2.5
2.2
2.0
100 mm
1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz1234567890
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890
1.0 mm
1.5 mm
2.0 mm
A5

DOCUMENT RESUME

ED 290 764 TM 011 004

AUTHOR Redfield, Doris L.
TITLE Expected Student Achievement as a Potential Factor for Assessing Teacher Effectiveness.
PUB DATE Nov 87
NOTE 17p.; For a related document, see TM 011 005. Paper presented at the Annual Meeting of the Mid-South Educational Research Association (16th, Mobile, AL, November 11-13, 1987).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Elementary Secondary Education; Evaluation Criteria; Evaluation Methods; Predictive Measurement; Standardized Tests; Teacher Effectiveness; *Teacher Evaluation; Teacher Responsibility; Teaching Skills
IDENTIFIERS *Kentucky Career Ladder Plan

ABSTRACT

The Kentucky Career Ladder Commission has funded a special project on "expected student achievement," to study the evaluation of teachers while avoiding the indefensible use of standardized student achievement tests. As proposed, the plan uses student achievement as one aspect of evaluation. The problem is in determining the degree to which student achievement, however defined, is attributable to any particular source. The project found that teachers value general and specific academic and non-academic outcomes. A common core of student achievement goals might be developed for evaluating teachers through professional consensus with weighted significance for each goal. The project has identified many problems associated with using student achievement test results as it has begun to develop alternatives to the use of standardized test data for this purpose. The procedures piloted during the first year (1986-87) of the special project have potential for development as part of a teacher evaluation system which includes student achievement outcome data. (SLD)

Expected Student Achievement as a Potential Factor for Assessing Teacher Effectiveness

Doris L. Redfield [a]

Western Kentucky University

Running Head: Expected Student Achievement

Paper presentation. In R.S. Pankratz (Symposium Organizer), The Assessment of Teacher Performance Using Classroom Observations, Structured Interviews, Goal Attainment and Expected Achievement: Lessons Learned. 1987 Annual Meeting of the Mid-South Educational Research Association. Mobile, AL

[a] Doris L. Redfield is on leave from Western Kentucky University to the U.S. Department of Education's Office of Educational Research and Improvement (OERI). However, no official support or endorsement of this work by the U.S. Department of Education is intended or should be inferred.

Expected Student Achievement as a Potential Factor for Assessing Teacher Effectiveness

Steps 3 and 4 of Kentucky's Career Ladder Plan call for the inclusion of student achievement data in the evaluation of individual teachers. Measures of achievement are most often conceptualized as scores on standardized achievement tests. However, using students' standardized achievement test scores to evaluate individual teachers is not defensible (i.e., fair, reliable, valid) for a variety of reasons. To satisfy its mandated plan, while avoiding the indefensible use of standardized achievement tests, the Kentucky Career Ladder Commission funded a special project on "expected student achievement." The primary purpose of this paper is to consider the issues surrounding the Commission's decision. Project related procedures and findings are detailed elsewhere (Redfield, 1987a, Redfield, 1987b; Redfield & Craig, 1987a; Redfield & Craig, 1987b).

Why was Student Achievement Included in Kentucky's Proposed Teacher Evaluation Plan?

Both the Kentucky Career Ladder Committee responsible for writing Kentucky's Career Ladder Plan and the Kentucky Career Ladder Commission responsible for implementing the Plan agonized over the role of student achievement in teacher evaluation systems. They clearly recognized the controversial nature of student achievement. However, throughout the development of the Plan and the early stages of implementation, they may not have fully appreciated the complexity of the issues contributing to that controversy. The logic for including student achievement in the teacher evaluation plan might be described as follows. Student achievement is the goal of education. Teachers are educators. Therefore, teachers

are responsible for student achievement. Further, taxpayers need evidence that teachers are meeting their responsibility.

Why Was There a Separate, Special Project on Student Achievement?

The conclusion that teachers should be held accountable for student achievement raised at least two problems for the envisioned Career Ladder Pilot Study. One, it implied that teachers, rather than students are responsible for student achievement. Two, achievement was not explicitly defined; but, it seemed to have been implicitly defined as scores on the state mandated, standardized achievement test (viz., Kentucky Essential Skills Test). These two problems raised a series of issues.

Issues in Defining Achievement. Because the definition of "achievement" hinges on the student outcomes valued for particular students under particular circumstances, it has not, and probably never will be, universally defined. Clearly there are educational outcomes which are valued by teachers and parents but which are not measurable using traditional, standardized tests. Examples of such outcomes include critical thinking, motivation, self-discipline, self-esteem, positive attitudes, and acceptable behaviors.

The definition of achievement may not be limited to minimum competence in basic skill areas. In fact, not all teachers teach subject matter measured by standardized achievement tests currently available (e.g., Kentucky Essential Skills Test, Comprehensive Test of Basic Skills). For example, the Kentucky Essential Skills Test (KEST) measures essential skills in the areas of Reading, Mathematics, Writing, Spelling, and Library/Research/Reference. However, only 36% of Kentucky's teachers work in self-contained classroom situations with "normal" students

in the basic skills areas measured by the KEST. This fact adds fuel to the controversy over including student achievement data in the evaluation of teachers. Not all teachers' contracts hold them accountable for teaching what the KEST measures, standardized testing is not mandated (and usually not available) in non-basic skill content areas, and reasonable expectations of student achievement vary. Average performance or gain is not a defensible expectation for nonaverage students (e.g., handicapped, disadvantaged, gifted).

Teacher Responsibility Issues. Standardized tests of student achievement are designed to reliably assess students' performance, not teachers' effectiveness. The argument that student achievement test scores do not validly reflect teaching effectiveness is not meant to suggest that teachers cannot evaluate their students' accomplishments. As evidenced by various teacher incentive programs or pilot studies (Cornett, 1986), teachers can, and do, evaluate their students even outside the basic skill areas. However, even when teachers validly assess their students' achievements, that achievement data may not validly represent teachers' effectiveness. Collecting one type of information (i.e., student outcome data) intended for one purpose (i.e., evaluation of student achievement) and using that information for a second, unintended purpose (i.e., teacher evaluation) may seem logically appropriate; but, it may also be technically inadequate.

The problem, then, is not the inability of teachers to assess the quality of their students' accomplishments. Rather, the problem is determining the degree to which a student's achievements, however defined, are validly attributable to any particular source (e.g., teacher, program,

socioeconomic status, innate ability). This problem is often considered technical in nature. However, problem solution may be much more dependent upon adequately conceptualizing and understanding the problem than upon the perfection of some statistical technology. Clearly, there are many reasons why some students do not perform well; none of these reasons may relate to the teacher's performance. Unless a teacher is the sole influence on a student's learning, not all of a student's achievement may be attributed to a particular teacher.

Since we do not yet know how to directly measure teaching effectiveness, as defined by Medley (1987), we are relegated to the use of proxy measures. While student achievement offers one proxy measure, it is not a particularly defensible one. Good teaching may be a necessary but insufficient ingredient in student achievement and/or student achievement may be a necessary but insufficient indicator of teaching effectiveness. Compared to indicators of student learning, observations of teaching behavior may provide more defensible measures of teaching because: (a) they are more directly attributable to particular teachers and (b) there is empirical evidence that certain teaching behaviors are positively related to student achievement (e.g., Anderson, 1972; Brophy, 1981; Doyle, 1985; Good & Beckerman, 1978; Kounin, 1970; Lockwood, 1978; Luiten, Ames, & Ackerson, 1980; Medley, 1977; Peterson, Marx, & Clark, 1978; Redfield & Rousseau, 1981; Rosenshine, 1979).

Summary of Project Procedures/Findings and the Issues They Raised

The following paragraphs provide a brief summary of the special project on Expected Student Achievement. Project details are described elsewhere (Redfield, 1987a; Redfield, 1987b; Redfield & Craig, 1987a; Redfield & Craig, 1987b).

Procedures and Related Issues. In September, 1986. 26 teachers representing a wide variety of grade levels (kindergarten through grade 12) and teaching areas (special education, gifted, vocational arts, visual arts, social sciences, basic skill areas, etc.) were identified for participation in the Project.

From late September to early October individual teachers negotiated a set of Student Achievement Outcome goals with their principals. Using a five-point scale, each teacher and his/her principal reached agreement on: (a) the educational significance of each goal, (b) the difficulty of attaining progress toward each goal, and (c) the degree of relationship between each goal and the documentation proposed by the teacher for demonstrating progress toward the goal. Each of these three negotiation items represented a concern or issue. For example, one concern of project participants was that teachers migh be unduly rewarded for accomplishing trivial goals. Hence, using a five-point scale, each teacher and his/her principal attempted to guage the educational significance of each goal. A second concern of project participants was that the difficulty of attaining any particular goal be taken into account when assessing the degree to which the goal had been met. Therefore, using a five-point scale, each teacher-principal pair negotiated agreement on the difficulty of attaining each goal. A third issue, arising out of efforts to document goal progress, was the difficulty participants experienced in operationally defining their goals. This difficulty suggests a need for teacher, as well as principal, training.

Once a teacher and his/her principal negotiated a set of Student Achievement Outcome goals, the teacher worked toward those goals throughout the 1986-87 Project year. Near the end of the Project year, participating teachers again met with their principals to: (a) share the results of their efforts and (b) negotiate agreement concerning the degree to which each Student Achievement Outcome goal had been met.

Findings and Related Issues. The purpose of the first year (1986-87) of the Student Achievement project was to address the three questions described below. The data-based answers to each of these questions raised additional issues.

What sorts of learning outcomes do teachers want for their students? The data indicate that teachers value a variety of student outcomes, categorized as follows: (a) academic outcomes that are specific to the subject matter area in which a teacher teaches (e.g., mastery of math facts), (b) academic outcomes that are nonspecific or general and cut across subject matter areas (e.g., writing skills that can be used to write essays in english class as well as lab reports in science), (c) nonacademic outcomes that are specific to an individual teacher's teaching-learning situation (e.g., self-help skills such as tooth brushing among handicapped students), and (d) nonacademic outcomes that are nonspecific or general and that seem to be valued by most teachers (e.g., self-confidence, self-discipline, motivation to learn). All project participants worked toward goals in each of the four categories.

Which desired outcomes are common across subject matter areas and grade levels; and, which outcomes are unique to particular subject matter

areas and grade levels? The goal categorization system used by project participants lends itself well to addressing this second question. The goals they labeled as "specific," whether academic or nonacademic in nature, were viewed as peculiar to the unique subject matter or students being taught. Half (n=56) of the documented Student Achievement Outcome goals fit the "specific" category. The remaining documented goals (n=55) were considered "general" in nature, meaning that they would apply across grade levels, content areas, and student types. These findings suggest that there may be a common core of valued outcomes that might be included in the evaluation schemes for most teachers. The findings further emphasize the need to tailor evaluation schemes to validly reflect a wide variety of teaching-learning situations.

When standardized test scores cannot be defensibly used, how do teachers, particularly teachers in nontraditional teaching areas, document the degree to which desired student outcomes are accomplished? The data address this third question by showing that teachers use a variety of methods for assessing student outcomes when standardized tests cannot be defensibly used. Tables 1 and 2 summarize the types and formats of submitted documentation. A documentation concern of project participants was that documentation paperwork be kept to a minimum. For 63 of their 111 collective goals, participants were able to restrict documentation to one page or less. Documentation in excess of a page usually included work samples or information on multiple classes.

Insert Tables 1 and 2 about here

Discussion/Conclusions/Recommendations

There is professional and public agreement that teachers should be evaluated. There is also agreement that student achievement is the educational "bottom line." However, there is great controversy over the role that student achievement can, or should, fairly play in teacher evaluation systems. One source of controversy is that "achievement" has not been and, perhaps, cannot be universally defined. We might consider abolishing the term, "achievement." Instead, we might define specific terminology for varieties of achievement much as the Eskimos have a variety of expressions for particular kinds of snow.

Acknowledging the variety of student outcomes constituting achievement does not necessarily mean that every teacher's evaluation plan needs to be totally unique. In fact, data yielded by the Expected Student Achievement project indicate that all teachers, regardless of the students or types of subject matter they teach, work toward some similar or "general" Student Achievement Outcomes. Perhaps a common core of Student Achievement Outcome goals and documentation procedures could be developed. Goals "specific" to the needs of a particular teacher's students might be added to the common core. For example, both an elementary special education teacher and a secondary level science teacher might be expected to teach their students how to use classroom materials/equipment properly, communicate clearly, and behave appropriately. However, these two teachers may rightfully have different goals or expectations specific to their teaching situations. The elementary level special education teacher may teach students how to accurately write their names, addresses, and phone numbers; whereas, the secondary level science teacher may need to teach students how to write an accurate lab report.

Pursuing an evaluation system that includes a core of expected student outcomes, supplemented by outcomes associated with specific teaching-learning needs, suggests that aspects of evaluation are formative in nature. Such a system would allow teachers to risk less than perfect summative evaluations for the good of students' learning. Teachers would be able to revise their instructional and goal documentation plans as necessary, based upon their students' changing needs.

Ways to enhance the fairness of a flexible, formative evaluation system might include: (a) using professional concensus to identify minimally acceptable outcomes that are reasonable expectations for all teachers, (b) weighting the educational significance of each Student Achievement Outcome goal that is specific to the needs of a particular teaching-learning situation, (c) weighting the difficulty of attaining each Student Achievement Outcome goal that is specific to the needs of a particular teaching-learning situation, and (d) developing an appeals process that may be used (but not abused) by teachers, evaluating supervisors, parents, and/or students.

Kentucky's special project on Expected Student Achievement, implemented during the 1986-87 school year, was but one step in a proposed five-year research and development plan. A flexible, formative evaluation system allowing for both the professional development and accountability of teachers is worthy of development. Such a system cannot be developed quickly or inexpensively and it should not be viewed as a replacement for the appropriate use of standardized tests. Standardized test scores may provide one indicator of the degree to which a school or district teaches

basic skills or as one indicator, over several years, that a teacher is doing an inadequate job of teaching basic skills. To continue the development and testing of a teacher evaluation system that calls for teacher accountability and allows for professional development, at least the following events would need to occur over a continuing two to four year period of time:

1. Decide if the goal is to reward excellent teachers and/or remediate incompetence among teachers and/or dismiss incompetent teachers. (Despite the fact that Kentucky's Career Ladder Plan calls for the rewarding of excellence, the issue seems yet to be resolved).
2. Determine if a relatively large number of teachers and principals, given adequate training and support, are able to negotiate Student Achievement Outcome goals and appropriate assessments for goal attainment.
3. Determine if this relatively large number of teachers and principals could provide a sufficient variety of Student Achievement Outcome goals and assessment techniques for the development of a menu from which core goals and assessment techniques could be validated against professional concensus.
4. Determine the number of teachers with whom principals or other supervisors could reasonably work.
5. Test a system for taking into account the significance and difficulty of Student Achievement Outcome goals.
6. Determine the degree to which the process is able to differentiate good teachers from the best teachers.
7. Develop and test an appeals process.

8. Determine how to provide school personnel with the ongoing support needed for maintaining development efforts to enhance Student Achievement Outcomes.

9. Develop and test instruments for specifying, documenting, and evaluating Student Achievement Outcome goals.

10. Develop and test training programs for teachers and the supervisors responsible for assisting and/or evaluating them.

The primary problem with using student achievement data in the evaluation of teachers is not the "goodness" or "badness" of standardized tests or the inability of teachers to assess the quality of their students' accomplishments. Rather, the problem is the misuse of standardized test data and the inability to yet determine the degree to which a student's achievements, however defined, are validly attributable to any particular teacher.

Kentucky's special project on Expected Student Achievement has focused on understanding the problems associated with using student achievement data to evaluate teachers. This project also implemented and documented procedures that may serve as alternatives to the exclusive use of standardized achievement tests as indexes of student achievement and indicators of teacher effectiveness. Finally, the procedures piloted during the first year (1986-87) of the special project on Expected Student Achievement have potential for development as part of a teacher evaluation system which includes student achievement outcome data.

Bibliography

Anderson, R.C. (1972). Learning concepts from definitions. American Journal of Educational Research, 7, 385-390.

Berk, R.A. (1984, March). The use of student achievement test scores as criteria for allocation of teacher merit pay. Invited paper presented at the 1984 National Conference on Merit Pay for Teachers, Sarasota, FL.

Brophy, J. (1981). Teacher praise: A functional analysis. Review of Educational Research, 51(1), 5-32.

Cornett, L. (1986). 1986 - Incentive programs for teachers and administrators: How are they doing? Career Ladder Clearinghouse, Southern Regional Educational Board, Atlanta, GA.

Doyle, W. (1985). Classroom organization and management. In M. Wittrock (Ed.)., Third handbook of research on teaching. New York, NY: MacMillan.

Good, T.L., & Beckerman, T. (1978). Time on task: A naturalistic study in sixth grade classrooms. Elementary School Journal, 73, 193-201.

Haertel, E. (1986). The valid use of student performance measures for teacher evaluation. Educational Evaluation and Policy Analysis, 8 (1), 45-60.

Helmstadter, G., Moore, N., & Noggle, N. (1987, November). Regression analysis applied to the development and use of an index of teacher effectiveness. In D.L. Redfield (Panel Organizer), Perspectives on the use of student achievement data in the evaluation of teaching. American Evaluation Association Annual Meetings, Boston, MA.

Kounin, J. (1970). Discipline and group management in classrooms. New York, NY: Holt, Rinehart and Winston.

Luiten, J.L., Ames, W., & Ackerson, G. (1980). A meta-analysis of the effects of advance organizers on learning and retention, American Educational Research Journal, 17, 211-218.

Lockwood, A.L. (1978). The effects of values clarification and morals development curricula on school-age subjects: A critical review of recent research. Review of Educational Research, 48, 325-364.

Medley, D.M. (1987, May). Teacher education evaluation: Outcomes assessment. Invited paper, Tennessee Technological University's Center for Teacher Education Evaluation (CTEE) Conference, Nashville, TN.

Medley, D.M. (1977). Teacher competence and teacher effectiveness: A review of process-product research. Presentation, American Association of Colleges of Teacher Education (AACTE), Washington, DC.

Peterson, P.L., Marx, R.W., & Clark, C.M. (1978). Teacher planning, teacher behavior, and students achievement. American Educational Research Journal, 15, 417-432.

Redfield, D.L. (1987a, May). Pupil achievement: The weakest link in the evaluation chain. Invited paper, Tennessee Technological University's Center for Teacher Education Evaluation (CTEE) Conference, Nashville, TN

Redfield, D.L. (1987b, November). A comparison of the perspectives of teachers, students, parents, and principals concerning the influences of teaching on students and the use of student outcomes to evaluate teaching. Paper presentation, Mid-South Educational Research Association, Mobile, AL.

Redfield, D.L., & Craig, J.R. (1987a, October). The use of pupil outcome measures in the assessment of teaching when reliance upon standardized test scores is not defensible. In D.L. Redfield (Panel Organizer), Perspectives on the use of student achievement data in the evaluation of teaching, American Evaluation Association, Boston, MA.

Redfield, D.L., & Craig, J.R. (1987b, November). Identifying and documenting student outcomes for use in the evaluation of teachers when standardized achievement tests do not apply. In D.L. Redfield (Symposium Organizer), The hows and whys of using a variety of student achievement data to assess the effects of teaching and schooling: The state of the Mid-South Region, Mid-South Educational Research Association, Mobile, AL.

Redfield, D.L., & Rousseau, E.W. (1981). Meta-analysis of experimental research findings on teacher questioning behavior. Review of Educational Research, 51(2), 237-245.

Rosenshine, B. (1986). In P. Peterson and H. Walberg (Eds.). Research on teaching: Concepts, findings, and implications. Berkeley, CA: McCutchan.

Stiggins, R.J., Conklin, N.F., & Bridgeford, N.J. (1986). Classroom assessment: A key to effective education. Educational Measurement Issues and Practice, Summer 1986, 5-17.

Table 1

Types of Documentation Acutally Submitted

Type of Documentation	Number (%) Pertinent Goals	Number (%) of Teachers Selecting Documentation Type
*Anecdotal	33 (30%)	15 (58%)
Proportions (e.g., % correct)	23 (21%)	12 (46%)
Symbols (e.g., checkmarks, plus signs, minus signs)	12 (11%)	8 (31%)
Scores (other than standardized test scores)	19 (17%)	12 (46%)
**Work-Samples	4 (4%)	4 (15%)
Grades	4 (4%)	3 (12%)
***Standardized Test Scores	6 (5%)	2 (8%)
Office Records	1 (<1%)	1 (4%)

* Eight teachers representing 17 goals relied exclusively upon anecdotal records; the remaining seven teachers representing 16 goals additionally used other forms of documentations.

** Only two of these teachers relied exclusively upon worksamples as documentation.

*** Only two of these goals, both stemming from the same teacher, were exclusively reliant upon standardized test scores for documentation.

Table 2

Formats Assigned by Submitted Documentation

Format	Number (%) of Goals Documented Using Format	Number (%) of Teachers Using Format
*Narrative	55 (50%)	21 (81%)
Marked Calender Page(s)	2 (2%)	2 (8%)
Grade-book page(s)	20 (18%)	10 (38%)
List	13 (12%)	8 (31%)
Graphs/Histrograms	18 (16%)	8 (31%)
Tables/Charts	24 (22%)	14 (54%)
Checklists	11 (10%)	6 (23%)

* Twenty-two goals represented by ten teachers had narration as the exclusive form of documentation. However, most narratives provided a data based summary.