DOCUMENT RESUME

ED 139 627 SE 022 354

AUTHOR Okey, James R.
TITLE Using Student Performance to Measure Science Teacher Effectiveness.
PUB DATE Mar 77
NOTE 12p.; Paper presented at the annual meeting of the National Association for Research in Science Teaching (50th, Cincinnati, Ohio, March 22-24, 1977); Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Academic Achievement; Effective Teaching; *Evaluation; *Evaluation Methods; *Instruction; *Science Education; Science Teachers; Teacher Behavior; *Teacher Evaluation
IDENTIFIERS *Learner Outcomes; Research Reports

ABSTRACT

A rationale for using learner outcomes as the primary indicator of teacher effectiveness is presented. Specific procedures are outlined for collecting different kinds of pupil data to use in the evaluation process. Emphasis is given to the problems of making reliable inferences about teacher effectiveness and to the role that fellow teachers play in establishing reasonable standards of teacher effectiveness. (Author/MH)

USING STUDENT PERFORMANCE TO MEASURE SCIENCE TEACHER EFFECTIVENESS

James R. Okey
Department of Science Education
University of Georgia

A paper presented at the Annual Meeting of the
National Association for Research
in Science Teaching

Cincinnati, Ohio
March 23, 1977

USING STUDENT PERFORMANCE TO MEASURE SCIENCE TEACHING EFFECTIVENESS

How do you identify a successful automobile mechanic? How do you choose an effective lawyer? or butcher? or detective? Should we watch how skillfully the mechanic uses a wrench, how cogently the lawyer speaks, how artfully the butcher trims a roast, and how stealthily the detective stalks the criminal? Or should we examine how well the car runs, how successfully cases are prosecuted, the taste and appearance of meat, and the number of criminals apprehended?

How do you determine if a teacher is effective? Do you examine the quality and quantity of their questions, the eloquence of their lectures, the length of their hair or skirts, or the number of times they smile or frown? Or do you watch to see how many students are engaged in learning, how well they can read, whether they can measure or weigh, and whether the students seem confident and calm?

The questions don't seem difficult to answer. Skilled workers and professionals are effective when they accomplish the purpose of their jobs. We don't particularly care if mechanics are unshaven, if detectives are obtuse question askers, or if lawyers are flamboyant so long as they accomplish their purposes. Nor should we be excessively concerned with how teachers look and act when the influence they have on student achievement is a more important concern.

The purpose of this paper is to argue the case for using student achievement as the primary means of measuring science teaching effectiveness. Propositions to support this argument will be presented and defended. A number of specific procedures will be given for collecting data from students to use in assessing teaching effectiveness. Finally, some problems and cautions

when following the approaches suggested here will be examined.

THE CASE FOR USING STUDENT PERFORMANCE

Four propositions will be presented and defended that summarize the argument favoring the use of student performance as the principal means of assessing the effectiveness of teachers.

Proposition 1: Teachers are employed to help students learn and should be assessed accordingly.

A study was reported several years ago in which housewives, students and tradesmen were compared with teachers in their ability to teach students (Popham, 1971). The result was a standoff in effectiveness and, thus, a clear defeat for teachers. Teachers would be expected to do better at their jobs than persons without special training and experience, but this was not the case. Should we not expect certified teachers to be more effective than lay people and be willing to be assessed and rewarded by their effectiveness?

Few schools are organized to reward teachers for being effective. Instead the salary schedules provide increments in pay (admittedly small in many cases) for every teacher that stays employed and continues to aggregate course credit and advanced degrees. The teacher evaluations done in schools are often completed by a harried administrator or department head and focus on teacher characteristics (e.g., smiles, neat appearance) or actions (e.g., asks high level questions, reinforces student responses). The result of both the evaluation system and the pay schedule is to focus attention away from the primary function of the teacher -- helping students learn.

Proposition 2: Teachers more effectively devote their time to teaching when they are being evaluated by the results they produce.

It is difficult to operate a school in a highly efficient manner because many people with diverse opinions, purposes, and ideas need to work together. Yet that built in inefficiency is compounded by a teacher assessment system that focuses on the means of instruction (i.e., the procedures teachers follow) instead of the results of instruction. Rather than reduce the inefficiency by directing assessments and evaluations toward desired outcomes, we usually see a compounding of the problem by excessive attention to the appearance and actions of students and teachers.

What happens when teachers are held responsible for student outcomes? There have not been many studies on this topic but a report by McNeil (1967) illustrates what could be achieved on a wide basis if teacher assessment was focused on outcomes. McNeil assigned 44 elementary student teachers to two groups and directed them to work with pupils that had deficiencies in punctuation. One group of student teachers (control group) was told that they would be graded on their "professional characteristics and teaching methods". The other group (experimental group) was told that they would be graded on their ability to remedy some of the punctuation problems of their pupils. Following a two week period of instruction, the elementary school pupils were evaluated on their punctuation skills. Every measure of achievement showed that the pupils of the experimental teachers learned significantly more during the unit.

The result of focusing teacher evaluation on outcomes (student achievement) rather than procedures (teacher actions) is to put teachers clearly on task. That is, they devote their time and energies more to direct instruction

and less to peripheral and cosmetic concerns. When teachers are on task it is reasonable to assume that students are being given opportunities to learn and that is the most certain way to influence school achievement. A study in California (McDonald, 1976) showed that student opportunity to learn was strongly related to student achievement.

Proposition 3: Too little is known about the effects of teacher behaviors to use them as a basis for assessing teachers.

Several reviews have been made of studies that examined the relationships between what teachers do and what students learn (e.g., Rosenshine and Furst, 1971). Factors like teacher enthusiasm and clarity correlate positively and significantly with student outcomes. But correlational findings do not demonstrate causes and effects. Because teacher enthusiasm is associated with higher student achievement doesn't mean that it causes it -- and that is the problem. Only a few studies have been done that show the direct influence of teacher activities on student accomplishment (e.g., Gage, 1976, and Okey and Ciesla, 1973).

Experimental studies of the influence of teacher activities on student outcomes are needed to serve as a basis for planning teacher training and staff development programs. If teacher behaviors can be identified that influence what students learn, a great deal more direction is available for selecting behaviors and skills that teachers should spend time studying.

The situation we find today concerning links between teacher behaviors and student outcomes is one of hardly any experimental (i.e., cause and effect) findings and only a few correlational hints. Rosenshine (1976) has estimated that no more than 25 studies have been reported on even the most thoroughly

investigated teaching variable. The result is that the research base is too slim to serve as a reliable guide for selecting teacher behaviors for study and certainly too weak to serve as a standard for assessing teacher effectiveness.

Proposition 4: Student achievement is not assured even when validated teaching skills are used.

Suppose a number of experimental studies (like those just described) have been conducted and evidence is available to show that certain teaching skills influence what students learn. Can we then assess the skills of teachers and assert that they are effective if they demonstrate the appropriate actions and ineffective if they do not. We cannot -- and that peculiar condition is one we must live with even when many teaching skills have been validated (i.e., shown to influence student outcomes).

Even when teachers use validated teaching skills we can not be certain that they are effective in helping students learn. Conversely, when teachers fail to use validated teaching skills we can not be certain that they are ineffective. Said differently, a teacher can follow all the rules and students may still fail to learn or fail to follow the rules and yet produce good results. The reason this is so is because teaching is much like weather predicting. The weather predictor knows that frontal systems, air flow patterns, and terrain all influence the weather. But weather predictions are not perfect based on this information because there are many complex factors that can interact with one another. The same holds true for education. We may know that amount of practice, student motivation, and kinds of questions all influence what students learn. But the predictions of whether students will

learn in a given situation are less than perfect because the factors influencing learning are many and by no means all under the control of the teacher.

Procedures For Assessing Student Performance

Three procedures will be described that can be used to answer questions about the effectiveness of teachers. Each procedure involves collecting some kind of evidence from students as a means of assessing teachers. The procedures range from those with a successful history to those that show promise but have been little used.

Teaching Performance Tests

In a teaching performance test (Popham, 1971), teachers are given one or more objectives on a topic and told to prepare and deliver a short unit of instruction on the objectives. The instruction may be as short as a few minutes or last for several class periods. Post-tests on the objectives are provided along with a brief description of the topic. The teachers do not see the post test until instruction is completed although they should be able to infer the type of questions on it by examining the objectives. Several Likert-type items are also provided for assessing the attitudes of students toward the instruction they receive. Thus the teaching performance test is a compact way of assessing how able teachers are in bringing about student achievement for a short unit of instruction. The student scores on the post-test and responses on the attitude items provide evidence of both cognitive and affective effectiveness.

Teaching performance tests provide a controlled setting for assessing

the effectiveness of teachers in either pre-service teacher training programs or in on-the-job settings. Popham selects novel topics for his performance tests so that prior knowledge of teachers and students has a minimal influence on the results. Rezba, et al. (1975) have developed performance tests in science to assess the effectiveness of student teachers in a field-based methods course. Information on how to construct and use teaching performance tests is available (Popham, 1973).

Contract Teaching

McNeil (1967) describes an example of a teaching contract in which pre-service teachers obtained agreement from their supervising teachers on specific objectives that were to be accomplished by students. If most of the students achieved most of the objectives the student teachers were to be graded as effective. For lesser accomplishment with students, the student teachers were to be graded accordingly. The same tactic of establishing standards for student achievement can be used by inservice teachers for either single units of instruction or entire school terms. The key is to establish and agree on specific objectives to be accomplished and specific levels of student performance on the objectives. The agreement can be made between a teacher and either an administrator or department head. Both the level of the objectives and the level of student performance can be adjusted for the situation in which the teacher works. Some students are more teachable than others and contract expectations can be adjusted to take this into account.

Peer Expectations

The principal danger in using student achievement as a measure of teaching effectiveness is that unreasonable performance expectations may be set. In some schools it is relatively easy for even unskilled teachers to show top student achievement. In other schools the most diligent efforts may result in minimal learning. Another problem to be considered is the danger of teachers setting easy achievement standards if they know they will be judged on how well their students achieve. How are these factors to be taken into consideration when deciding if a teacher is effective?

A method to counteract the problems just mentioned is use of peers (fellow teachers who know both the teaching content and the instructional setting) to set reasonable standards for pupil achievement. How would this work? Suppose a ninth grade physical science teacher is to be assessed for effectiveness. Two or three colleagues of the teacher can examine the objectives the teacher has set for a unit and judge their quality. They need to determine whether they are important to the topic, comprehensive in coverage, and appropriate for the students. Secondly the colleagues need to set reasonable standards for achievement of the objectives with the students. If the students are both able and highly motivated they would expect higher performance than for students of less ability and lower aspirations.

Use of peer expectations is a procedure that may get us out of the dilemma of how to establish expectations for teacher effectiveness. Peer expectations can take into consideration what is being taught, who is being taught, and where instruction is taking place.

Some Final Considerations

We have for too long avoided direct measures of teaching effectiveness. The result is a stagnant profession that passes out rewards to all teachers without consideration for their different abilities and efforts. It has resulted too in teachers that are less able than they could be.

A rapid change to a teacher accountability system based on student achievement is not likely and unwise considering the problems in assessing teachers (cf., Berliner, 1976; Soar, 1977). But we could and should begin to move in the direction of assessing the effectiveness of teachers based on measures of what learners are able to do.

REFERENCES

Berliner, D. A status report of the study of teacher effectiveness. Journal of Research in Science Teaching, 1976, 13, 369-382.

Gage, N. A factorialy designed experiment on teacher structuring, soliciting, and reacting. Journal of Teacher Education, Spring 1976, 27, 35-38.

McDonald, F. Report on phase II of the beginning teacher evaluation study. Journal of Teacher Education, Spring 1976, 27, 39-42.

McNeil, J. Concomitants of using behavioral objectives in the assessment of teacher effectiveness. Journal of Experimental Education, Fall 1967, 36, 69-74.

Okey, J. and Ciesla, J. Designs for the evaluation of teacher training materials. AV Communication Review, 1973, 21, 299-310.

Popham, W. Performance tests of teaching proficiency: Rationale, development, and validation. American Educational Research Journal, 1971, 8, 105-117.

Popham, W. Evaluating instruction. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Rezba, R., Lahnston, A., Lapp, D., and Willcutt, R. Pupil growth in classification skills as a consequence measure of learning site on preservice elementary teachers. Paper presented at the National Association for Research in Science Teaching Annual Meeting, Chicago, April 1974.

Rosenshine, B. Recent research on teaching behaviors and student achievement. Journal of Teacher Education, Spring 1976, 27, 61-64.

Rosenshine, B. and Furst, N. Research in teacher performance criteria. In B. O. Smith (Ed.), Research in Teacher Education. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Soar, R. Teacher assessment problems and possibilities. In G. Borich (Ed.), The Appraisal of Teaching. Reading, Massachusetts: Addison-Wesley, 1977.