DOCUMENT RESUME

ED 269 425 TM 860 251

AUTHOR Ediger, Marlow
TITLE Appraising Student Achievement in Mathematics.
PUB DATE [86]
NOTE 18p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Check Lists; Criterion Referenced Tests; Educational Testing; Elementary School Mathematics; Elementary Secondary Education; *Informal Assessment; *Mathematics Achievement; *Mathematics Tests; Parent Teacher Conferences; Rating Scales; Secondary School Mathematics; Sociometric Techniques; Standardized Tests; Student Evaluation; *Teacher Made Tests; Test Construction; Testing Problems; Test Reliability

ABSTRACT

This paper briefly describes a number of techniques for evaluating students' achievement in mathematics: (1) checklists and rating scales; (2) standardized tests; (3) criterion referenced tests; (4) anecdotal records; (5) student products, including homework and school assignments; (6) teacher-written tests; (7) conferences with students; (8) parent-teacher conferences; (9) oral tests; and (10) sociometric devices. The advantages and limitations of these techniques are briefly discussed, and use of a variety of tests is suggested. Guidelines for the construction of teacher-written tests are also included. (GDC)

APPRAISING STUDENT ACHIEVEMENT IN MATHEMATICS

Marlow Ediger
Northeast Missouri State University

APPRAISING STUDENT ACHIEVEMENT IN MATHEMATICS

In an era of accountability for teachers, it is vital that quality procedures in evaluation of learner progress in the mathematics curriculum be in evidence. Which techniques of appraisal then need to be emphasized?

Using Checklists and Rating Scales

A checklist may be utilized if progress can be recorded on an either/or basis. For example, a student who correctly finds the area of a square, rectangle, or triangle can receive a checkmark on his/her checklist to indicate that the following objective has been attained: Given the dimensions of each side of a square, rectangle, or triangle, the pupil can compute the area correctly.

The name of the involved pupil needs to appear on the checklist. The date for achieving the objective needs to be recorded next to the objective on the checklist. Relevant sequential objectives only, should appear on the checklist.

The rating scale should be used when general objectives are utilized in ongoing units. It is difficult to state selected objectives in measurable terms. And yet, the following are indeed worthwhile in any unit of study in mathematics: being able to think critically, creatively, and solve problems. Thus, a student's proficiency in each of these higher cognitive level skills may be evaluated on a five point scale, such as doing excellent, very good, average, below average, and poor in thinking critically, creatively, and being able to solve problems.

No doubt, the checklist and its use appraises students objectively, independent of the evaluator, whereas the rating scale stresses subjectivity, in degrees. Interscorer reliability in utilizing the rating scale can be emphasized by having a colleague assist in developing the ratings on a five point scale. If the two raters of student achievement in critical thinking, creative thinking, and problem solving agree, the chances are reliability with its consistency of results is in evidence.

Using Standardized Tests

In using standardized tests to appraise student progress, teachers of mathematics in a school system need to notice validity correlation coefficients, as well as reliability values. Certainly, the test used to measure pupil achievement must appraise student progress in terms of what it purports to measure and in this situation achievement in mathematics. Validity in the measurement instrument is then inherent. Consistency of results is also wanted if a test, for example, were taken over again by the student. Otherwise, what worth would a test have if the same pupil taking it two times would rank on the 70th and 25th percentile respectively? Inconsistent results are then in evidence. Ranking on the 70th percentile is high since out of every 100 taking the test, thirty would be above and seventy below. Being on the 25th percentile is low. Thus, out of every 100 students taking the test, seventy five would be above and twenty-five below for the person taking the test and ranking on the 25th percentile.

Standardized achievement tests are norm referenced. When viewing the results from standardized achievement tests in mathematics, pupils scores are spread out from high to low. This is typical of norm referenced testing. Thus,

a student's results can be compared with others in a classroom or schools. The range in percentile ranks could be from 99th to the first percentile. Or, if grade equivalents are given for each students, the range could be from twelve to grade three, as examples, for a seventh grade student.

Norm referenced tests have been standardized on a set of students which should possess similiar traits and characterstitics, as the general population in society. Otherwise, the concept of external validity would be lacking. For example, if one's present classroom of students were all gifted by definition and the norm group on which the achievement test was standardized were heterogenous (mixed achievement levels), it would mean that the two sets would be highly incongruous. Results of the gifted students can be compared to the norm group. However, the chances are that the standardized test will not measure as accurately as one would wish the results to be for the gifted students. Why? There would be relatively few students in the norm group who were gifted, compared to average achievers. In the norm group, most students will have measured close to the mean or arithmetical average. One standard deviation (34% of the norm group) below the mean and one standard deviation above the mean (another 34% of the norm group) make for 68% of the students in the norm group. These students represent a range of percentiles from the 16th percentile to the 84th percentile. Thus, most of the students in the norm group have been taken into consideration with the $\pm$ one standard deviation above the mean. Gifted students in our classroom would tend to score above or well above the 84th percentile. Just sixteen percent of the students in the norm group scored higher than one standard deviation above the mean. Add another standard deviation above the 84th percentile (one standard deviation above the mean is the 84th percentile approximately) and a student is on the 98th percentile. Thus, two standard deviations above the mean and two standard deviations below the mean account for 96 percent of the norm group, leaving two percent three

standard deviations above the mean and two percent three standard deviations below the mean. Thus, one's own gifted class of students in comparison with the gifted on the norm group would find few in number of the later. Most students in the norm group would cluster more so toward the mean. Thus, an inadequate number of students three standard deviations above the mean are available to compare results with the teacher's own gifted class. Better comparisons can be made with average and talented students in the norm group if students in the teachers classroom are also of similar abilities. An adequate number of any category of pupils be they high, average, or slow learners need to be in any norm group so that students in classrooms who come in any one of these categories may have adequate numbers for making comparisons.

Criterion Referenced Tests (CRT)

Norm referenced test results provide for a range of achievement among students. Thus, a soread of scores are involved from high to low. On the other hand, CRT's may not provide for much of a spread of scores among students. CRT philosophy is much different than norm referenced generalizations. In CRT's, whether teacher, school, or commercially developed, measurably stated objectives are emphasized. The objectives exist prior to instruction of pupils. Either students attain or do not attain the chosen ends. No guesswork is involved in these testing situations.

The teacher can announce to students at the beginning of a class session which objective(s) a student is to attain. The mathematics teacher may then teach to the stated objective. The learning activity or activities are valid in that they harmonize directly with the objective. Finally, the instructor appraises to see if each student has or has not achieved the specific objective. The teacher then stresses a similar sequence, announce the measurable

objective to learners prior to instruction, teach so that students are on course to directly attain the precise end, and measure to notice which pupils have been successful in goal attainment. If a student did not achieve a measurable objective, additional learning activities need to be provided so that the involved learner can be successful in attaining the chosen end(s).

Time is the variable in CRT. Thus, slow learners will need more time as compared to average and gifted students to achieve worthwhile objectives. A computer printout of sequential objectives may be sent home with the pupil to parents. Along with each objective on the printout are the textbook, workbook, and/or worksheet pages so that the parent might know exactly which learning activities assist students to attain an end. On the elementary school level, in most cases, it is possible in homework for parents to help their offspring achieve objectives.

Thus, slow, average, and fast learners can achieve the same sequential objectives. More time and assistance must be given to the slower achiever to attain sequential ends.

From CRT results, there might well be a spread of scores in achievement from high to low. However, that is not the point. The goal in CRT is to assist each student to achieve as many specific objectives as is reasonably possible. Absolute standards are then emphasized. Either a student has or has not achieved a sequential objective.

Anecdotal Records

Teachers of mathematics need to take time to record representative behavior of each student. Unless behaviors are recorded, they can be forgotten by the teacher. Representative behavior of each student needs to be recorded.

Biased statements written for any pupil should be omitted. Loaded words also should be left out of written anecdotal statements. If a teacher records observed behavior for two students each school day, it does not take long before the rounds have been made one time. For example, with twenty four students in a classroom, it should take the teacher twelve school days to complete writing the acecdotal statements for involved learners. The teacher needs to continue writing the anecdotal statements throughout the school year to notice patterns of behavior for each student.

Which statement(s) might be written as an example for a student?

September 7, Lois completed her assignments with no errors in the completed work.

This anecdotal statement is factual and verifiable. No loaded terms were used. The statement indicated the kind of behavior exhibited by Lois.

Another example of an anecdotal statement would be the following:

September 8, Albert looked around the room for five minutes before starting on his mathematics assignment. He missed fifteen out of the thirty addition computations.

Again, a factual statement of Albert's achievement has been recorded. Other observers should be able to verify or refute Albert's observed behavior.

Student Products

Homework and schoolwork assignments provide an adequate supply of student products in terms of completed work. Pupils with teacher guidance may diagnose and remediate errors made by the former. From learner products in mathematics, the teacher may notice the kinds of mistakes students make in ongoing units.

The errors may be due to:

1. human factors in that perfection does not reside within the individual.
2. carelessness on the part of the student.
3. computational errors.
4. not understanding an operation or process in mathematics.
5. a lack of readiness within the pupil.
6. not perceiving purpose or reasons for learning.
7. poor sequence in instruction.
8. not perceiving interest in learning.

By diagnosing student progress in mathematics, the teacher can make appropriate judgments in terms of which objective should come next in sequence. Feedback from student products provides teachers with needed information on which learning activities in mathematics need to be provided to learners. Success in learning is basic for students to achieve optimally in the mathematics curriculum.

TEACHER WRITTEN TESTS

Periodically, the teacher will wish to appraise student progress through testing. Test results from learners can provide excellent information to the teacher as to which goals need to be emphasized within diverse units of study.

One kind of teacher written test item appropriate for measuring mathematics achievement is the multiple choice item. The following is an example:

The formula for finding the area of a circle is

(a) $r^2\pi$ (b) $\frac{1}{2}bh$ (c) lw (d) s^2.

Criteria to emphasize in writing multiple choice items include the following:

1. content needs to be clearly written so that either a, b, c, or d is the correct response. Sometimes more than a single response is correct in a

multiple choice item. In the directions for taking the test, clarity is important in the printed content. Students should know precisely how to take the test as a result of having read the directions.

2. the distractors should be plausable. Too frequently, teachers have written ridiculous distractors whereby students need minimal knowledge to eliminate the bizarre, such as

The formula for finding the area of a circle is
(a) 13 + 7.

3. no clues should be given as to which response is correct or incorrect. The following violates this standard:

The Pythagorean Theory, in finding the diagonal of a right triangle was developed by
(a) Pythagoras (b) Plato (c) Aristotle

4. unnecessary words should be eliminated in any test item. The following is an example of useless wording:

The formula $r^2\pi$ is used
(a) to determine the area of a square.
(b) to determine the area of a circle.
(c) to determine the area of a triangle.
(d) to determine the area of a rectangle.

To eliminate excess words, the multiple choice item should be rewritten in the following way:

The formula $r^2\pi$ is used to determine the area of a
(a) square (b) circle (c) triangle (d) rectangle.

5. the determiners "a" and "an" need to be carefully used in written multiple choice items. For example, in the following test item the article "an" provides the clue as to which is the correct answer:

The formula "a=lw" is not applicable in finding the area of an
(a) square (b) rectangle (c) ellipse

The onl· correct answer would be an ellipse.

True-false items can be used to evaluate pupil achievement in mathematics, such as in the following clearly written test item:

The formula for finding the circumference of a circle is $d\pi$.

The answer to the above true-false item is clearly true. If a true-false item is false, the student could be asked to correct the part that is false. For example, supposing the following item is on the test:

The formula for finding the area of a square is $bh=2$. The underlined part if false may then be corrected by students so that it reads s^2.

A matching test can be developed by the teacher to measure factual learnings acquired by students. It is significant to follow the following criteria when developing matching tests:

1. have more items in one column compared to the second column to match. Thus, the process of elimination cannot be used extensively to complete the matching test.

2. have phrases or single concepts in one column to match with the second column. Column two may also contain phrases or single concepts. However, if both columns contain lengthy sentences, the matching test may be complex indeed. One column may have sentences of reasonable length.

3. develop a test of moderate length so that fatigue does not set in on the part of the student taking the test. The mathematics teacher is attempting to measure conceptual learning and not necessarily endurance in taking the matching test.

4. use a single topic when developing test items. If items pertain to diverse topics, it may be relatively easy for a pupil to match selected items in column A with Column B. Thus, a matching test may deal with the topic of formulas to determine areas of diverse geometrical figures. A single topic is then utilized in developing the matching test. If a simple addition problem

were added to the test involving the above named formulas, it would be relatively easy for the learner to notice that only one possible matching could be made in relating column A with column B of the test items, e.g. a numeral plus a numeral equals a numeral.

Short answer or completion items may be utilized in testing. The following is an example of a short answer test item:

A+B=B+A emphasizes the _____ property of addition. There are relevant criteria to follow in writing short answer test items.

1. Adequate information must be provided in the short answer test item so that students know what is wanted in terms of responses. The following short answer item lacks needed subject matter in terms of responses wanted:

_____ and _____ are the _____ of _____.

2. The blanks in short answer test items should be of equal length so that clues are not given to the test taker as to which the correst answer is.

3. The blank spaces should be numbered sequentially to make for ease of scoring.

4. It is important to write subject matter clearly so that the involved test taker interprets the short answer test items accurately.

5. Learners need to have an adequately developed writing vocabulary to respond correctly to the blanks in short answer test items.

6. The teacher needs to give credit to correct responses even though they differ from the right answer written on the teachers' own originally developed key.

Essay items as a fifth type of teacher written test item may be utilized to appraise student progress. Thus, an essay item might be written to have students clarify and explain thinking involved in solving a word problem. The following is an example:

Mr. and Mrs. Brown and their two children, ages seven and ten, took a vacation trip in which 1,500 kilometers were traveled. The car averaged ten kilometers traveled for each liter of gasoline used on the excursion. The price of the gasoline was 26¢ per liter. The total cost for lodging was $280 for the vacation trip which lasted seven days. The average cost per day for meals for each of the four family members was $12. How much did the total vacation trip cost? In your answer, give reasons for using each numeral and/or number name used.

Here students need to analyze their thinking in terms of how each numeral and number name is to be utilized. In the analyzat , selected values will not be used in solving the problem. The cognitive level of analysis may then be appraised by the teacher.

The mathematics teacher may also appraise if pupils can apply that which had been learned previously. Thus, are students able to utilize the cognitive level of application within the framework of problem solving? If a pupil cannot use what has been acquired previously, perhaps meaning and purpose were omitted on the learner's part in learning experiences prior to the essay test item that needs its required responses.

In any essay response, the teacher can notice the quality of spelling, handwriting, punctuation, usage, capitalization, and sequence of ideas. The mechanics of writing described above, should be appraised separately from content or subject matter needed to solve the problem.

In writing essay tests to appraise student progress in mathematics, the teacher must

1. write adequately delimited test items. The following essay item is too broad: Discuss mathematics.

Volumes have been written and will continue to be written on mathematics.

The following essay item is adequately delimited:

Describe the meaning of each symbol in the formula $A=\frac{1}{2}bh$.

2. write essay items which are not too factual in terms of needed student responses, such as:

"What is the formula for finding the area of a parallelogram?"

A quality true-false, multiple choice, matching, or completion (short answer) test item can deal in a more effective way with factual content, as compared with the essay test. Thus, instead of the essay item

"What is the formula for finding the area of a parallelogram?"

the following completion item would suffice:

"The formula for finding the area of a parallelogram is ____."

Conferences with Students

The mathematics teacher can assist student achievement with the use of conferences. The teacher may meet with one or more (a small group) students to discuss common errors made in ongoing learning activities. Diagnosis is an important concept to emphasize in the conference setting. Which specific errors did one or more learners make? The following are examples of common mistakes:

1. not copying a problem correctly from a textbook or workbook.

2. a lack of understanding of the terms carrying & borrowing. Another name for the same kinds of errors would be regrouping & renaming.

3. inability to recall answers to basic addition, subtraction, multiplication, and division number pairs.

4. not being able to apply formulas in mathematics to concrete situations in life.

5. a lack of proficiency in reading content, such as story problems in mathematics. If a student cannot identify, approximately, ninety percent of the words in word problems, comprehension will tend to go downhill.

6. inability to analyze in terms of needed content as compared to the unneeded in solving word problems.

7. lacking the desire to evaluate personal achievement in ongoing lessons and units in mathematics.

Thus, the conference method in an atmosphere of respect can assist individual students to achieve at a more optimal rate of progress.

Parent - Teacher Conferences

An adequate number of parent-teacher (PT) conferences should be held in any given school year. Mathematics teachers need to inform parents of their son or daughter's progress in a face to face situation. Too frequently, report cards alone are utilized to report pupil progress to parents. However, report card results are a one way street of communication. Parents, no doubt, have questions pertaining to items on a report card. Sometimes PT conferences are held right after the report card has been issued the first time to a pupil in a given school year.

To achieve readiness in having a PT conference, the teacher needs to:

1. have work samples pertaining to completed daily assignments of the involved student. The parents need to see the quality of work done by their offspring. Seeing directly products of a student is better than merely attempting to describe how well a pupil is achieving in mathematics.

2. be knowledgeable about the capacity and general achievement level of the student in the area of mathematics.

3. understand attitudes possessed by the learner toward mathematics.

In a PT conference, the teacher must:

1. accept parents as human beings having a sincere desire in wanting their son or daughter to achieve well in the mathematics curriculum.

2. respect the thinking of parents. Nothing is gained by exhibiting feelings of hostility and mistrust.

3. work together with the involved parents in guiding each pupil to achieve optimally in mathematics.

Certainly, PT conferences are vital in improving the mathematics curriculum!

Oral Tests

For blind or partially sighted students, oral tests may be utilized to appraise progress. The normal learner, in selected instances, may also benefit from oral testing. The use of oral tests can eliminate the reading

factor in measuring student achievement. Sometimes, a mathematics test appraises reading skills, such as in word or story problems. However, teachers should attempt to ascertain students' progress in computation, concepts, and problem solving. Oral tests can be valid and reliable to measure achievement in mathematics.

Content validity is involved in selected oral items to utilize in appraising student progress in mathematics. Each item selected should relate directly to the precise, measurable objectives emphasized in teaching and learning. Thus, the effectiveness of the teacher's proficiency in the teaching of mathematics is being appraised. Did the involved student attain the precise ends? It is an either or situation. Either the pupil was or was not successful in goal attainment. If a learner has not been successful in achieving an objective. diagnosis of the situation is necessary to determine causes and remediation methods.

Test items, orally administered, that relate directly to the statement of precise objectives should be valid. Thus, learning activities have been provided by the teacher which guided students to achieve the specific ends. After instruction, the teacher measured the involved pupil's achievement to notice if the stated objectives had been attained. In situations such as these, the test is valid, if items are clearly stated orally, since the learning experiences guided students to achieve the precise objectives, and the measurement procedures harmonized directly with the statement of objectives.

Consistency of results from students who have been administered an oral test in mathematics is important. One way of ascertaining internal consistency in administering a test is to compare odd versus even numbered items. Did those learners who scored high on the even numbered items also score high on the odd numbered items? The teacher may wish to rank each pupil in the class from high to low in the even numbered items on the oral test. The same also needs to be done for the odd numbered items. A reliability coefficient can

then be computed for internal consistency. In any teacher determined test be it oral or written, internal consistency reliability can be computed when comparing students in class with responses correct to even versus odd numbered items.

Sociometric Devices

Periodically, the mathematics teacher will wish to have pupils work cooperatively within committees. To determine committee membership, the teacher may want to utilize the sociometric device. Sociometric devices attempt to evaluate social, not academic growth.

The teacher may have students list on paper their first, second, and third choices in working on a committee. Learners need to be assured that their responses will be kept strictly confidential. Also, the mathematics teacher needs to mention to students that the results of the sociometric device will be utilized to determine committee membership.

What might the teacher appraise from the responses given by students as to whom they would prefer to work with on a committee?

1. students that are chosen frequently by others.

2. learners who are on the fringe area in that they are chosen, perhaps, only once and that being a third choice.

3. individuals who are complete isolates.

Guidance can be provided by the mathematics teacher in assisting isolates to becoming increasingly accepted by others. Changes here generally will take place slowly. Students who are isolates can be placed in committees in which other learners are highly accepting of others. Certainly an isolate should not be placed in a committee or small group of cliques.

The goal for any committee to attain is to achieve as much or more using this method of teaching as compared to other procedures.

In Closing

There are numerous means available to ascertain learner progress in the mathematics curriculum.

These include using:

1. checklists and rating scales.
2. standardized tests.
3. criterion referenced tests (CRT).
4. anecdotal records.
5. student products.
6. teacher written tests.
7. conferences with students.
8. parent-teacher conferences.
9. oral tests.
10. sociometric devices.

Each evaluation technique has its strengths with selected limitations also. The mathematics teacher needs to utilize a variety of procedures to appraise student progrress. With quality evaluation results from students, the teacher may truly develop a sequential mathematics curriculum from which all students might benefit optimally.