University of Northern Iowa

UNI ScholarWorks

Graduate Research Papers   Student Work

1988

A study of contemporary practice in teacher evaluation

Dennis LaVerne Washington
University of Northern Iowa

Let us know how access to this document benefits you

Recommended Citation

A study of contemporary practice in teacher evaluation

Abstract

With the passage of the Stull Act in California, the race to evaluate was on! Assembly Bill 293 called for the following provisions: Boards of education must have written guidelines which contain (a) standards of expected student progress, (b) assessments of certificated personnel competencies as related to the standards for each classification, (c) assessments of other duties as adjuncts to regular assignments, and (d) the establishment of procedures and techniques for ascertaining that the employee is maintaining proper control of the learning situation. (California Board of Education, 1972). This bill gave impetus to the already well developed trend toward quantifying methods of teacher evaluations. The passage of the Stull Act reflected the increased public interest in knowing even more clearly, just what was going on inside the schools. The very large sums of money being spent for education and educational programs are reason enough to expect specific accounting from those responsible for its spending.

A STUDY OF CONTEMPORARY PRACTICE

IN TEACHER EVALUATION

———————

A Research Paper

Presented to

the Department of Educational Administration

and Counseling

University of Northern Iowa

———————

In Partial Fulfillment

of the Requirements for the Degree

Masters of Arts in Education

———————

by

Dennis LaVerne Washington

August 1988

This Research Paper by: Dennis LaVerne Washington

Entitled: A study of Contemporary Practice in
            Teacher Evaluation

has been approved as meeting the research paper requirement

for the Degree of Master of Arts in Education.

Norman McCumsey

6-24-88
_____
Date Approved

(
_____
Advisor/Director of Research Paper

Robert H. Decker

6-28-88
_____
Date Approved

_____
Second Reader of Research Paper

Dale R. Jackson

6/28/88
_____
Date Received

_____
Head, Department of Educational
Administration and Counseling

With the passage of the Stull Act in California, the race to evaluate was on! Assembly Bill 293 called for the following provisions: Boards of education must have written guidelines which contain (a) standards of expected student progress, (b) assessments of certificated personnel competencies as related to the standards for each classification, (c) assessments of other duties as adjuncts to regular assignments, and (d) the establishment of procedures and techniques for ascertaining that the employee is maintaining proper control of the learning situation. (California Board of Education, 1972). This bill gave impetus to the already well developed trend toward quantifying methods of teacher evaluations. The passage of the Stull Act reflected the increased public interest in knowing even more clearly, just what was going on inside the schools. The very large sums of money being spent for education and educational programs are reason enough to expect specific accounting from those responsible for its spending.

Teachers, however, have been evaluated for effectivenesss by researchers for the last 70 years. {James Popham, 1973}. No one became really excited about

it, as there seemed to be an understanding that no one really knew what characteristics or techniques were needed to produce an effective teacher. This set of "loose guidelines" for determining teacher effectiveness allowed for great lattitude in judgement. Specifically, in California, not one teacher was dismissed on grounds of incompetence. {Popham 1973}.

What were the conditions which brought about the increased public interest and demand for accountability in the education community that led up to the Stull Act? It appears to have started during the Sputnik era of the 1950s. The public was frightened and looked to education as a means of reestablishing our scientific world leadership. Funding became generous and programs were approved quickly with the hope of "catching up" and regaining our past scientific excellence.

When vast amounts of public monies are spent, taxpayers tend to expect results and an accounting. The 1960s brought criticism of education as restless, frustrated students engaged in demonstrations which were generally unpopular. Reports of lowered scores on national tests also contributed to public disillusionment with the whole educational process. As inflation intensified, the American people increased their

questioning of educational funding and became more and
more demanding of a sharper scrutiny of education and
educators.

The critics charged that education was being run in
an inefficient manner, and suggested that business
methods could well be incorporated, including the use of
industrial evaluation systems. The business management
folks agreed that more structured systems using
industrial instruments and quantified data would
streamline education and make it more efficient.
Educators were willing and some were anxious to
substitute industrial management techniques for
traditional administrative practices. The problem
accompanying this system are already beginning to be
heard as educators deemed to be creative and autonomous
react to the seemingly stifling structure of "management
by objectives." As a result of the Stull Act a change in
focus for evaluative ratings had been implemented.
Whereas, informal procedures were used in the past,
present formal evaluative procedures are now a reality
for teachers. A serious problem resulting from this
change in focus, however, is the lack of adequate devices
or systems for evaluating teacher effectiveness. The
research to be cited, reflects the

attempt of various individuals, local organizations,
state departments, colleges and universities, and federal
departments, to respond with solutions to the subject of
teacher evaluation.

The focus of this paper then is to reflect the
present state of the art of teacher evaluation as seen
from two points of view. First, evaluation for the
purpose of improvement of instruction and second,
evaluation for use in promotion, demotion and
termination. Presently, these two approaches are not
clearly delineated, but in order to clarify the teacher
evaluation process such a delineation is essential.

What should be the primary reason for teacher
evaluation? Most studies stated "improvement of
instruction" as the basis for their evaluative
strategies. Although this is the stated basis, teachers
are skeptical of this purpose because they see
questionable evaluative ratings being used as the basis
for promotions, demotions and terminations. Poliakoff
(1973) found in her study of trends, that over half the
schools studied used evaluations for promotion or
dismissal. Also, she concluded that the line between new
and old approaches to evaluation is drawn harder in
literature than in practice.

In the remainder of this literature review, various evaluation studies will be reviewed. Studies cited were chosen to demonstrate the many and varied approaches taken by researchers. Moreover, in order to organize this vast literature, the following criteria were used: 1. Presage criteria--What are the personality characteristics of the teacher before the teaching act begins? 2. Process criteria--What are the methods of instruction, interaction patterns and verbal behaviors in the classroom? 3. Product criteria--What are the measurable student learning resulting from the instructional process?

The majority of the studies were found to be various combinations of the three categories. Some were specifically class three. None were found to deal with class 1 exclusively. The discussion will begin with studies which represent combinations of presage and process criteria. These were found to be numerous. Studies in which product criteria was the focus will be discussed next, and finally a section on peer evaluation will be included because this crossed all three criteria. Most studies overlap somewhat the stated organizational classes, ie. teacher personality characteristics found their way into many of the studies treated under presage and probably also influenced product results.

Shibata and Roberson (1969) have developed a teacher
self appraisal system which uses the video tape recorder
as a vechicle to focus on specific teacher behaviors as
directly related to instructional objectives. In this
system, the co-operator and teacher work together in an
intimate, self expressive relationship to change teacher
behaviors where indicated. Micro teaching sessions are a
part of this. The success of this program, as stated,
depends on a sincere commitment of time and money by the
board of education.

Teacher evaluation may well be the most difficult
kind of people performance evaluation for two reasons.
First, there is no way of furnishing teachers with
identical or nearly identical students so one can measure
the effects of two different methods of teaching.
Second, there is an unfavorable climate for constructive
work with teachers resulting from the way schools are
organized--teachers are left alone in the classroom.
(Ohio Commission 1972).This report also stressed the
importance of bridging the gap between administrators and
teachers and working toward a cohesiveness, with all
personnel joining in to improve student learning.

The Redfern model was similar to the Teacher Self
Appraisal System because it focused on performance
objectives which may be used as a basis for appraisal at

the culmination of the year's work. (Ohio Commission p. 12). This system did not use video tapes, but it did use a teacher-appraiser who works closely with the teacher as an ally in seeking solutions. A sensitive appraiser is essential, according to the report.

The Battelle Self Appraisal Instrument uses a very complete listing of incidents of effectives and ineffective teaching performance. (Ohio Commission p.13). These "critical incidents" are used as performance criteria against which a teacher's performance is measured. An appraiser-counselor works very closely with the teacher to identify strengths and weaknesses. This system was stated to be very cumbersome becasue the list of "critical incidents" is very comprehensive. Poliakoff (1973) found in her trend study that consistent relations between teacher performance and student learning occur when the "critical incidents" form is used.

Blommel and Babel (1972) present similar points of view in their studies which stress the industrial management model. The teaching process cannot be approached from a global point of view, but the teaching act must receive the focus. Blommel's supervisory program begins with the teacher's request for an observer to focus on specific definable technical skills. The

teacher chooses the time for the observation and the principal deals with the acts of the teacher, not the personality. Babel's system is based on "job targets," which are goals set out by the teacher at the beginning of the year and are used as a basis for the end of the year appraisal by the principal. This same system was described by Poliakoff (1973) to be used in the evaluation of school administrators. It begins with the person and his tasks, not with the technique of evaluation. A more traditional approach would begin with performance standards rather than goals for one's self, so this system focuses on the needs of the evaluatee, rather than the standards set for the job. Because of this focus, this program could be used as inservice. Babel (1972) favors a supervisor-counselor for each teacher as presented by Battelle's "critical incidents system."

Evaluation of teachers by their students has not been considered below the college level. Some writers felt students' reactions may be very valuable because they are the only people who monitor the teacher's performance continuously; however, trend studies do not emphasize participation by subordinates, generally. Stemnock (1970) calls this a "client-oriented" evaluation system; that is the subordinates do the evaluating. She

found a strong correlation between the validity of student evaluations and pupil achievement under a specific teacher.

Several studies mentioned micro-teaching as an effective way for teachers to learn very specific techniques. (Ohio Commission p. 15-19). This involves working on a small segmented task with a small group, usually fellow students. The lesson in use is video taped and rerun so a teacher can see his/her mistakes and continue to work on them until the specific skill is perfected.

The following studies present an orientation to product criteria. Paul Geisert's (1973) system puts the focus on product rather than process, and presents a plan for using a discrepancy system similar to contingency contracting used in behavior modification systems with younger students. This study suggests that principals with no training or background in evaluation have no place in evaluation schemes. This plan simply compares intents with accomplishments. Because today's public is looking at "what the system produces," rather than "how the system produces," the emphasis should be on the student product. Moreover, the teacher is provided feedback within this system which provides for better ways to improve student product.

James Pophan (1972) gives the "final alternative" to teahcer assessment--the use of teacher performance test. These would measure a teacher's ability to accomplish pre-specified instructional objectives. In this test, the teacher writes the objective, then teaches a mini-unit to a small group. The students are tested at the end. This, then, stands as the teacher's assessment.

Popham states there are fatal defects in each of the present methods of teacher evaluation. Different raters have different notions of what it is that constitutes good teaching. Observational systems do not pick up the subtleties of interaction effects and are subject to "faking." Pupil test performance compares learners, not teachers, and are insensitive to high quality instruction. (Popham 1973).

In a study concerned with perceptions which lead to ratings, Jenkins and Bausell "suspect that discrepancies in conceptions of teacher effectiveness may be at the root of the strong feelings raised by the accountability issues (1974, p. 72). This study was a survey to probe the way both teachers and administrators view the criteria upon which teachers are evaluated. Their respondents rated each of 16 criteria according to its importance in determining teacher effectiveness. Their ratings were quite similar. The criterion receiving the

highest rating was the teacher's relationship with his class. In the above study, what one does was as important as the outcomes of the doing. Student learning received only the eleventh highest rating. David Darland, Instruction and Professional Development Department Chairman of NEW was quoted by Pierce and Smith in the evaluation study: "Observation by a single rater as an exclusive approach to teacher evaluation should be grounds for bringing a charge of imcompetent administration against rater or school (1974, p. 33-40). Darland also feels "Use of student test scores and national norms as a basis for teacher evaluation should be grounds for dismissal (1974, p. 35). These same authors state that practice shows the group or team process is most effective, and the team should include a peer, administrator, parent and student. Scott Edwards (1974) suggests a rather interesting lottery system in which college teachers' names are drawn and then matched with three others for the purpose of evaluating colleagues across disciplines. This system as outlined provides for evaluator anonymity and an objective commentary on the instructor's teaching ability. Randomly chosen evaluators should be free from prejudice and all reports would eventually be read by the subject.

The following quotes from Buck's (1973) survey
study may be helpful in summing up:

The research on teaching effectiveness has
produced deplorably few results. The persistent
problem of determining what teaching leads to what
learning continues to plague administrators and
teachers alike. Years of research and experience have
left many educators with serious doubts concerning the
efficacy of teacher evaluation efforts. No
correlation appreciably different from zero was
discovered between the evaluations of the teachers on
the different rating scales and the evaluations based
on the achievement of pupils in the subject matter
areas. No teacher is more effective that another
except as someone so decides and designates. No
single specific observable teacher act has yet been
found whose frequency or percent of occurrence is
invariably and significantly correlated with student
achievement. Anderson (1954) concluded that ratings
by principals, peers, pupils, self, outside agencies
and by pupil achievement scores were all invalid as a
comprehensive measure of either general or specific
teacher ability. Stepheen Hencley (1971) sees the
intense concern for accountability as philosophically
in conflict with the child centered schools of
Silberman, Glasser and others.

The summary of the following authors seem to be appropriate in summary of evaluation. Dieter (1972) states:

Teacher evaluation has been one of the most

researched areas in the field of education,

and in spite of the fact that a variety of

approaches to study the topic have been

devised, efforts to find suitable methods

for evaluating teacher competence have been

quite disappointing.

The disagreements over evaluation concern purposes of evaluation, strategies employed, uses made from the evaluations, person or group given the responsibility of evaluating and other sub-questions.

Hopefully, the outcomes of all the frenzied activity on this subject will be the improvement of classroom instruction to the end that children will benefit from improved teaching. However, it is obvious that present methods of evaluation are far from satisfactory. Richard Richardson (1983, p. 64) concludes:

Present methods of evaluation are contradictory. They seek to motivate faculty to improve instruction, while at the same time collecting data which threatens their security and financial status. The element of threat in the evaluative process focuses on strengths in

the hope of concealing weaknesses. This situation can hardly work toward the improvement of instruction, but as long as evaluation is an administrative responsibility, faculty members will feel compelled to conceal their weaknesses rather than exposing them for correction.

From the research studies, it is apparent that the two points of view stated in the introduction are incompatible. The element of threat which is present when administrators perform the evaluation makes the improvement of instruction a superfluous point. If peers could satisfactorily do the job of evaluating, then perhaps we could get on with the task of improving instruction. It is to this end that the following proposed study is directed.

References

Anderson, H. M. "A study of Certain Criteria of Teacher
    Effectiveness," Journal of Experimental Education.
23:41-71, September 1954.

Babel, John, Jr. "Teacher Appraisal: How to Make It More
    Meaningful," Paper presented at the American
    Association of School Administrators Annual Convention,
    1972.

Blommel, George R. "Initial Evaluation of a System Approach
    to Supervision," Paper presented at the American
    Educational Research Association Annual Meeting, 1972.

Buck, James., and James F. Parsley, Jr. Compliers. The Way
    We See It: A Survey of Teacher Evaluation Policies and
    Practices Operant in the State of Washington. Seattle,
    Washington: School information and Research
    Service, 1973.

Buhl, Lance. Innovation Teaching: Issues Strategies and
    Evaluation. Cleveland: Cleveland State University
    Center for Effective Learning, June 1973.

Buhl, Lance, ed. and Sam H. Lane, ed. Improving
    Academic Performance through Evaluation of
    Teaching. Ohio Center for Affective Learning,
    Cleveland State University, June 1973.

California State Board of Education. California State Board
    of Education Guidelines for School Districts to Use in
    Developing Procedures for Evaluating Certificated
    Personnel. Sacromento, California: The   Board, 1972.

Commission on Public School Personnel Policies in Ohio.
    Teacher Evaluation to Improve Learning. Fourth Report.
    Cleveland, Ohio: The Commission. (736 National City
    Bank Building, 44114), 1972. p. 12-15.

Crandall, Curtis R., and Kenneth E. Shibata. A Guide to
    Implementing the Video Inservice Program (VIP).
    Nebraska Educational Service Unit 6, Milford. Bureau
    of Elementary and Secondary Education, (DHEW-OE),
Washington, D.C., 1969.

Darland, David. "Colleagues as Raters of Classroom
    Instruction," Journal of Higher Education, May-June
    1972. p. 33-40.

Dieter, Donn L. Teacher Evaluation: A Strategy for Science
     Teachers. Paper presented at the annual meeting of the
     North Carolina Science Teachers Association, Asheville,
     North Carolina, November 1982. p. 36.

Edwards, Scott., "A Modest Proposal for the Evaluation of
     Teaching." Liberal Education, October 1974.

Elliott, Velma L. "Peer Evaluation for Teachers? Why Not?"
     Elemenatry English, May 1984.

Ernest, Bill., Can You Eat? Can You Sleep? Can You Laugh?
     The Why and How of Evaluating Principals. Clearing
     House; v58 n7 p.290-292. March 1985.

Frye, Helen B. Development and Initial Validation of the
     Teaching Situation Reaction Test, (Elementary Form), A
     Measure of Teacher Effectiveness. Dayton, Ohio:
     Dayton University Research Institute, 1972.

Geisert, Paul. A Discrepancy Evaluation System for
     Elementary School Teachers. Laramie: University of
     Wyoming Center for Research, Service and Publication,
     1973.

Harrison, Willaim C.; Peterson, Kent D. Complexities in the
     Evaluation of Principals: The Relationship between
     Satisfaction with Evaluation Processes, Criteria, and
     Sources of Information 1987.

Harrison, Willian C.; Peterson, Kent D. Pitfalls in the
     Evaluation of Principals. Urban Review; v18 n4 p.
     221-235. 1986.

Hencley, Steephen P. "Impediments to Accountability,"
     Administrator's Notebook, 20 December, 1981.

Hess, Fritz. The Socialization of the Assistant
     Principal: From the Perspective of the Local
     School District. Education and the Urban Society; v18
     n1 p. 93-106. November 1985.

Jenkins and Bausell. "How Teachers View the Effective
     Teacher: Student Learning is Not the Top Criterion,"
     Phi Delta Kappan, April 1974. p.572.

McGee, Glenn W. The Instructional Leader as Middle Manager:
     The Principal's Role in Implementing a Technological
     Innovation. 1986.

McKenna, Bernard H. and others. Teacher Evaluation: An
     Annotated Bibliography. Washington, D.C., ERIC
     Clearing house on Teacher Education, 1971.

Murphy, Joseph; and others. The Administrative Control of
     Principals in Effective School Districts: The
     Supervision and Evaluation Functions. v18 n3 p.
     149-75. 1986.

Persons, Jane. Concerns of Teachers--Recent Research on Two
     Assessment Instruments. Paper presented at the Annual
     Meeting of the American Educational Research
     Association, Chicago, Illinois, April 1974.

Poliakoff, Lorraine L. Evaluating School Personnel Today.
     Washington, D. C. ERIC Clearinghouse on Teacher
     Education ERIC. Document ED 073 045, 1973. p.5.

Popham, W. James. "Found: A Practical Procedure to Appraise
     Teacher Achievement in the Classroom," Nation's Schools
     89:59-60, May 1973.

Redfern, George B. Techniques of Evaluation of Principals
     and Assisstant Principals: Four Case Studies. NASSP
     Bulletin; v70 n487 p. 66-74. February 1986.

Richardson, Richard C. "Can Faculty Evaluation Improve
     Instruction?" Community and Junior College Journal.
     p. 64.

Shibata, Kenneth E., and E. Wayne Roberson. Teacher Guide
     for Self Appraisal, Video Inservice Program. (ESEA,
     Title III). (OEDHEW) Washington, D. C., 1969.

Stemnock, Suzanne K. The Evaluatee Evaluates the
     Evaluator. ERS Circular No. 5. Educational
     Research Service, Washington, D. C. ERCI Document ED
     044 378, 1970.

Stephens, Pat; Valentine, Jerry. Assessing Principal
     Nonverbal Communication. Educational Research
     Quarterly; v10 n3 p. 60-68. 1986.

Valentine, Jerry W.; Bowman, Michael L. Audit of
     Principal Effectiveness: A User's Technical
     Manual. Designed and Tested for Principal
     Assessment in Elementary, Middle and Secondary Schools.
     Revised. 1986.

Valentine, Jerry W. Performance outcomes Based Principal
     Evaluation: A Summary of Procedural Considerations.
     1986.

Watson, Pat; and others. The School Makes a   Difference:
     Analysis of Teacher Perceptions of Their Principal and
     School Climate. 1985.

William, John; Campbell, Lloyd. Evaluating Principals for
     Your School District--20   Questions. NASSP Bulletin;
     v71 n501 p. 113-118. October 1987.