DOCUMENT RESUME

ED 172 367 EA 011 779

| | |
|---|---|
| TITLE | Teacher Evaluation. Research Action Brief Number 4. |
| INSTITUTION | Oregon Univ., Eugene. ERIC Clearinghouse on Educational Management. |
| SPONS AGENCY | National Inst. of Education (DHEW), Washington, D.C. |
| PUB DATE | Jul 79 |
| NOTE | 5p. |
| AVAILABLE FROM | ERIC Clearinghouse on Educational Management, University of Oregon, Eugene, Oregon 97403 (free) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | Effective Teaching; Elementary Secondary Education; Learning Characteristics; Research Methodology; Research Problems; *Teacher Evaluation; Teaching Methods |
| IDENTIFIERS | *Beginning Teacher Evaluation Study; Information Analysis Products |

ABSTRACT

Research into teacher evaluation techniques has not provided conclusive answers due to a number of research problems, including the subjectivity of traditional teacher observation techniques and the lack of sensitivity of newer, more objective techniques. A five-year research project in California, the Beginning Teacher Evaluation Study (BTES), tried to avoid these problems through the use of nonstandardized testing, teacher questionnaires and interviews, and multiple objective behavioral observations of interaction between students and teachers. A wide range of classrooms and schools were sampled to allow for socioeconomic differences. The researchers found that students learned best when spending sufficient time deeply engaged in a task with a high success rate. Teachers were most effective at creating such learning situations when they did a good job of diagnosing student needs, prescribing appropriate tasks, structuring presentation of the tasks, monitoring student progress, and providing adequate feedback. While BTES, despite its name, never directly addresses itself to teacher evaluation, its findings imply answers to the questions of what to measure when evaluating teachers and how to measure it. The study also points the way to future definitive research of a more immediately practical nature. (Author/PGD)

ERIC
CLEARINGHOUSE
ON EDUCATIONAL
MANAGEMENT

# RESEARCH ACTION BRIEF

Number 4 July 1979



Each *Research Action Brief* reports the findings of significant empirical research studies on a topic in educational management. From these findings implications are drawn for the operation of today's schools, thus serving as a guide for enlightened administrative action.

This *Research Action Brief* was prepared by the ERIC Clearinghouse on Educational Management for distribution by the National Association of Secondary School Principals.



## Teacher Evaluation

Although report card time can be agonizing for teachers, the dilemmas they face when evaluating students look rather easily solvable when compared to the dilemmas administrators face when evaluating teachers. Teachers have to puzzle over only what grades to give their students. Administrators, if they are thoughtful and honest, are not even sure what teacher characteristics they ought to be evaluating or what criteria to use.

What kind of report card should the teacher get? Do the A's go to the "hard" teacher who makes sure the homework is difficult and does not give many high marks or to the teacher who makes sure the students have a lot of success experiences? What is good teaching and how can it be recognized?

## What and How to Measure

Discovering the best way to evaluate teachers rests on the answers to two very basic questions. What do we measure, and how do we measure it? For many years, teacher evaluation research has attempted to answer these questions, but useful information has been slow in emerging.

Problems with research on evaluating teachers have been so serious and so pervasive that some critics dismiss all this research as just so much worthless paper. Research reviews are filled with cogent criticism. Traditional teacher observation methods that examine such characteristics as personality or management skills are subjective and imprecise. Newer performance-based approaches attempt objectivity but have problems of their own.

The standardized tests that performance-based methods sometimes use are unable to separate effects of teaching from other effects like those of watching TV or having a high IQ. These tests also cannot reflect the unique content actually taught in each classroom and are not sensitive enough to record some students' gains over a period as short as one year.

Both observation and performance-based systems often disregard powerful socioeconomic factors and base evaluations on only a small number of lessons, thus ignoring conclusive findings that many teacher behaviors vary a lot from lesson to lesson.

So when researchers undertake a study on teacher evaluation that addresses these problems, the results are worth more than a second glance. The Beginning Teacher Evaluation Study (BTES) commissioned by the California Commission for Teacher Preparation and Licensing has made some impressive strides toward answering the questions of what to measure and how to measure when evaluating teachers. Although it is certainly not the only tightly run and intelligently planned teacher effectiveness study, it is worth careful scrutiny if only for its scope alone.

Made up of three phases and numerous substudies, the research stretched over five years and generated more than fifty papers and reports by the end of 1978. Involving more than 200 teachers in thirteen California school districts, the study examined many more variables than can even be touched on here.

Although its title specifies "beginning" teachers, the vast majority of subjects were experienced teachers, and the study's implications apply to all teachers regardless of years of experience. The study focused on only two grades, but the findings are broad generalizations about the behavior of teachers and students that appear to be applicable to virtually every teaching and learning situation.

## Avoiding the Usual Problems

The BTES study avoided some of the usual research problems in a number of ways. First, researchers did not rely on traditional subjective teacher observation methods but instead attempted to correlate measurable teacher behaviors with student achievement. In Phase III of the study researchers began by identifying student behaviors conducive to learning and moved from these to identify teacher behaviors that fostered the desired student behaviors. To measure student and teacher behavior, the research team developed two classroom observational procedures.

Marliave and his colleagues have explained that one observation method used in Phase III of the study was a method of coding classroom interaction using a time-sampling procedure. The time coding forms measured the time span in minutes for each specific classroom activity. In addition, they indicated such things as particular content area of instruction, level of difficulty of instruction, and level of engagement of students in each task.

The other observation method uses rating scales to measure fifteen different variables. These scales rate such variables as classroom atmosphere, academic orientation of instruction, and clarity of instructional communication. In addition to observational procedures, four teacher interviews and one teacher questionnaire were developed. These methods are important because, although not intended specifically for teacher evaluation, they suggest new and more valid ways to evaluate teachers.

Unlike some teacher evaluation researchers, BTES researchers did not use standardized tests to measure student learning but instead developed their own achievement tests. Filby and Dishaw explain that these tests, initially developed in Special Study A of the BTES research, have several advantages over standardized achievement tests. First, they do not merely test an overall subject matter area but instead are broken down into many subtests covering subareas of the course content. With such subtests, it is easier to determine if what is being tested was actually taught in the classroom. The devised tests are also more sensitive than standardized tests to learning over short time periods. The items on these instruments were carefully pretested to ensure that they would reflect that learning had actually taken place. They were also examined to determine whether they were measuring learning or ability.

Many teacher evaluation studies look at the effects of teaching for one class over one class period. The socioeconomic status of the class and the different effects of teaching on different socioeconomic groups are often ignored. Thus, it is unclear whether the findings apply to any other classrooms or to socioeconomic groups other than the one assessed. Data for the final phase of the BTES research, however, were drawn from fifty different classrooms. Students in the sample came from both urban and suburban schools, and high socioeconomic status schools were excluded from the study. The sample was about one-third nonwhite.

Although many studies that attempt to identify teacher behaviors correlated with achievement are based on only one class period, BTES research was based on observation and

### References


1. California Commission for Teacher Preparation and Licensing. "An Overview and Future Activities." Beginning Teacher Evaluation Study *Newsletter*, (October 1978), entire issue, 8 pages.

2. California Commission for Teacher Preparation and Licensing. "Beginning Teacher Evaluation Study. Phase III Technical Reports and Notes." Sacramento, California: 1977.

3. Filby, Nikola N., and Dishaw, Marilyn. *Development and Refinement of Reading and Mathematics Tests for Grades 2 and 5. Beginning Teacher Evaluation Study. Technical Report Series. Technical Report III-1. Continuation of Phase IIIA.* San Francisco: Far West Laboratory for Educational Research and Development, 1975. 234 pages. ED 151 375.

4. Fisher, Charles W., and others. *Teaching and Learning in the Elementary School: A Summary of the Beginning Teacher Evaluation Study. Beginning Teacher Evalution Study. Technical Report Series. Report VII-1.* San Francisco: Far West Laboratory for Educational Research and Development, 1978. 50 pages. ED number not yet assigned.

5. Marliave, Richard. *Observable Classroom Variables. Beginning Teacher Evaluation Study. Technical Report I-2.* San Francisco: Far West Laboratory for Educational Research and Development, 1976. 33 pages. ED 146 162.

6. Marliave, Richard; Fisher, Charles W.; Filby, Nikola N.; and Dishaw, Marilyn. *The Development of Instrumentation for a Field Study of Teaching. Beginning Teacher Evaluation Study: Technical Report Series. Technical Report I-5.* San Francisco: Far West Laboratory for Educational Research and Development, 1977. 273 pages. ED 151 378.

7. Tikunoff, William J.; Berliner, David C.; and Rist, Ray C. *Special Study A: An Ethnographic Study of the Forty Classrooms of the Beginning Teacher Evaluation Study Known Sample. BTES Technical Report Series No. 75-10-5.* San Francisco: Far West Laboratory for Educational Research and Development, 1975. 264 pages. ED 150 110.


Do not order from the Clearinghouse

Copies of all the above references are available from the California Commission for Teacher Preparation and Licensing, 1020 "O" St., Sacramento, CA 95814. Free.

assessment of several months of classroom activity. In Phase III of the study, most classes were observed a total of twenty times.

## Findings

Some of the study's preliminary findings appear to be almost as interesting as the final findings. For example, Marliave, in reviewing the findings of Phase II of the study, reports that one of the many findings to emerge was that quantity of teacher preparation is unrelated or perhaps even negatively related to amount of student learning. Teachers who spend long hours poring over lesson plans are not necessarily better teachers than those who do not. Marliave concluded that quality and not quantity of teacher preparation is related to effectiveness.

Phase III's Special Study A revealed several behaviors that discriminated between teachers who had been identified as either more or less effective at teaching an experimental teaching unit. One of Study A's many findings reported by Tikunoff, Berliner, and Rist was that more effective teachers did not often belittle or berate their students to make them behave, and they did not punish the whole class for misbehavior by a few.

Summary reports indicate that BTES researchers are most excited about the findings of Phase III-B, the final phase of the study. These findings represent the culmination of five years of research. This part of the study identified both teacher behaviors and student behaviors that correlate with long-term student achievement.

**Indicators of student learning.** BTES researchers developed a measure of student learning called "academic learning time" (ALT). For a student to accumulate ALT, three things have to occur. The student must be spending time on an academic task; the student must be really engaged in and paying attention to the task; and he or she must be performing the task with high success. The BTES study revealed that the more ALT a student accumulates, the more the student is learning.

As Fisher and his colleagues explain, part of the value of ALT lies in its ability to determine whether learning is occurring at the time it actually is occurring without relying on achievement tests. It is also useful for specifying why learning is not occurring. When students are not learning, we can conclude that the cause is either lack of engagement in an academic task or a low success rate.

The implication of ALT that is most likely to be surprising to educators is that students learn more when classwork seems rather easy to them. If students give a large number of correct answers, their subsequent achievement gains in that area will be higher. Fisher and his colleagues found that "if the task is very difficult and the student produces few correct responses during the task, then the activity will not yield much learning for that student." These findings may come as a surprise to those who believe that students are learning more when classwork is very difficult.

Fisher and BTES researchers qualify these findings with the warning that not all of a student's time should be spent in high success activities. If all tasks are easy and no new challenging material is presented, students cannot progress. The researchers suggest particularly that older students or good students may benefit from spending quite a bit of time at a medium success level.

**Indicators of teacher effectiveness.** In addition to identifying ALT, another major outcome of the study was the identification of things teachers do that increase student learning. The researchers called these behaviors *diagnosis, prescription, presentation, monitoring,* and *feedback.*

Regarding diagnosis, students learn more when teachers know more about what their individual students can and cannot do. Achievement scores are high when teachers have a good grasp of each student's ability. Regarding prescription, a teacher's ability to match instruction with the needs and skill levels of individual children was related to a high student success rate. Students learn more when teachers assign them appropriate tasks.

Regarding presentation, students made fewer errors when teachers spent more time structuring the lesson and giving directions. Observers of particularly successful classes often mentioned that the teacher had a regular routine of beginning each lesson with a presentation in a group setting.

Regarding monitoring, students learn more when teachers keep track of student progress on instructional tasks. The major form of this monitoring is teacher questioning. It also takes the form of surveying the work of students who are engaged in seatwork. Regarding feedback, students learn more when they frequently learn whether their answers are right or wrong. Feedback may come from checking homework, from oral question and answer sessions, or even from programmed texts. All kinds of feedback had a stronger and more consistently positive relation to achievement than did presentation or monitoring.

To summarize, all these variables—diagnosis, prescription, presentation, monitoring, and feedback—were found to be strong indicators of student learning and, therefore, of truly effective teaching.

A third important finding to emerge from Phase III-B concerns classroom environment. Students learn more when a teacher fosters an environment where academic learning is valued. BTES researchers maintained that a teacher who is committed to the goal of producing academic learning will instill that goal in students.

Students also appear to learn more when they are willing to work cooperatively and take responsibility for their classroom work. When a teacher creates an environment that encourages student cooperation and responsibility, higher achievement scores result.

## Implications

The BTES study, belying its title, never directly addresses itself to teacher evaluation. Nevertheless, it is full of implications useful for evaluators. The findings point toward answers to the questions of what to measure and how to measure when doing teacher evaluation.

First, the findings of Phase III-B on important teacher behaviors address the problem of what to measure when evaluating teachers. They suggest the kinds of behaviors to look for in effective teachers. Are teachers committed to and capable of diagnosing the learning needs of each student? Do they prescribe appropriate learning activities for each student? Do they present lessons to the class through

structured explanations and good directions? Do they monitor their students' work by asking them questions and checking their progress in seatwork? Finally, do they provide frequent feedback about the accuracy of their written work and oral answers?

Fisher and other BTES researchers have warned that there are many different, equally effective ways to accomplish each of these five functions. For example, teachers may provide feedback in many different ways—question and answer sessions, marking homework, or reward systems. No particular type of feedback seems better than any other. The important thing is that all these functions are accomplished in some way by the teacher and that he or she believes they are important.

Findings from Phase III-B on classroom environment, too, have implications for teacher evaluation. Is the classroom environment one in which academic learning is valued and encouraged? Are students cooperative and responsible for their academic work? If the answer to these questions is yes, then it is more likely that the teacher is an effective one.

The concept of academic learning time suggests possible answers to the question of how to measure teachers' performance. Although BTES researchers so far make no claims about the value of using ALT to measure teacher effectiveness, the implications seem obvious. The observational techniques used in the BTES study to measure ALT might be used in classrooms to determine how much learning actually is occurring. How much time do students spend on academic activities? Do they appear to be really engaged in and paying attention to what they are doing? Are they experiencing a high rate of success?

Many of the substudies' findings also have implications for teacher evaluation. For example, the findings from Phase II on teacher preparation suggest that the length of time teachers spend preparing for class is not a valid teacher evaluation measure.

Many findings of Special Study A also have implications for teacher evaluation. To take just one example, the findings about methods of punishment suggest that positive ratings in this area ought to go to teachers who do not berate students for mistakes or shortcomings and who do not habitually punish the whole class for misbehavior by a few students.

Of course, the BTES findings are not a panacea and leave many questions unanswered. Some of these questions concern low achieving students. Even when it is clear that overall ALT in a class is low and that a teacher does not spend much time in diagnosing, prescribing, presenting, monitoring, and giving feedback, it is not clear where the problem lies. Is it in the teacher or in a class that is so large or has students with such diverse abilities or is so hard to handle that the teacher is unable to give the individual attention necessary to every student?

No, the BTES does not hold all the answers, and furthermore BTES researchers have warned that the study is a seminal one whose models need to be refined and whose findings need to be replicated. It is nevertheless a remarkable study both in its findings and in the careful and thorough way it attempted to avoid methodological problems of previous studies. The study points the way, perhaps more clearly than any previous study, toward identifying teacher effectiveness. And it is on a clear definition of teacher effectiveness that all teacher evaluation must rest.

The Educational Resources Information Center (ERIC) is a national information system operated by the National Institute of Education. ERIC serves educators by disseminating research results and other resource information that can be used in developing more effective educational programs. The ERIC Clearinghouse on Educational Management, one of several such units in the system, was established at the University of Oregon in 1966. The Clearinghouse and its companion units process research reports and journal articles for announcement in ERIC's index and abstract bulletins.

Research reports are announced in Resources in Education (RIE), available in many libraries and by subscription for $42.70 a year from the United States Government Printing Office, Washington, D.C. 20402. Journal articles are announced in Current Index to Journals in Education. CIJE is also available in many libraries and can be ordered for $62 a year from Macmillan Information, 100D Brown Street, Riverside, New Jersey 08075.

Besides processing documents and journal articles, the Clearinghouse prepares bibliographies, literature reviews, monographs, and other interpretive research studies on topics in its educational area.

This publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to the National Association of Secondary School Principals for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either the National Association of Secondary School Principals or the National Institute of Education.

EA 011 779

ERIC Clearinghouse on Educational Management
University of Oregon
Eugene, Oregon 97403