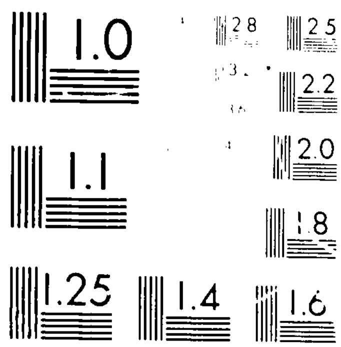
1.0
2.8
2.5
2.2
1.1
2.0
1.8
1.25
1.4
1.6

DOCUMENT RESUME

ED 169 101 TM 008 510

AUTHOR Anderson, Lorin W.; And Others
TITLE Using Errors to Improve the Quality of Instructional Programs.
PUB DATE [Mar 78]
NOTE 18p.; Revision of a paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Algorithms; Calculation; Elementary Education; *Elementary School Mathematics; Elementary School Students; *Formative Evaluation; Mastery Learning; *Mathematical Logic; *Mathematics Instruction; Program Evaluation; Student Evaluation; *Test Interpretation; *Test Results

ABSTRACT

Clinchy and Rosenthal's error classification scheme was applied to test results to determine its ability to differentiate the effectiveness of instruction in two elementary schools. Mathematics retention tests matching the instructional objectives of both schools were constructed to measure the understanding of arithmetic concepts and the ability to perform computations and algorithmic operations. Inter-school comparisons were made with respect to the types of errors made: computational, algorithmic, and omission. Fifth grade students in one school made significantly fewer computational errors than in the other school. In the remaining grades, there were no significant differences. Additional practice exercises were recommended to correct this error. Other students made significantly more algorithmic errors and a four-step approach to teaching algorithms was outlined to alleviate this weakness. Significant inter-school differences in omissions errors were noted in grades 5 and 6--the students concerned would also benefit by a more systematic approach to algorithm teaching. In sum, the error classification scheme can be used to evaluate instructional programs and to suggest instructional improvements. (CP)

Using Errors to Improve the Quality of Instructional Programs

Lorin W. Anderson

University of South Carolina

Corinne C. Scott

National Institute of Education

Amelia Reynolds

University of South Carolina

Abstract

The study examined the usefulness of Clinchy and Rosenthal's error classification scheme for the purpose of using test results to improve instructional effectiveness. Students in two instructional programs were compared with respect to the types of errors made on a test of mathematics retention. Three types of errors were identified: computational, algorithmic, and omissions. Differences in the types of errors made by the students in the two programs were found. In addition, prescriptions for improving the programs based on the nature of the errors can be offered.

Using Errors to Improve the Quality of Instructional Programs

Evaluation results are used to determine whether to maintain, terminate, or modify instructional programs. The decision to maintain or terminate an instructional program is typically made on the basis of the relationship between the actual performance of a group of students and the "ideal" performance of students as specified by the program. If the actual performance of students surpasses the "ideal" performance standard, the decision is made to maintain the program as is. On the other hand, if the actual performance of students is far below the standard specified by the program, the decision is made to terminate the program.

The problem occurs when the discrepancy between the actual performance of students and the level of performance specified by the program is small, but noticeable. Typically, the decision in this situation is to modify the program in an effort to improve its effectiveness. While a decision to modify implies that the program is salvageable, specific ways in which the program can be improved are not immediately evident.

Perhaps one reason for the lack of suggestions for improvement of programs comes from the manner in which the initial decisions are made. Since the decisions to maintain, terminate or modify programs are made based on the size of the discrepancy between actual performance and "ideal" performance, the most appropriate measure of actual performance is the number of correct items. While the number of correct items can be used to judge the extent to which students attain or surpass a particular standard, it is of little use in determining ways of improving instructional programs. Different information seems necessary in order to accomplish this latter task.

One such source of information is the nature of errors made by the students in the program. Since errors typically indicate misunderstandings that need to be corrected, the type of error made may provide a clue to alterations of instructional programs that may lead to some kind of improvement. What appears necessary, then, for errors to be useful in this regard is some typology of errors. Once errors have been classified according to a typology, prescriptions for the correction of these types of errors can be incorporated into subsequent revisions of the programs.

Clinchy and Rosenthal have proposed a classification scheme for learner errors.[1] According to Clinchy and Rosenthal, three categories of errors exist: intake errors, organizational errors and executive errors. These three types of errors can be defined briefly as follows. Intake errors refer to "errors in the (learner's) conception of the goals and the data of the task."[2] Organizational errors refer to errors made in the "selection of a method for manipulating the available information in order to reach a solution."[3] Finally, executive errors refer to errors in the "actual performance of the operations demanded by the program."[4]

The purpose of the present study was to examine the feasibility of using Clinchy and Rosenthal's error classification scheme for the improvement of instructional programs. More specifically, the study was designed to answer two questions.

1. Is the error classification scheme useful in differentiating the effectiveness of instructional programs?

2. If the error classification scheme is useful in the above context, what prescriptions for program alteration can be made based on the types of errors made by learners in the programs?

Method

Sample

Two elementary schools, employing two different instructional programs, were used in the study. Both schools were located in the same school district and both schools contained students in grades one through six.

The first school (School A) contained 135 boys and 141 girls. Eighteen per cent of the students received free or reduced lunches. Approximately fifteen per cent of the parents were classified as being in professional occupations. Virtually all of the students were Caucasian.

The second school (School B) contained 239 boys and 220 girls. Thirteen per cent of the students received free or reduced lunches. Approximately twenty-two per cent of the parents were classified in the professional occupational category. Again, virtually all students were Caucasian.

Since two different schools were used in the study, an investigation of the initial cognitive differences of the students was conducted. Raw scores on the Metropolitan Readiness Test (MRT) administered at the beginning of first grade were available for students in grades two through five. Stainine scores on the Otis-Lennon Mental Abilities Test (OLMAT) administered in grade four were obtained for the grade six students.[5]

The mean scores of the students in the two schools at each grade level were computed and the differences compared using a one-way analysis of variance. Statistically significant differences favoring School B were found at grades three and six. Differences at grades two, four, and five also favored School B but were not statistically significant.

Based on these findings, a decision was made to form matched pairs of students at each grade level, the matching being made on the student scores on the MRT or the OLMAT.[6] The matching process produced 22 pairs of students in the second grade, 21 pairs in the third grade, 22 pairs in the fourth grade, 22 pairs in the fifth grade, and 23 pairs in the sixth grade. The means and standard deviations of the two groups of students at each grade on either the MRT or OLMAT were virtually identical. These matched pairs formed the sample used in the study.

Instructional Programs

School A teachers implemented an elementary mathematics program based on Block's description of group-based mastery learning.[7] The School A teachers identified instructional objectives, designed original and supplementary instructional activities and materials for each objective, constructed formative tests, and designed a system for recording and reporting individual pupil progress with respect to the objectives. The teachers began each instructional unit by informing the students of the objectives to be learned. The teachers then presented the original instructional materials and activities relevant to the particular subset of objectives to the entire class of students. Through the use of informal in-class questioning and worksheets as well as more formal tests, student progress *vis à vis* the objectives was periodically checked and recorded. Errors, once identified, were remediated through the use of large-group and small-group re-teaching (or corrective) sessions. After every student in the class had demonstrated a preset level of "mastery" on either the unit test or a parallel form of the unit test, the entire class moved to the next instructional unit.

School B teachers employed an instructional program believed to be representative of the conventional instructional programs in elementary schools. More specifically, the instructional program in School B was lacking in explicity stated objectives, formative tests, re-teaching or corrective activities and materials, and a whole-class mastery performance standard. The study lasted one academic year.

Instrumentation

Tests appropriate to each grade level were constructed for the purpose of assessing the effectiveness of the two instructional programs. Since program effectiveness was defined in terms of student retention (rather than immediate achievement) the tests were designed to be administered early in September. The majority of the items on the test were constructed to represent the instructional objectives of the instructional program operating in School A. Additional items were written to parallel items on the California Achievement Test. Since the majority of the items were based on the objectives of School A's instructional program, one might suspect that the test might be differentially valid for the two instructional programs. In order to avoid this problem, a visit was made to School B where the inital set of objectives were discussed with the teachers at each grade level. Objectives that were not taught by teachers at both schools were identified and eliminated from further consideration. Teachers also examined the items and indicated those items they did not believe measured student attainment of the stated objectives. Such items although few in number, were eliminated from the tests.

The items were written to test students understanding of arithmetic concepts, ability to perform arithmetic computations, and ability to perform algorithmic operations. All items were written in the short-answer completion format. The internal consistency reliabilities of the five tests ranged from 0.80 to 0.92. The median internal consistency coefficient was 0.87. In general, then the tests seemed to possess an adequate degree of content validity and internal consistency reliability.

Scoring

Following administration of the tests, the student responses were categorized according to the type of error made, one type of error for each correct answer.

Three categories of errors were established based on Clinchy and Rosenthal's error classification scheme.

Computational errors most closely resembled Clinchy and Rosenthal's category of executive errors. Computational errors refer to those errors which indicated that the student had performed the correct arithmetic operation inaccurately. That is, the student used the appropriate algorithm in solving the problem but made a mistake on one or more of the steps. An example of a computational error would be the following. A student sees the problem 14 + 17, adds, but arrives at the answer 41.

What may be termed "algorithmic errors" most closely resembled Clinchy and Rosenthal's category of organizational errors. Algorithmic errors refer to those errors which indicated that the student made use of an incorrect operation on algorithm. For example, the student may have added when he should have subtracted. Or, a student may have failed to "invert and multiply" when confronted with a problem dealing with division of fractions.

Errors of omission, or omissions, were thought to most closely resemble Clinchy and Rosenthal's category of input errors. The reasoning behind this equation of error types was that students who failed to conceptualize the "goals and data of the task" would be likely not to attempt the task at all. Omissions, therefore, included those problems the student made no apparent attempt to answer.

In sum, then, each student response to each item was classified as to the type of error made: computational, algorithmic or omission. Where no mistake was made, of course, no error category was indicated.

Results

The first question was concerned with the utility of the error classification system in differentiating the two instructional programs. Tables 1 through 3 display the means, standard deviations, and correlated t-tests for the three categories of errors made by students in the two programs.

Insert Tables 1, 2, and 3 About Here

An examination of the differences in computational errors ma e by the students in the two programs (see Table 1) indicates that School B students made significantly fewer computational errors than School A students in grade five. In the remaining grades, no significant differences in the number of computational errors were found to exist.

Quite a different picture is presented with respect to algorithmic errors (see Table 2). With the exception of students in grade four, students in School A made significantly fewer algorithmic errors than students in School B. In grade four the difference was not significant.

Finally, Table 3 presents the results with respect to student omissions. In the first three grades, no significant differences existed in the number of items omitted by students in the two instructional programs. In grades five and six, students in School B omitted significantly fewer items than students in School A.

In general, then, the error classification scheme appears to be a useful way of distinguishing the instructional effectiveness of the two programs. The major difference between the two programs is in the area of algorithmic errors with the difference favoring instructional program A. Few differences between the programs exist in the areas of executive errors and omissions. What differences do exist, however, tend to favor instructional program B.

Discussion

The second question underlying the present study was concerned with the extent to which differences in types of errors could be used to improve instructional programs. The search for an answer to this question began with an examination of what each type of error might mean about the nature of students' learning. The following speculations were offered.

Computational errors generally indicate that while students know how to think about the problem, they either are careless in performing the algorithms or are unaware of one or more of the mathematics facts necessary to perform the algorithm. On the other hand, algorithmic errors suggest the students, for some reason or reasons, selected the wrong algorithm to use in attempting to answer the question or solve the problem. Finally, omission errors typically indicate that the students do not know how to think about the problem. In other words, they do not even know where to begin.

Based on these possible meanings, different alterations can be recommended for the different instructional programs yielding different types of errors. The major instructional weakness in the School B program was in the area of algorithmic errors. Two suggestions can be offered for such programs.

First, a more systematic approach to the teaching of algorithms is recommended. This approach would include 1) an explicit display or presentation of the steps involved in the algorithm; 2) a discussion of each of the steps, emphasizing the learner's understanding of each step; 3) a demonstrated application of the algorithm to a particular problem, complete with "answers" that follow from the successful application of each step; and 4) practice exercises which allow the student to apply the algorithm in a variety of situations and which require the student to write the partial answers after performing each step. The emphasis on the understanding of each step and the performance of the algorithm in a variety of situations would help the learners to see the limitations of the algorithm in solving problems. This may help the students to avoid over-application of the algorithm.

Second, since algorithmic errors include errors in which the student adds instead of subtracts and divides instead of multiplies, it would be wise for the unit tests in such programs to be somewaht cumulative in nature. That is, unit tests could present addition and subtraction exercises, and multiplication and division exercises, in close proximity to each other. This will allow students practice in discriminating the algorithms that are necessary to order to solve the problems correctly.

In contrast to programs in which algorithimic errors are the major problem are those programs in which computational errors constitute the major problem. In the present study, this problem seems apparent only in the fifth grade program in School A. Whereas algorithmic errors seem to require alterations in the actual teaching and testing of the material or skill (as indicated above), computational errors seem to require alterations in the amount or type of practice provided to the learners. Thus, instructional programs in which executive errors occur frequently should provide the learners with additional practice exercises or different types of practice exercises. In this regard, students should check their answers to the exercises (or have their answers checked) frequently so that they do not overlearn incorrect computations.

Instructional programs plagued by omissions would seem to benefit from a more systematic approach to the teaching of algorithms. The emphasis in this teaching should be on the relationship between problem situations and appropriate algorithms rather than on the application of the algorithmic *per se*.

In conclusion, the use of evaluation as an aid in the improvement of instructional programs seems promising if, and only if, attention is focused on the types of errors made by the students on the tests. The error classification scheme suggested by Clinchy and Rosenthal seems extremely

useful in this regard. Applications of the error classification scheme can aid in the identification of differences in the effectiveness of instructional programs which can, in turn, be used to suggest specific changes which can be made in order to improve the programs.

Bibliography

Block, James H., ed. Mastery Learning: Theory and Practice. New York: Holt, Rinehart and Winston, 1971.

Clinchy, Blythe, and Rosenthal, Kristine. "Analysis of Children's Errors." In Psychology and Educational Practice, pp. 90 - 129. Edited by Gerald Lesser. Glenview, Illinois: Scott, Foresman and Company, 1971.

Footnotes

1 Blythe Clinchy and Kristine Rosenthal, "Analysis of Children's Errors," in Psychology and Educational Practice, ed. Gerald Lesser (Glenview, Illinois: Scott, Foresman, and Company, 1971), pp. 90 - 129.

2 Clinchy and Rosenthal, *op. cit.*, p. 94.

3 Clinchy and Rosenthal, *op. cit.*, p. 94.

4 Clinchy and Rosenthal, *op. cit.*, p. 95.

5 Since the dependent variable was retention the appropriate test was administered in September. Consequently, only data for students in grades two through six were obtained. The first grade students were just beginning the instructional programs.

6 Once again the authors were faced with the dilemma of comparing initially noncomparable groups or comparing matched groups of students. In view of the questions under investigation, the choice of matched groups appeared to be preferable.

7 James H. Block (ed.), Mastery Learning: Theory and Practice (New York: Holt, Rinehart and Winston, Inc., 1971).

Table 1

Means, Standard Deviations, Sample Sizes and Correlated t-tests for Number of Computational Errors Made by Students in the Two Programs in Grades 1 Through 5

Grade Level	n	School A Mean	School A S.D.	School B Mean	School B S.D.	Correlated t-test
2	22	3.1	2.4	3.3	4.1	0.0
3	21	7.6	3.7	6.4	3.5	1.2
4	22	5.4	3.3	4.6	2.4	0.7
5	22	6.8	2.8	4.8	2.8	5.9*
6	23	4.3	2.[illegible]	5.0	3.3	0.6

Note---The following convention for level of significance is used throughout this paper: .05 level = * and .01 level = **

Table 2

Means, Standard Deviations, Sample Sizes and Correlated t-tests for Number of Algorithmic Errors Made by Students in the Two Programs in Grades 1 Through 5

Grade Level	n	School A		School B		Correlated t-test
		Mean	S.D.	Mean	S.D.	
2	22	0.2	0.6	3.4	7.7	3.8*
3	21	0.0	0.0	3.0	6.1	5.1*
4	22	3.4	2.7	3.6	4.5	0.0
5	22	5.3	4.4	13.1	9.8	11.7*
6	23	10.3	8.3	17.0	11.3	5.2*

Table 3

Means, Standard Deviations, Sample Sizes and Correlated t-tests for Number of Omissions by Students in the Two Programs in Grades 1 Through 5

Grade Level	n	School A		School B		Correlated t-test
		Mean	S.D.	Mean	S.D.	
2	22	0.6	2.6	0.6	2.8	0.0
3	21	1.0	2.1	0.2	0.4	2.9
4	22	0.6	1.7	0.2	0.9	0.8
5	22	4.0	5.0	1.5	3.0	3.9*
6	23	12.3	10.8	6.3	7.3	4.9*