Reforming Teacher Evaluation: Naturalistic Alternatives

Thomas J. Sergiovanni

"Teacher evaluation is today shrouded with a false sense of scientism." Assumptions underlying the prevalent technical/rational approach appear, to this author, to be less than adequate. Suggested here are some naturalistic assumptions that may be worth trying.

Let us assume that you are at a workshop for supervisors and others with responsibility for teacher evaluation. The workshop leader asks you to write three statements, each of which expresses a personal feeling or belief you have about teacher evaluation. First, write a public statement that you would be willing to share with teachers, board members, and parents; next a confidential statement that you would be willing to share only with trusted colleagues; and finally a private statement that you would not readily share with others.

Chances are that doubt in the credibility of present practices characterizes your confidential and/or private statements. Don’t feel guilty: many others with responsibility for teacher evaluation share your doubts. Indeed, the quandary I describe turns to quagmire when one realizes that the same teacher evaluation practices that raise doubts of credibility are today being implemented with more fervor and on a wider scale than ever before. Let us assume for a moment that my hypothesis—"by-and-large supervisors and others responsible for teacher evaluation privately view the procedures as lacking in credibility"—is correct. What are the likely effects of participating in a system characterized by such doubts? I’ve noticed that the system takes on a certain artificial or mechanical quality, a routine function that becomes an end in itself.

Prevailing and Alternate Assumptions

This article reviews some of the assumptions basic to present teacher evaluation practices. It also contrasts these with alternate assumptions and practices that I believe hold promise for increasing meaning in the process of teacher
evaluation and that seem more naturally compatible with the work of educational enterprises. At present, the dominant view of teacher evaluation is characterized by a commitment to technical/rational values. These values are expressed in the form of predetermination and the scientific method. Predetermination is evidenced by establishing, before a teaching episode, outputs such as specific objectives and competency levels to be exhibited, and by otherwise specifying the rules of the game or the blueprint for evaluation.

The scientific method is evidenced by an emphasis on objective design characteristics in the evaluation process and on a primary concern for precision in measurement. Rating scales are emphasized as means to measure predetermined competencies, and effectiveness in teaching is defined as the accomplishment of predetermined intents, sticking to predetermined rules or displaying predetermined behavior.

In recent years, a number of prominent program evaluation experts have developed and begun to test alternatives to this technical/rational approach that rely far less on the scientific method and far more on the intuitions, aspirations, and capabilities of those involved at both ends of the evaluation. Theirs is a more naturalistic approach that sees value in discovering as opposed to determining and in describing as opposed to measuring. Though the primary focus of this pioneering work is on program evaluation, its underlying assumption, characteristics, and design features apply to teacher evaluation as well.

Let us contrast some key assumptions and practices (Figure 1) associated with technical/rational approaches to teacher evaluation with those associated with more naturalistic approaches.

Why is it important to describe prevailing assumptions behind teacher evaluation practices? Technologies are associated with ideologies, and the language and values of science (objectivity, rationality, reliability, and precision) have been found to be irresistible. Present classroom observation and evaluation technology is shrouded with a sense of scientism often not even found in the more legitimate sciences. This phenomenon flies in the face of what most educators down deep believe—that teaching is a far more artistic enterprise than scientific. That being the case, it may be that we have adopted a technology of teacher evaluation ill-suited to the nature of the educational enterprise. In the sections which follow, work being done to develop naturalistic alternatives to teacher evaluations more compatible with the nature of education is highlighted.

**Connoisseurship and Criticism**

It would be difficult to discuss naturalistic alternatives to present teacher evaluation practices without reference to the work of Elliot Eisner.³

"Clinical supervision can be geared to discovering and understanding rather than determining, and in that sense it has naturalistic potential."

Eisner is concerned with developing, in supervisors and teachers, the qualities and skills of appreciation, inference, disclosure, and description.

---


**MAY 1977 603**
tion. He refers to these qualities as the cultivation of educational connoisseurship and criticism. It is through the art of connoisseurship that one is able to appreciate and internalize meanings in classrooms, and through the skill of criticism that one is able to share or "disclose" this meaning to others. Eisner uses references to wine connoisseurship and art criticism as illustrative of these concepts. The art of appreciation is the tool of the connoisseur and the art of disclosure the tool of the critic.

Cross uses the example of sports commentators and writers to illustrate the combined application of connoisseurship and criticism.

Most of us are familiar with some of the techniques employed by commentators in describing and remarking on well-executed plays or potentially victorious strategies. Athletic plays executed with finesse are often seen in stop action, instant replay, slow motion, or are recounted in stirring detail on sports pages. One of the major contributions of these commentators is their great knowledge of sports, familiarizing them with possibilities so they know whether a flanker reverse, off-tackle run, screen pass, or drawplay was used or has potential for gaining yardage in a given situation, or when the bump and run, blitz, or single coverage was used or likely to prevent gain. Knowledge about educational potentials is also necessary. The potentially worthwhile tactics of teaching or those in use—the bump and runs or flanker reverses of schooling—need to be described and conveyed.4

Teachers should develop a portfolio—a file or collection of artifacts, records, and other materials designed to represent some aspect of the classroom program and teaching activities.

The commentator’s ability to render play-by-play action in a fashion that permits us to see and feel the game as he/she does depends upon a feel of intimacy with the phenomena under study not permitted by mere attention to game statistics and other objective information and a quality of disclosure more vivid than a box score.

Clinical Supervision

Many forms of clinical supervision resemble naturalistic approaches. Such forms are naturalistic when they rely heavily on developing a complete representation of a teaching episode and when they use this representation as a basis for making inference and building understanding of events. Videotaping is the most common method of representation associated with clinical supervision. Clinical supervision uses the data at hand (actually generated from the environment and activities being evaluated) rather than data that fit a preconceived rating form or a set of instrument specification and which place the teacher in a key role as generator, interpreter, and analyst of events described.5

Sometimes clinical supervisors take too seriously the need to “scientifically” and “objectively” document events. Sometimes they focus too intensely on the step-wise or work flow aspects of clinical supervisors. Sometimes they rely too heavily on predetermined objectives or on specifying detailed blueprints and plans that subsequently determine the direction of the evaluation. But clinical supervision can be geared to discovering and understanding rather than determining, and in that sense it has naturalistic potential. Alan Simon in another article in this issue of Educational Leadership describes work on developing a hybrid of clinical supervision that incorporates naturalistic features.6 (See page 580.)

5 Clinical supervision, generally associated with Morris Cogan’s book, Clinical Supervision [Boston: Houghton Mifflin Co., 1973], developed as a result of the pioneering work of Robert Anderson, Morris Cogan, Robert Goldhammer and others first at Harvard in the late 1950’s and early 1960’s and now at Pittsburgh and other locations.

Artifacts Analysis and Portfolio Development

Videotaping can provide a useful and readily accessible representation of teaching episodes and classroom activities. But because of the selective nature of lens and screen, this technique can also frame perception and evoke slanted meanings. Further, what the screen shows always represents a choice among possibilities and therefore provides an incomplete picture. And finally, some aspects of classroom life do not lend themselves very well to lens and screen and could be neglected.

Artifacts analysis and/or portfolio development, when used in conjunction with videotaping, can help provide a more complete representation of classroom life and therefore can increase meaning. These approaches, however, can stand apart from videotaping and indeed can stand apart from each other.

Imagine a classroom deserted suddenly 20 years ago by its teacher and immediately sealed. Everything else remains exactly as it was at the moment of desertion—desks, chairs, interest centers, work materials, test files, homework assignments, reading center sign-up lists, star reward charts and other “motivational” devices, bulletin boards, workbooks, student notebooks, grade books, plan books, library displays, teacher work-room arrangements, student lounge area arrangements, and so on.

Twenty years later, you arrive on the scene as an amateur anthropologist intent on learning about the culture, way of life, and meaning of this class (its goals, values, beliefs, activities, norms). As you dig through this classroom, what artifacts might you collect and how might you use these artifacts to help you learn about life in this classroom? Suppose, for example, you were interested in discovering what was important to this teacher, how this teacher viewed his/her role in contrast to that of the students, what youngsters seemed to be learning and/or enjoying, and how time was spent? In each case, what might you collect? What inferences might you make, for example, if you were to find most of the work of students to be in the form of short answer responses in workbooks or on ditto sheets, no student work displayed in the class, all student desks containing identical materials, and a teacher test file with most questions geared to the knowledge level of the taxonomy of educational objectives?

Portfolio development represents a teacher evaluation strategy similar to that of artifacts analysis but with some important differences. The intent of portfolio development is to establish a file or collection of artifacts, records, photo essays, cassettes, and other materials designed to represent some aspect of the classroom program and teaching activities. Though the materials in the portfolio should be loosely collected, and therefore suitable for rearrangement from time to time to reflect different aspects of the class, the portfolio should be designed with a sense of purpose. The teacher or teaching team being evaluated are responsible for assembling the portfolio and should do so in a fashion that highlights their perception of key issues and important concerns they wish to represent.

Like artists who prepare a portfolio of their work to reflect a point of view, teachers prepare a similar representation of their work. Together, supervisor and teacher use the collected artifacts to identify key issues, to substantiate teacher educational platform dimensions, and to identify serendipitous but worthwhile outcomes. A portfolio collection could be used for example to examine such issues as:

1. Are classroom activities compatible with the teacher’s espoused educational platform and/or that of the school?
2. Do supervisor and teacher have compatible goals?
3. Are youngsters engaging in activities that require advanced cognitive thinking or is the emphasis on lower level learning?
4. Do youngsters have an opportunity to influence classroom decisions?
5. Is the classroom program challenging all of the students regardless of academic potential or are some youngsters taught too little and others too much?

See: Patricia Scheyer and Robert Stake. “A Program’s Self-Evaluation Portfolio.” Center for Instructional Research and Curriculum Evaluation, University of Illinois at Urbana-Champaign, undated, mimeographed, for a discussion and application of this concept for program evaluation.
6. Are the youngsters assuming passive or active roles in the classroom?
7. Is the teacher working hard? That is, is there evidence of planning, care in preparation of materials, and reflective and conscientious feedback on students' work or are short cuts evident?
8. Does the teacher understand the subject matter?
9. What is the nature and character of the hidden curriculum in this class?

Though portfolio development and artifacts analysis share common features, the most notable of which is the collection of artifacts, portfolio development is the responsibility of the teacher. The teacher decides what will be represented by the portfolio and the items to be included in its collection. Together, the teacher and supervisor use this representation to identify issues for discussion and analysis.

I have argued that teacher evaluation is today shrouded with a false sense of scientism. Approaches following this thinking have simply not been effective. The proposed solution is to tighten up existing procedures, to get serious, to increase objectivity—or to otherwise emphasize even more present technical/rational procedures. What we may not realize, however, is that the defects of technical/rational views are not just in the procedures, but are inherent in underlying assumptions. One cannot promise that practices based on naturalistic assumptions will be more effective, but, if my original hypothesis is true, the assumptions suggest alternatives worth trying.