneous response as the target. Consequently, the mere existence of a correspondence between answers of the observer and the target is not sufficient to sustain a charge of cheating.

There are also cases when one can be certain that the observer did not copy from the target. For example, whenever the target provides a correct response and the observer makes an error, or vice versa, the error must occur due to some cause other than copying the correct response from the target. Also, when both the observer and the target make different errors, there is evidence that the observer did not copy from the target. Given that the performances may contain corresponding responses not due to cheating, how can the joint responses be used to determine a measure of cheating?

The matrix of joint responses by the target and the observer is shown in Table 1. Within the matrix, the first

Figure 1. The theory of motivated cheating assumes that when a student does not know an answer, the student may answer either by guessing or by copying from a nearby student.

Table 1
Conditional Probabilities of Various Response Types

Target	Observer	
	Correct	Errors
Correct	$k_O + (1 - k_O)[c_O + (1 - c_O)(1/m)]$	$(1 - k_O)(1 - c_O)(m - 1)/m$
Error	$k_O + (1 - k_O)(1 - c_O)(1/m)$	$(1 - k_O)(1 - c_O)(m - 2)/m$ $(1 - k_O)[c_O + (1 - c_O)(1/m)]$
		Different Same

row contains the conditional probabilities of either a correct or an erroneous response by the observer given a correct response by the target. The second row gives the conditional probabilities of making a correct response, a different error, or an identical error when the target makes an error. These conditional probabilities provide a basis for estimating the unknown probability of the observer's knowing an answer and the probability of cheating.

Estimation and Interpretation of Parameters

The evaluation of the best estimates of these parameters we leave for another occasion. However, for the purpose of discussion, note that an estimate of k may be obtained through a method similar to correcting multiple-choice tests for guessing (Link, 1982). That is,

$$k_O = \frac{\Pr\{\text{observer correct} | \text{target correct}\} - \Pr\{\text{observer makes same error} | \text{target error}\}}{1 - \Pr\{\text{observer makes same error} | \text{target error}\}}, \quad (1)$$

subject to the constraint that $0 < k_O < 1$.

The theory postulates that, given the opportunity to cheat, the observer cheats when unsure of an answer. A suitable measure of cheating is the estimate of the probability of cheating. Several methods yield estimates of varying quality for this unknown parameter. A fairly direct algebraic computation shows that

$$c_O = \frac{\Pr\{\text{correct} | \text{correct}\} - \Pr\{\text{correct} | \text{error}\}}{1 - \Pr\{\text{correct} | \text{correct}\} + \Pr\{\text{same error} | \text{error}\}}, \quad (2)$$

again subject to the constraint that $0 < c_O < 1$. Or, as an alternative to estimating c_O ,

$$1 - c_O = \frac{m[\Pr\{\text{same error} | \text{error}\} - \Pr\{\text{correct} | \text{correct}\} + \Pr\{\text{correct} | \text{error}\}]}{m - 1}. \quad (3)$$

There are three degrees of freedom in the joint response probability matrix and, therefore, several possible approaches to the estimation of two parameters. Even so, in the most interesting extreme cases, there are often no observations of a particular kind and, thus, the number of degrees of freedom is limited.

Beyond the estimation of the two unknown parameters, there is also a test of the theory that follows from the observation that

$$\frac{\Pr\{\text{different error by the observer} | \text{error by the target}\}}{\Pr\{\text{error by the observer} | \text{correct by the target}\}} = \frac{m - 2}{m - 1}. \quad (4)$$

When the numerator of Equation 4 is plotted against the denominator, the result should be a linear function with zero intercept and slope equal to a predetermined value, $(m - 2)/(m - 1)$, known in advance because the number of possible response choices, m , is known in advance.

When the observer cheats, the probability of an error must depend upon the performance of the target. To see how the parameters of the target and observer determine the (marginal) probability of an error for the observer, remember that under the usual correction-for-guessing model (Link, 1982), the observer knows the answer with probability k_O and will make a correct response with probability 1.0. Otherwise, with probability $(1 - k_O)$, the observer does not know the answer and either will not cheat with probability $(1 - c_O)$ and will guess an erroneous answer with probability $(m - 1)/m$ or will cheat with probability c_O . If the observer cheats from the target, the target will have made an error with probability $(1 - k_T)(m - 1)/m$. Thus,

$$\begin{aligned} \Pr\{\text{error for observer}\} &= (1 - k_O)[(1 - c_O)(m - 1)/m + c_O(1 - k_T)(m - 1)/m] \\ &= (1 - k_O)(1 - c_O k_T)(m - 1)/m. \end{aligned} \quad (5)$$

When c_O equals zero, the observer does not cheat. When k_T equals zero, the target knows nothing and always guesses. Either case produces a probability of error identical to that for the usual model of correction for guessing. Another extreme case occurs when the observer always cheats when uncertain of an answer—that is, $c_O = 1$. Then, the probability of an error for the observer depends directly on the probability k_T that the target knows the correct answer. If the target has perfect knowledge, the observer can reduce the probability of an error to zero by cheating. Presumably, it is this consideration that leads to cheating in the first place.

What can be done with these parameter estimates? First, they are not bidirectional as is a correlation coefficient. Instead, the parameters depend upon who is the target and who is the observer. One expects that, in a case of moti-

vated cheating, the observer will exhibit a relatively low value of k and a relatively high value of c . The meaning of *relative* must be defined with respect to the norms of a particular class. If estimates of a student's values of c and k deviate substantially from the norms for the class, then further investigation of where the observer and target were seated is warranted.

An Example

Imagine the following actual scenario. A class of some 1,500 introductory psychology students takes a 40-item multiple-choice test. For each item, there are five alternative choices but only a single correct alternative. To determine extreme cases, all students may be compared against each other in a full paired comparison design. In this way, each student is taken to be an observer of each of the other students, and each student is taken as a target of all other students. All pairs of targets and observers contribute a five-celled joint response-frequency matrix from which parameter estimates are derived.

Table 2
Raw Data From 2 Students Who Took
Identical Multiple-Choice Exams

Questions	Students		Master Key
	B	A	
1	5	5	5
2	2	2	2
3	1	1	1
4	5	5	5
5	2	2	2
6	1	1	1
7	1	1	1
8	4	4	4
9	1	1	1
10	3	3	3
11	2	2	2
12	2	2	2
13	5	4	5
14	3	3	1
15	4	4	4
16	2	2	1
17	5	5	3
18	1	1	1
19	5	5	4
20	4	4	4
21	4	4	4
22	5	5	2
23	1	1	3
24	3	3	3
25	2	2	2
26	3	3	3
27	4	4	4
28	3	3	4
29	5	5	5
30	1	1	1
31	1	1	1
32	2	2	2
33	4	4	4
34	2	2	5
35	3	3	3
36	4	4	4
37	3	3	3
38	5	5	4
39	2	2	3
40	2	2	4

Table 3
The Joint Response-Frequency Matrix When Student A
is the Observer of Student B

Target (Student B)	Observer (Student A)		Total
	Correct	Error	
Correct	28	1	29
Error	0	11	11
		Different	Same

Table 4
The Joint Response-Frequency Matrix When Student B
is the Observer of Student A

Target (Student A)	Observer (Student B)		Total
	Correct	Error	
Correct	28	0	28
Error	1	11	12
		Different	Same

Of course, we are mostly interested in the most extreme cases. To illustrate the point, we analyzed the test scores of students enrolled in an introductory psychology course. Although the course has 1,500 students, requiring the computation of some 2,248,500 joint response-frequency matrices, the task was greatly simplified by calculating the joint response-frequency matrices within representative classes of students. Of the various pairings of students within one class, one pair produced the data reported in Table 2, where the question numbers, answers from 2 students, and the correct answers are recorded.

The joint response-frequency matrix shown in Table 3 treats Student A as the observer and Student B as the target. Of the 40 responses, 28 are jointly correct. The target has a slight advantage in overall performance by showing a total of 29 correct responses versus the 28 correct for the observer (see Table 4). On one item, when the target responded correctly, the observer responded in error. Otherwise, the target and observer made 11 identical errors!

There are no correct responses by Student A when Student B made an error. Furthermore, all joint errors are identical. From the response probability equations and the parameter estimates available from Equations 1, 2, and 3, we see that the admissible value of k_O derived from these extreme data is $k_O = 0$. Given this value of k , the only admissible value for c_O is 1.0. The subject knows nothing at all and cheats with probability 1.!! Examining the data when Student B is the observer and Student A is the target also produces estimates of $k_T = 0$ and $c_T = 1$. Thus, for these 2 students, it appears that their performance is virtually identical, and neither knew much about the test material. The fact that all errors are identical suggests that the students copied from each other—that is, they were in collusion. The suggestion of collusion would gain added support if the times at which the responses were made were also available. However, no data on the timing of responses were available.

FRONT OF ROOM

INVIGILATOR'S DESK

AISLE

STUDENT 1			STUDENT 2		STUDENT 3		STUDENT 4	
STUDENT 5		STUDENT 6		A I S L E	STUDENT 7		STUDENT 8	
			STUDENT 9		STUDENT 10		STUDENT 11	
STUDENT 12		STUDENT 13	STUDENT 14		STUDENT 15			STUDENT 16
	STUDENT 17		STUDENT 18	A I S L E	STUDENT 19		STUDENT 20	
	STUDENT 21		STUDENT 22		STUDENT 23		STUDENT 24	
	STUDENT 25		STUDENT 26		STUDENT 27			STUDENT 28
STUDENT 29	STUDENT 30		STUDENT 31		STUDENT 32	STUDENT 33		STUDENT 34

BACK OF ROOM

Figure 2. The seating assignments for 34 introductory psychology students taking a 40-item multiple-choice test. Two students are conjectured to have copied from each other.

This conjecture gathers support when we analyze the seating arrangement for the students taking this test. The seating arrangement is shown in Figure 2. For this exam, the instructor was seated at the front of the examination room. Students A and B, who are numbered as 32 and 33, were seated next to each other at the very back of the room.

The analysis presented here came too late to be of use to the instructor seated at the front of the testing room. If the testing procedure were computer controlled so that the indicators of cheating were determined during the time of the test, a warning signal identifying the two students could be flashed to the instructor. Furthermore, the knowledge that an on-line program was monitoring the occur-

rence of cheating might do much to reduce the incidence of cheating.

Summary

Certainly this theory provides only a first step toward a rigorous analysis of cheating. Although some limitations exist in the application of these methods when there are no errors, the general scheme of analysis suggests a new approach to the analysis of interactions between test-taking students. There is a serious question surrounding the meaning of the estimation of the joint probability matrix when the time ordering of the responses is not known. To some extent, the loss of time-ordered response infor-

mation must reduce the value of the probability estimates. In spite of these reservations, the theory of motivated cheating shown in Figure 1 is a step forward. It provides conditional probability statements that lead to testable consequences (as in Equation 4), provides methods for estimating the probability that cheating occurs, and shows how the performances of the observer and the unsuspecting (perhaps) target are interrelated.

REFERENCES

- LINK, S. W. (1982). Correcting response measures for guessing and partial information. *Psychological Bulletin*, **92**, 469-486.