CHAPTER 10

BEYOND CLASSROOM WALLS
Indirect Measures of Teacher Competence

JEAN A. KING
Tahoe University

A common metaphor in educational writing compares a teacher to a gardener. Like gardeners working in greenhouses, teachers nurture growing things, watching carefully over their tender charges, adjusting conditions to provide for optimal growth, clipping back leaves and buds that grow in an improper direction, and so on—an idyllic image, perhaps, but a workable one.

The metaphor also provides a useful starting point for considering the use of indirect measures in the evaluation of teachers. In evaluating a gardener, the most direct and ultimate criterion is obvious: either the seedlings prosper and bloom, bringing abundance to the greenhouse and success to the gardener, or the plants wither in the soil, the greenhouse is barren, and the gardener fails at the one test that counts. The fact that the same gardener handles a roto-tiller with ease, serves as president of the local gardeners' union, and earns straight A's in courses on biodynamic farming simply does not excuse repeated failures in the plot. If the plants do not thrive, it does not matter, finally, if the tools used are clean and well cared for, if the compost added is wonderfully rich, or even if U.S. Department of Agriculture soil tests suggest that the growing conditions are the best to be found. Apart from the success of the crop, none of the variables listed above is either necessary or sufficient to determine the competence of the gardener.

Determining the competence of teachers in one sense parallels the case of the gardener. Many would say, in this era of accountability, that the one and only indicator of teaching effectiveness should be a direct measure, the ultimate growth and improvement of students. Without growth, regardless of the conditions under which the instruction occurs, the teacher is viewed a failure. But, as anyone who has taught will readily see, the teacher's case is not that simple, and, for a number of reasons, the gardener metaphor is finally inadequate to apply to the complexities of teacher evaluation. For one
thing, the goals of education are far more complicated and difficult to agree upon than those of a garden. A second weakness in the metaphor is the classic difficulty of measuring growth in an educational sense. Whereas a gardener can detect small signs of growth immediately, a teacher often cannot. Third, unlike plants, students are not unmovable beings who react passively to their surroundings, but are, rather, active participants in the process of education. A further weakness of the metaphor is that, unlike the controlled atmosphere of a greenhouse, the context of learning is often beyond the teacher's control, and forces before which the teacher is helpless affect students, classrooms, and schools in ways that impede learning.

Others can surely think of additional inadequacies in the metaphor, but one point is clear. The evaluation of teachers will be insufficient if the sole criterion of success is the final achievement of students. It is for this reason that indirect measures of teacher competence, used with an awareness of their inherent limits and deficiencies, are not only appropriate but also of value in teacher evaluation.

**VARIABLES INVOLVED IN DETERMINING TEACHER COMPETENCE**

Figure 1 presents five sets of variables that can be involved in determining teacher competence, roughly placed on a continuum from direct to indirect measurement. The product variables are the most direct measures of teaching competence because they involve actual measurement of student growth. Indeed, as was suggested earlier, the strict gardening view of teaching would hold that the final products are all that matter in such evaluations.

Next are the process variables, more proximate indicators of teaching competence, including measures of in-class behaviors and activities (see Chapters 7 and 8), as well as assessments of teacher-produced instructional materials (see Chapter 6). The traditional importance of studying process and product variables in the evaluation of teachers is evident, and until recently few would have questioned the theoretical propriety of evaluating teachers on the activities and outcomes of their classrooms. But the measurement of product and process variables, although most direct, often proves difficult in practical settings, and other variables must often play a part in the evaluation of teaching competence.

One such set of variables involves professional activities, that is, nonclassroom work that marks the teacher as a member of a school staff and of the teaching profession. As a staff member, a teacher might chaperone dances, do hall duty, hold fund drives, or work on curriculum committees. The commitment of the teacher as a professional might include participation in professional organizations at the local, state, or national level, attendance
at inservice or continuing education courses, and, especially at the college level, publishing activity.

Encompassed in the next set of variables, here labelled presage variables after Mitze (1960), are the traditional teacher characteristics variables, some of which are more directly related to teaching than others. In this early summary of research on teacher effectiveness, Mitze pointed to four categories of such variables in common use: (a) teacher personality attributes, (b) characteristics of teachers in training, (c) teacher knowledge and achievement, and (d) in-service teacher status characteristics (p. 1484).

The "preoperational" variable categories given in a more recent work by Borich (1977) include teacher personality, aptitude/achievement, attitude, and experience. Other presage variables, sometimes called "employee" variables, include teachers' credentials and popularity.

The last set of variables includes personal activities of two types: community work, for example, active involvement in the Red Cross, American Civil Liberties Union, or a church; and more purely personal pursuits, for example, cello lessons, daily running, vegetable gardening, or part-time employment. The appropriateness of these additional nonteaching measures in evaluation depends clearly on what is considered "good" and what is considered "teaching."

As McNeill and Popham (1973) have noted, indirect measures of the type described above allow the value preferences of the individual and the local community to operate. This can be limiting: in former days presage and personal variables bearing little relation to instruction, for example, teacher dress codes, expected church attendance, and rules forbidding teachers to smoke, drink, or (in the case of women) even to be married, sometimes determined who would be in the classroom. In all fairness, however, it must be noted that until quite recently direct measurement of student gain was just not feasible, and evaluation had little choice but to rely on such indirect measures.

THE LIMITATIONS OF RESEARCH ON INDIRECT MEASURES

Given the longstanding use of indirect measures in the evaluation of teachers, one would hope and assume that research has supported this practice. Unfortunately, this has not been the case. In a lengthy review of almost 200 studies on the personality and characteristics of teachers, Getzels and Jackson (1963) came to the discouraging conclusion that

Despite the critical importance of the problem and a half-century of prodigious research effort, very little is known for certain about the nature and measurement of teacher personality, or about the relation between teacher personality and teaching effectiveness. The regrettable fact is that many of the studies so far have not produced significant results (p. 574).

They pointed to three persistent obstacles facing researchers in this area: (1) the problem of defining "personality," (2) the problem of choosing appropriate instruments for measuring it, and (3) the problem of identifying and measuring the criterion (pp. 574–575). These obstacles continue to affect research on teacher characteristics and other presage variables.

The most elaborate project to date has been the Teacher Characteristics Study, a collection of approximately 100 separate studies over a period of six years, involving more than 6,000 teachers, in 1,700 schools and 450 school systems across the country (Ryan, Note 1). In the terms of Figure 1, the study worked to relate significant process and presage variables to each other. The study had three general objectives:

(1) To identify, analyze, and describe some of the patterns of teachers' classroom behavior and teachers' attitudes, viewpoints, and intellectual and emotional qualities. (2) To isolate and combine into scales significant correlates of some of the major dimensions of teacher behavior—scales which might be used in evaluating and predicting important teacher characteristics. (3) To compare the characteristics of various groups of teachers when they had been classified according to such conditions as age, experience, sex, size of school, cultural climate of the community, and the like (p. 369).

It is worth noting, however, that the correlates of teacher classroom behavior calculated for various groups (by age, sex, extent of teaching experience, marital status, etc.) are simply correlates; although they relate variables, they do not demonstrate causation. The fact that older teachers as a group differed from younger teachers in certain ways (they scored lower on all variables except "systematic and businesslike classroom behavior" and "learning-centered, traditional educational viewpoints") does not necessarily mean that getting older will lead to predictable behaviors or attitudes in teaching. More important, the relation of any of these variables to actual student growth is uncertain because, although such growth is not entirely ignored, the study places its emphasis elsewhere. Although the points it makes are made well, by examining the correlation between preage and process variables, the Teacher Characteristics Study focuses on the middle of the direct-to-indirect continuum and can, therefore, say little about the relation between teacher characteristics and student growth.

A theoretical limitation of research on indirect measures was noted over 20 years ago by Mitze (1960), who wrote that
Presage criteria, so-called here because of their origin in guessed predictions, are from a logical standpoint completely removed from the goals of education... In a sense they are pseudo criteria, for their relevance depends upon an assumed or conjectured relationship to other criteria, either process or product [p. 1484].

The assumptions implicit in research on indirect measures—professional and personal as well as presage—are presented in Figure 2. The required leaps of faith suggest why empirical research on these variables is conceptually problematic. For the presage, teacher-as-a-person variables, an assumption links a teacher's personality traits, knowledge, or attitudes to certain behaviors in the classroom. These, in turn, are assumed to facilitate students' learning. Not that all variables in this category are equally important to the final criterion; some, like knowledge of subject matter and voice control, are necessary minimally for success; as Hildebrand (1973) noted, as in bathing and doing dishes, their neglect is noticed. If, for example, a teacher lacks the requisite diction and volume to present a subject so that all can comprehend what is said, student growth in the end may be impossible, regardless of what else the teacher does. Research on this point is not appropriate because the results would be predetermined; a tightrope walker without balance is bound to fall.

Other variables, like sense of humor, are not necessary in the same way—we have all known effective teachers who made few jokes—but they somehow seem supportive of the learning process. A teacher who jokes in class may create an atmosphere more conducive to learning than one who does not joke. However, given the tenuous chain of assumptions linking this characteristic and student growth, and given what we now know about the multidimensional arena called a classroom, expectations that these variables will relate significantly to effective teaching seem unrealistic.

Figure 2 also presents the assumptions involved in researching professional and personal indirect variables. Again, the likelihood of empirical research showing that these variables affect student growth seems small. Depending on how teaching is defined, these indirect variables can be viewed in one of two ways. On the one hand, a teacher doing work outside of class may, by definition, be rated good (or bad). Some would say that participation in professional organizations makes a teacher good simply because good teachers must be part of the professional community. Others, however, would not want to condemn teachers who do nothing beyond the tasks required in their contract and would not label teachers inadequate because they worked at Woolworth's in the evenings. In either case, research would not alter such opinions.

A second way to view professional and personal variables is to assume that work of whatever type outside the classroom somehow encourages
positive changes in or provides support for certain process variables, thus leading to enhanced student growth. A teacher who develops a new curriculum may implement the change and somehow convey newfound enthusiasm to students. Or, by having a life outside the classroom, a teacher who has a hobby—water-skiing, swimming, or collecting helmets—may avoid the burnout that other teachers experience. But even assuming that it is true theoretically, to demonstrate this chain of events empirically would be monumental.

Upon reflection, no one should be surprised that research to date has been inconclusive. As Ryan's (Note 1) suggested and Medley (1973) reminded us, there is no single set of skills, attitudes, interests, and abilities that all good teachers have and that all poor teachers lack. Effective teaching means different things for different students at different times, and research that has sought to discover one set of characteristics for effective teachers has doomed itself to failure.

POSSIBLE ROLES FOR INDIRECT MEASURES

Evaluations that include indirect measures of teacher competence have three inherent weaknesses. First, some teachers reject out of hand any definition of teaching that goes beyond classroom activities or contractual requirements. The development of teacher militance makes even the suggestion of supplementary requirements for teachers a potential grievance. In addition, some characteristics, like age, cannot be changed. A second weakness is that indirect measures have not been even minimally validated by research, and their use is therefore based more on custom and intuition than on knowledge. As McNeill and Popham (1973) pointed out, indirect measures may predict "retention in a teaching position," rather than instructional effectiveness (p. 233).

Related to the second is a third weakness; even the theoretical validation of indirect measures presents some difficulties. One way to validate the use of any measure is to consider the three general attributes of criteria for teacher competency presented in McNeill and Popham (1973). The first of these is the reliability or consistency of a measure. The second is a neutral orientation, that is, the ability of a measure to be used successfully by people with a variety of different instructional viewpoints. The final attribute is its having an assignment indicator, an ability to yield "information about the types of instructional situations in which a given teacher functions best" (p. 238).

Applying these attributes to the three types of indirect variables—professional, presage, and personal—suggests that the use of these measures in teacher evaluation requires careful consideration. First, the reliability of certain measures may present a problem; although information on professional and personal variables may be reliably collected, personality variables are sometimes difficult to measure. The second attribute, a neutral orientation, presents another problem for indirect measures, although not in a usual sense. Because they do not involve observation in the classroom, indirect measures are neutral in McNeill and Popham's sense. However, because they measure nonclassroom aspects of a teacher's life, indirect measures in another sense are not neutral, but are, instead, biased toward a view of teaching that reaches far beyond a single classroom. As for yielding information about where a given teacher might function best—the third general attribute—professional and personal variables may appear useful but should be of secondary importance because they concern activities outside the classroom. As these arguments suggest, people who oppose the use of indirect measures in evaluating teacher competence have reasons to support their position.

The fact that these reasons exist, however, is not sufficient argument for ignoring or dismissing indirect measures of teacher competence. Despite the limitations given above, their use can be validated conceptually by considering two of the six additional attributes that McNeill and Popham (1973) suggested for discriminating among criterion measures. These two attributes, differentiation among teachers and adaptation to teachers' goal preferences, suggest the value of including data from indirect measures in teacher evaluation.

McNeill and Popham gave as one attribute of a good criterion measure its ability to discriminate among teachers. Indirect measures can easily differentiate among teachers, not in the way that more direct measures can, by showing whose students learn more, but rather in a preconceived, conceptual way. Given a broad definition of teaching that includes nonclassroom activities, teachers can be distinguished on the basis of what they know, what they are like, and how they choose to spend their time outside of class.

This strength of indirect measures leads to the second attribute of a good criterion measure, the adaptation to teachers' goal preferences. If teachers look at their overall abilities and activities (being, of course, careful not to ignore their effectiveness in the classroom) they may then choose areas in which to improve themselves. Faculty development will thus be encouraged, and teachers will continue to learn while they teach. Given these criteria, indirect measures have a viable role in the evaluation of teachers because they can effectively discriminate among teachers while adapting to teachers' goal preferences.

The question that then arises is how to use such measures in real life. The criteria just mentioned suggest two possible uses for indirect measures.

First, because they do not directly address instructional concerns, they are not immediately appropriate for improving instruction. They may, however, be appropriate for decisions of hiring and firing because they help to dis-
minate among teachers. In many job interviews, little beyond transcripts, recommendations, and applications is available to administrators, and hiring decisions have traditionally relied almost totally on pre-service and professional variables (see Chapter 4). The same has frequently been the case in retention decisions, though the move toward the use of more direct process and product variables in teacher evaluation may now begin to add new validity to such decisions (see Chapters 7–9). Whether retention decisions should be entirely based on process and product measures is, for now, a moot point. In any case, the use of indirect measures in such decisions must be tied to a predetermined definition of good teaching accepted by the faculty and administration for a given school and community. As Gephart (1979) has pointed out, “Evaluation of teaching rests heavily on the values in a given setting. If it is to be a productive activity, those structuring values must be made known to all people involved” (p. 3). Where such a definition of values exists—and the process of definition itself can be worthwhile—indirect measures can play an important role in evaluating teaching competence.

In addition to their use in personnel decisions, indirect measures can also play a viable part in self-evaluation and professional growth by adapting to teachers’ goal preferences. Making teachers aware of their professional profile can allow them to make conscious decisions about their role in education both inside and outside of the classroom. There are, after all, many different ways to be a competent teacher. The procedure could be as simple as completing a form similar to the outline of Figure 1, with data provided in each category. Looking at such a form, a teacher might see, for example, that he has done a great deal of professional work within the school, but nothing at all in the community. Or a teacher might realize that she has learned nothing new in her subject area since completing her most recent certification courses. Whether this makes a difference will depend on the teacher, the variable, and the school: cooperation and the possibility of and a desire to change are requirements in this process. But given these, the explicit use of indirect measures in teacher evaluation can be a positive way to allow for individual goal differences among teachers, schools, and even educational levels.

This use of indirect measures already occurs to some extent. At the college level, for example, the publish-or-perish ethic with its requisite professional duties has for years placed emphasis on professional variables, although sometimes to the exclusion of process and product variables. At the elementary and secondary levels, more often, emphasis on professional activities at the school, for example, sponsoring clubs, serving in the P.T.A., or participating in in-service workshops, has demonstrated a commitment to teaching or to working with students.

The key to the use of indirect measures in evaluation is the consent of the evaluated. Teachers must be willing to look for ways to change their professions, including what they are and what they are doing outside the class, as well as more direct measures of what life is like in their classroom. If they reject this concept, indirect measures can have no place in their evaluation.

Are teachers willing to be evaluated using indirect measures? I think so and offer three reasons to support a view that, to some, may seem overly realistic. First, teachers are accustomed to such measures. Historically their use is commonplace. Second, most teachers are more concerned with becoming better at what they do than with just getting by in their jobs. Few teachers want to be mediocre, and the possibility of continual improvement in some areas should be welcome to the teacher looking for ways to improve. Third, it seems to me that even the most militant teachers would be hesitant to be evaluated solely on process and product measures, given the current reality of evaluation in the schools. Indirect measures can represent positive input into an evaluation process that has the potential to become too cut-and-dried. Although indirect measures cannot replace more direct measures, they can provide supplementary information about a teacher’s effectiveness.

There is a wonderful educator’s fable that details the trials of a young man named Nathan who wants to become a teacher (Nagle, 1977). The wisest, most respected counselor in the land tells Nathan that to become a teacher he must first overcome three obstacles: the Sea of Children, the Mountain of Paperwork, and the Country of Duties and Commitments. Undaunted, Nathan sets out to face these challenges and, after much memorizing of names, writing of objectives, correcting of papers, and duties of myriad types, he returns wearily to the counselor to ask if he is now worthy of the title Teacher.

“Why, Nathan,” began the counselor, “you have been a teacher all along.”
Nathan protested, “But I have not stimulated any minds. I have not guided anyone down the road to knowledge. I have not had any time to teach.”

“Oh, you say you want to TEACH! I thought you said you wanted to be a TEACHER. Well, that is an entirely different story” [p. 19].

As the wise counselor knows and Nathan sadly learns, doing the work of a teacher is not always the same thing as teaching. The professional activities demanded of teachers, and even many activities performed in the classroom, may not seem to affect the growth of students directly. Evaluation schemes that look only at the most direct criteria may ignore or belittle the often exhausting demands of being a teacher. Indirect measures such as those discussed above are needed because they take these other demands into account.

This is not to say, however, that indirect measures can or should replace more direct measures. Evaluation of teaching competence must ultimately
look at the outgoing students, regardless of how teaching is defined. Even if indirect measures are used, this must never be forgotten. But the more research on teaching that is done, the more it is becoming evident that there is no single way to be a good teacher. Teaching is a multivariate process within which goodness is often relative. To return to the metaphor that began this chapter, schools are not, finally, like greenhouses or gardens. Because they are social institutions, schools need teachers who are aware of personal and professional goals outside of the classroom as well as within. The value of indirect measures in teacher evaluation is that they make as explicit as possible the several domains of activities involved in teaching. Whether they are used in personnel decisions or in self-evaluation for professional growth, indirect measures suggest an appropriately multidimensional model for teacher evaluation. Because they can provide valuable information unavailable from more direct sources, indirect measures may prove invaluable.

NOTE

1. It must be noted that at the base line are certain contractual and humane considerations. A teacher who repeatedly misses first period, who makes no attempt to prepare or present lessons, or who beats students should not be evaluated positively regardless of what else he or she is doing within or outside the school. That this will necessarily exclude some people who have much to offer is sometimes lamentable, but unavoidable.

REFERENCE NOTE


REFERENCES


