
◆

HANDBOOK OF RESEARCH ON TEACHER EDUCATION

Second Edition

A Project of the Association
of Teacher Educators

John Sikula
National University
Senior Editor

Thomas J. Buttery
East Carolina University
Editor

Edith Guyton
Georgia State University
Editor

Macmillan LIBRARY Reference USA
Simon & Schuster Macmillan
New York

Prentice Hall International
London Mexico City New Delhi Singapore Sydney Toronto

AUTHENTIC ASSESSMENT

Kip Tellez

UNIVERSITY OF HOUSTON

Tests tell us who we are when we are not quite sure.

—Hanson, *Testing Testing*

Educational connoisseurs recognize the signature that individual teachers give to their work. . . . We need to recognize the pervasive qualities of teaching as they are displayed in some form and a judgment—one that is difficult to ask—of how teaching might be enhanced.

—Eisner, *The Enlightened Eye*

The changing contexts of education, schooling, and teaching and the emphasis on reflective practice in progressive, innovative teaching pose special problems for the assessment of teaching. Few would argue that what is considered “excellent” teaching has remained a singular and static construct. Fewer still would argue that contemporary assessments of teaching that are traditional in perspective are adequate to capture the subtleties of innovative instructional practice (Grover, 1991; Smith, 1990). Those who find traditional measures lacking are searching for more contextually and personally sensitive ways to assess the complex professional practices of teachers. The purpose of this chapter is to honor the spirit of this search by examining theoretical and practical perspectives on authentic teacher assessment.

When teachers and teacher educators understand their work, they understand their role as members of a complex profession, which, in turn, promotes an examination of the deeper meanings of practice and the assumptions behind those practices. The term *authentic assessment* denotes those assessments of practice that emerge from context-sensitive understandings of pedagogical and personal principles that underpin the work of

teaching. Centrally, the proponents of authentic assessment want to engage teachers in thoughtful, self-conscious, and ongoing examinations of the important problems of their work in the situations where they work. Such examinations rest on teachers’ descriptions of their own practices. Asking the question, “What am I doing?” provides the context and perhaps the motivation needed for teachers to stand back and examine their practices. If teachers can describe what it is that they do in a day or over time—independently or with the assistance of others—they may be better positioned to think about central questions that define their work. The ebb and flow of professional life occasions the opportunity to think about another central question, “How am I doing?”

ASSESSING TEACHING

Provided that teachers have sufficient information about the “what” and “how” of their practice, they may be compelled to ask a third, equally important, question, “How can I do better?”

I thank Tom Bird (Michigan State University), Ardra Coles (Ontario Institute of Studies in Education), Pat Holland (University of Houston), and Carol Mullen (Ontario Institute of Studies in Education) for their formative reviews of this chapter. A special thanks goes to J. Gary Knowles (University of Michigan) who provided assistance in earlier drafts of this chapter. I am, however, solely responsible for any inaccuracies or shortcomings.

or "How can I enhance my practice?" Teacher educators, those who promote professional practice in education, are fundamentally concerned with the processes by which emerging professionals ask and answer these questions. From the perspective of a teacher educator, these central questions can be examined this way: The job of professional preparation is to facilitate (1) the degree and intensity with which emerging teachers ask "what" and "how" they are doing, and (2) the insights and enhanced views of practice that come from internal assessments of "how" they can do better. University teacher educators and school administrators responsible for teacher evaluation face similar predicaments in their external assessments of teachers. As a teacher educator working in a program of initial teacher preparation, the author has sought to make assessment more authentic within a program designed to prepare teachers for urban schools and acknowledges that even his most concerted efforts still fall short of the ideal. However, as an emerging discipline, authentic assessment of teaching will require a level of indeterminacy and vagary as it comes into its own.

To clarify the use of terms in this chapter, *emerging teacher* encompasses both preservice teachers who are seeking teaching certification and beginning teachers, although it is acknowledged that many experienced teachers may consider themselves as emerging, as lifelong learners, much like emerging teachers. The term *student* is reserved for pre-K-12 children and youth. *Teacher* is used as a general term for both in-service and preservice teachers.

Recently, much emphasis in the professional literature has been placed on teachers' interest in and concern with asking about how they are doing. In a broad sense, this focus is central to the work of reflective practice. The opposite of the reflective teacher is one who rarely or never considers in depth the "what" or "hows" of practice. Such a teacher functions uncritically, even within ever more complex contemporary educational, social, and political landscapes.

The assessment of teaching is problematic, riddled with tensions and challenges. On the one hand, the centrality of reflective practice in innovative teaching is more suitable to inquiry and dialog rather than to formal assessments and evaluations. On the other hand, professional practice itself is in need of reassessment, as are theories about, and stances on, assessment. In other words, one cannot be exclusively concerned with the promotion of particular professional practices for and by teachers. This chapter's focus is on assessment from the perspective of "productive diversity rather than standard uniformity" (Eisner, 1991, p. 79) within the multiple, overlapping landscapes of teachers' work and individual/personal signatures; innovative and sensitive methods/modes of practice; and documentation of the increasingly varied/complex topography of assessment in research-based theory and technique.

Even a representation of teacher assessment is itself problematic. For example, this chapter often relies on a language borrowed from a heritage of teacher evaluation that might be considered "inauthentic," primarily because the language of authentic assessment is emerging. This dilemma draws attention to the minimal work pursued, as evidenced in the educational literature, on what has recently been called authentic forms of assessment. The author proceeds with full recognition that there

exists a rich foundation of professionally enhancing teacher assessment, but the focus of this chapter (on the advice of the editors) is on the recent literature in the area. Therefore, the chapter grew into less of a formal review of literature and more of an exploration of possibilities.

The unevenness in the assessment literature draws attention to a distinction between what is relatively easy to measure but not very interesting and what is extraordinarily difficult to understand and document but highly important. Schön (1987) points out this tension in *Educating the Reflective Practitioner*:

In the varied topography of professional practice, there is a high hard ground overlooking a swamp. On the high ground, manageable problems lend themselves to solution through the application of research-based theory and technique. In the swampy lowland, messy confusing problems defy technical solution. The irony of this situation is that the problems of the high ground tend to be relatively unimportant to individuals or society at large, however great their technical interest may be, while in the swamp lie the problems of greatest human interest. (p. 3)

Schön's description, as a suggestive analogy, portrays the current state of research in teacher assessment. The high ground in traditional teacher evaluation is represented by lesson-length observations in which the evaluator arrives at the classroom with a checklist of "teaching behaviors" to be found. Although sometimes useful, such an evaluation obscures the day-to-day, encompassing, and behind-the-scenes work of a teacher and says little about his or her influence. For instance, typical and traditional assessments rarely explore and document the extent to which a teacher has created a democratic classroom where issues are raised about race, class, and gender. How many traditional assessments of teaching strive to capture the finely textured meanings and patterns of practice? Rarely does a checklist examine the educational experience of students, for example, or how teachers' core assumptions about practice are evidenced in actions, activities, and resources.

Gitlin and Smyth (1989) can be turned to for their distinction between "educative" and "dominant" views of teacher evaluation. The dominant view, they argue, serves to perpetuate the notion of teaching as a semi-profession. They cite Gitlin and Goldstein (1987) to support this stance:

These abrupt observation visits are initiated with little sense of the classroom's history and upon completion are not integrated into its ongoing history. In making these judgments, the administrator is usually armed with a summative rating scale which lists any number of desirable teaching outcomes. . . . The evaluator acts as an expert who knows the script and score and has in mind how it can be best realized. The teacher satisfies or does not satisfy the expert in varying degree. . . . The activity is essentially monologic, essentially a process of communiqués, of one way declarations about the state of things. (p. 7)

It is not a question about whether traditional, dominant constructs are theoretically sound but, rather, about the sensitivity with which educational professionals and their practices, contexts, and situations are understood. Traditional assessment works well if the teaching practices are themselves traditional. In a classroom and student context, for example, if a teacher's instruction is designed so that students remember the date of the signing of the Declaration of Independence, then traditional

multiple-choice type tests serve a useful purpose; however, if a teacher hopes for students to understand deeply the concept of democracy and how this peculiar system of government is realized in modern cultures, then most traditional assessments are likely to fail. Similarly, if a teacher educator facilitates and encourages emerging teachers' practices to exhibit use of anticipatory sets in setting up a lesson, then traditional teacher assessments may be potentially very useful; however, if the instructional goal is more slippery and elusive and focuses on such concepts as "reflective teaching" or culturally sensitive teaching, teacher behavior checklists are undeniably inadequate.

However, the difficulty in assessing teaching should not prevent teacher educators from attempting to uncover the critical issues in teaching practice. Just as the direction of instructional assessment of students has recently loosened the bonds that held it closely to multiple-choice, true-false, and other "objective" measures, so, too, assessment connected to teacher education must continue to explore and formalize authentic ways to assess teaching.

Some readers may be disappointed to discover that this chapter does not review the salient literature in the field and present the "best form" or forms of authentic assessment. Even if the literature on authentic assessment were large enough to provide a more substantial, analytical review, authentic assessment of teaching is a constantly emerging concept. As an evolving concept, authentic assessment is, perhaps most appropriately, configured as both an ideal to which teacher educators aspire and an attitude or way of thinking about professional practice rather than a "measure" of teaching performance. This review discusses some of the ways in which people are using the term *authentic assessment*, some of the ideas on which authentic assessment is based, and some of the things people are doing under the heading of authentic assessment. Readers should also note that because teacher evaluation is a value-laden activity, one cannot entirely eliminate individual bias.

Teacher Assessment in the 1980s and 1990s

The "crisis" reports on education in the early 1980s motivated changes in teacher preparation and assessments of the teaching profession (Sikula, 1990). The interest in teacher preparation and development led to increased attention to teacher assessment from professional bodies: existing teacher education-related organizations and at least two significant organizations whose aims included enhancing teaching professionalism. Well-known is the energetic effort of the Educational Testing Service (ETS) in developing statewide certification tests for both preservice and in-service teachers (Dwyer, 1993). For instance, the Praxis Series for beginning teacher assessment measures academic skills, subject-matter knowledge, and classroom performance (Educational Testing Service, 1992). With the addition of the classroom performance assessment portion of the series, it seems that ETS is poised to provide teacher evaluation services to schools.

Perhaps the most visible of the newly formed groups is the National Board for Professional Teaching Standards (NBPTS). The NBPTS's own standards, reported in the document, *Toward High and Rigorous Standards for the Teaching Profession*, were published along with the Board's perspectives about the assess-

ment of teachers (National Board for Professional Teaching Standards, 1986). Their perspectives on assessment, perhaps better interpreted as expectations, represent the position that the assessment procedures, encompassing a variety of methods, have an effect on the teacher's role, on student learning, and on the public's perception of schools and education more in general. The model proposed by the NBPTS suggests that the assessment of teachers involves two modules (Baratz-Snowden, 1993). The first centers on data collected from an actual teaching setting such as videotapes of classroom instruction and student artifacts (such as projects, student portfolios, and essays). The activities of the second module take place in an "assessment center" where teachers engage in interviews, simulations, and written tests designed to corroborate evidence gained at the school site. Exactly who would evaluate the evidence created in both modules is not specified.

Although the efforts of the NBPTS might be considered a substantial innovation in teacher assessment, they do not, in the author's view, reflect authentic assessment procedures. Both the model proposed by the NBPTS and the language used, even in the context of this representation, are reminiscent of common traditional methods and philosophies. The NBPTS's standards and certification procedures may be useful in many ways but this work does not reflect authentic assessment. However, it must be recognized that the work of the NBPTS is clearly difficult and may reflect the conceptual tension that arises when traditional perspectives on assessment are used to capture new and progressive practices. The NBPTS standards are but one example of recent attempts to renew the debate on teacher assessment. As other professional groups and practitioners explore new methods of teacher assessment, the issue of authentic assessment will receive additional attention.

The Call for Renewed Assessment in Teacher Education

Highly visible and diverse organizations, such as the NBPTS and the ETS, are currently exploring new forms of teacher assessment and evaluation. The vigorous funding and support of such groups and their efforts imply, among other things, a dissatisfaction with the ways in which teachers are presently assessed.

The call for renewed forms of teacher assessment turns on the distinction between "bureaucratic" and "professional" teacher evaluation (Darling-Hammond, 1986). As the descriptor suggests, a bureaucratic view of teacher assessment assumes that the teacher's work is highly rule-governed and prescribed and is tantamount to ensuring that personnel (i.e., teachers) perform tasks supported by the larger organizational structure. However, a professional view of teacher evaluation suggests that teachers are thoughtful about their practices and that the contours of their profession necessitate that they frequently modify those practices. It also assumes that teachers engage in professional practice in spite of the bureaucratic routinization and constraints of their work.

Mumane and colleagues (Mumane, Singer, Willet, Kemple, & Olsen, 1991) have broadened the audience of those interested in teacher preparation and assessment, especially in

the area of certification and licensing. As a policy analyst, Mumane has argued that traditional tests of licensure such as the National Teacher Exam (NTE), with its core battery and professional knowledge sections, fails to discriminate between levels of teaching skill and eliminates a disproportionate number of minority prospective teachers from the pool. He also argued that performance-based licensing should supplant traditional licensing examinations.

For many years, teacher educators have acknowledged that standardized tests or traditional assessments of teachers have little content validity with respect to actual teaching practice (Grover, 1991). Therefore, what is surprising for some is not the call for reform of teacher assessments but the intensity of involvement of the manifold policy groups in asserting their interests. Recall that Mumane is not a teacher educator but a policy analyst. The licensure and maintenance of licensure of teachers impacts a wide range of professional and policy groups that have vested political, membership, and financial interests. Attention paid since the mid-1980s to teacher evaluation appears more vigorous than in earlier years.

Also critical in the renewed attention for teacher evaluation is who is participating in the reform agenda. For instance, to what extent have teacher associations become involved in the process? This feature is critical at least for in-service teacher evaluation in strongly unionized states. Peterson and Chenoweth (1992) have demonstrated that teachers must have some control over the evaluation development and process if they are to consider an assessment system both valid and useful and, more important, meaningful and valuable.

Defining "Authentic"

Although the term *authentic assessment* is emerging in the educational literature and in the field, not all who use the term agree on its meaning, nor do those who agree on its meaning necessarily practice a similar form of authentic assessment. The strength of authentic assessment is revealed in situations involving sensitivity to complex contexts. Another strength rests in its potential for acknowledging and exploring relational qualities. The very lack of consensus of meanings attributed by professionals may, in fact, provide evidence of a contextual responsiveness at work. In spite of the confusion surrounding the term, teacher educators and others seem to be in agreement that if the type of assessments typically in use are not authentic, then new ones are required.

It is important to point out that providing definitions of *authentic* is not the primary purpose of this chapter. However, in order to discuss what teacher educators and others are calling authentic, a common but provisional understanding must be created. What might authentic mean with respect to assessment and evaluation? Authenticity implies that an assessment is "genuine," "real," "uncompromised," "natural," or "meaningful." Although these descriptors clearly outrun common dictionary definitions of authentic, they are appropriate when referring to this alternative form of assessment.

Assessments are authentic according to the degree to which they are meaningful to and helpful for teachers in the exploration of their practices. The role of the individual teacher involves negotiation of desirable methods of assessment, it is not to satisfy others involved in the assessment processes. Rather,

entering into dialogues about alternatives to existing practice gets at the heart of authentic assessment processes. Authenticity turns on teachers and their classroom practices—and the histories of those practices—as well as their own perceptions of roles, experiences, and work more in general. However, it is important to point out that authentic assessment practices should not be associated with benevolence toward the learner or, in this case, teacher. Authentic assessments may indeed be better received by teachers than less authentic forms of assessment, but that charge is not being extolled here. Teachers, for instance, may claim that authentic assessments divert their attention from matters that more directly impact students. In the next section authentic assessment in the context of teacher evaluation is more fully described.

Separating "Authentic," "Alternative," and "Traditional"

Two emerging terms represent the new category of assessment used in many educational settings: a distinction may be drawn between "authentic" and "alternative" forms of assessment and "traditional" assessments. Again, the point here is not to provide specific definitions of terms but to explore the potential use of the terms *authentic* and *alternative*. Of course, as emphasized earlier, not all educators would agree with this usage, nor would the author expect such agreement. The discussion, therefore, is to illustrate, not to define.

Authentic assessment in the context of teacher evaluation may include variations on alternative assessments that meet two criteria: (1) that teachers have a voice in how they are assessed and in creating the climate that is conducive to assessment, and (2) that the assessment is embedded in the specific contexts of teachers' work, including their perceptions of roles, experiences, and practices. As the teacher education community dialogs about the place of "authentic assessment" in theory and in practice, changes in meanings of the term will occur. The following examples further describe the use of these terms.

Alternative assessments represent any evaluative process that varies significantly from traditional forms of assessment. Alternative assessments, then, deviate from traditional assessments such as "objective" evaluations and checklists of teaching behaviors. For instance, the state of Texas, like nearly all others, requires all those seeking state certification or specialist endorsements to pass professional development tests. This battery of tests is known as the Examination for the Certification of Teachers in Texas. Before 1993 all tests were in a multiple-choice format, measuring discrete learning outcomes, representing traditional assessment forms, and a number of the tests still represent this view. The following example is taken from the Professional Development Test study guide for the reading specialist endorsement:

Which of the following factors is considered most important in top-down reading models?

- decoding
- textual input
- syntax
- prior knowledge (Study Guide 45, Reading, National Evaluation Systems, 1990, p. 23)

The form of this assessment is traditional and is also inauthentic for several reasons. It follows the common form of a multiple-choice item, measures a discrete and disintegrated learning skill, requires no experiential basis in order to answer, and has one and only one correct answer. Although not authentic as defined in this chapter, this test question and others like it may serve a useful purpose. Certainly, the multiple-choice format continues to be widely used in every type of educational setting.

The current elementary comprehensive test, however, is moving toward an alternative assessment and requires emerging teachers to read a series of cases and to respond to several questions regarding teacher thinking. Fifteen "competencies" frame these dimensions and the test items correspond to each. The following item is taken from the study guide.

Each student in Ms. Burgess's third-grade class has been working on writing a story for the past week. Ms. Burgess observes that several of her students are spending their daily writing period adding on to their stories, making them longer but not necessarily better, and doing no revision except occasional corrections of misspelled words. She wants to encourage these students to take a broader, more exploratory approach to revision—to review and evaluate their work and then reshape it according to their new insight. Which of the following teaching strategies would be most effective in achieving this goal?

(a) asking students to think about what parts of their story are most important and whether they have described these parts clearly and effectively.

(b) encouraging each student to place an appropriate limit on the length of his or her story based on the number of characters and events the student intends to include.

(c) having students brainstorm words related to the subject of the stories they are writing and decide which words might be incorporated in their work.

(d) suggesting that students begin each writing period by drawing an illustration that depicts the main story idea they wish to convey that day and then resume work on their writing. (National Evaluation Systems, 1993, p. 41)

The new form of the ExCET tests might be considered an alternative to the traditional assessment. The latter example is not "traditional" in that it attempts to test teacher knowledge in a more integrated way, but because it remains insensitive to (1) emerging teachers' prior individualized knowledge, (2) the context of preparation to teach and work, and (3) the resolution of lived problems, it is regarded as less than authentic. Its authenticity is questioned not only because it uses a multiple-choice (one correct answer) format but also because it does not consider the implicit theories held by teachers in relation to their practices. Again, this type of assessment may be useful in a certain context. The purpose of the ExCET is to screen thousands of potential teachers, making more context-sensitive assessments prohibitively expensive.

The two previous questions provided examples of assessments divorced from the actual context of the classroom. How might assessments conducted in the classroom be categorized? Sikorski, Niemiec, and Walberg (1994) provided an example of a traditional performance-based assessment: a checklist of teaching behaviors based on what the authors maintain are the best teaching practices. The instrument is divided into five sections, including "Presenting the Lesson" and "Student

Participation." The observer is asked to respond "yes," "no," or "not observed" to approximately seven statements in each of the five categories, yielding a total of 39 items. An example taken from the lesson presentation section requires the observer to ascribe a "yes," "no," or "not observed" to the statement, "Asks higher-order questions." This teaching checklist clearly represents a traditional and inauthentic assessment of teaching. The teacher being observed was not invited to participate in either the formation of the categories or the items themselves. The context is perhaps authentic, but the method of data collection discourages any kind of personal exploration. Calling this assessment "traditional" may imply that all other teacher evaluation programs before today have been equally inauthentic. This implication is not intended. Indeed, the use of a behavior checklist in the classroom could engage both the observer and the teacher in a discussion of the merits of the checklist itself. Teaching checklists have also been used as a starting point for discussing what happened in a lesson. However, claiming to have discovered best practices in teaching and vesting their qualities in a checklist seems both inauthentic and unlikely.

In summary, the changing grounds on which the various definitions of assessment rest must be emphasized. Many authors make no fundamental distinction between the concepts of alternative and authentic assessments, for example, and the resulting lack of clarity provides readers and reviewers of research reports with considerable challenges.

Recent Descriptions of Authentic Assessment

Authentic assessments represent those measures that ring as being true to the learner. Herman, Aschbacher, and Winters (1992) suggested that authentic assessment (although they used the term alternative), "requires [learners or teachers] to actively accomplish complex and significant tasks, while bringing to bear prior knowledge, recent learning, and relevant skills to solve realistic or authentic problems" (p. 2). Using this definition, which is not without its problems, it is evident that the old form of the ExCET is clearly inauthentic and the new form troublesome as well.

Zessoules and Gardner (1991) suggested that authentic assessment meets four criteria not typically associated with other assessments: (1) nurtures complex understandings, (2) develops reflection as a habit of the mind, (3) documents learners' (teachers') evolving understandings, and (4) uses assessment opportunities as a moment of learning. All of these features can be understood as representing growth in learning, a central element of and in the authentic assessment of teaching. Growth in learning and self-growth together loosely provide the conditions for educative experiences (Dewey, 1938) as those that promote both growth in general and the conditions for further growth. The metaphor of "growth as education and education as growth" might serve as the epistemological basis of authentic assessment processes. In addition, with respect to the final criteria, as noted by Zessoules and Gardner, the opportunity for teacher assessments is not often conceptualized as an occasion when teachers learn more about prospective practice in what might be described as a "teachable moment." In a similar vein, Marzano and Kendall (1991), drawing on their explorations

of the scant literature on authentic assessment, noted several features of authentic assessment, one being the notion of personal relevance as related to personal professional goals. Returning to the earlier example, and using these definitions, the Professional Development ExCET tests and checklists of teaching behaviors are determined as being inauthentic.

The Continuum of Assessment

The temptation to view assessment as a dualism may divert attention away from the fundamental issues in the evaluation of teaching. Because there are no well-accepted definitions for authentic assessment, teacher educators may be inclined to accept that some assessments are authentic simply because their form is different from traditional assessments. For instance, the use of portfolios has emerged as a form of authentic assessment, but simply using portfolios in no way promises that the assessment is authentic. Documents in individual portfolio files prepared by teachers (in this case) represent and articulate, through various meaningful media, crucial elements of their work. Portfolios, whose writers help shape the foci, therefore, are compilations of documentary evidence that illustrate teachers' practices and work activities in their complexities and in their contexts. Similarly, portfolio assessments of students' work potentially give both teachers and students control over the representation of students' learning and performance. As a worst-case scenario, it is easy to imagine a teacher's so-called portfolio containing nothing more than a series of external lesson-length observations by an observer unfamiliar with either the teacher, the classroom, or the students. The focus should not be on the labels for evaluation styles but rather on the substance of what is being evaluated. This task, however, may prove difficult to accomplish.

Humankind's penchant for thinking in dualisms has been well documented by Dewey (1938) and others. Dualisms often make discussions lively and concepts easier to reckon, but they often hide critical issues embedded in the dualism itself. As the educational research community has debated the lines between and the assumptions behind qualitative and quantitative research methodologies (Howe, 1992), new lines are being established in the debate over authentic and traditional assessments (Cizek, 1991, 1993; Shepard, 1993; Wiggins, 1991). With respect to the debate about quantitative and qualitative research, Eisner (1991) argued that the "line" between them is not unambiguous: some forms of qualitative research involve quantification, whereas some forms of quantitative inquiry make use of qualities. Perhaps it is possible that the major issues in qualitative inquiry—"generalization, objectivity, ethics, the preparation of qualitative researchers, validity, and so forth" (p. 7)—are no less shared in quantitative inquiry. Issues of assessment, then, are illuminated by the expansiveness of lines and divisions, not by the steadfastness of dualisms. Just as an educational inquiry process in and of itself is neither valid nor invalid, the success of an assessment process is determined by a number of dimensions, including the potential for improvement of practice through reflectivity; the potential for dialog (involving facilitation, negotiation, and decision-making processes) among participating parties;

and the forms of representation within the assessment climate and the assessment literature itself.

This author takes the view that external assessments are less likely to be authentic, especially when the teacher is left out of the design and implementation of the "instrument" or when the teacher does not know of or agree with evaluator's version of good teaching. Even internal assessments or some forms of self-evaluation might be inauthentic under certain circumstances. More strongly put, self-deception is recognized as an epistemological orientation to both experiencing and reconstructing the self (Crites, 1979). We can all become convinced that we are doing poorly or well in spite of evidence to the contrary. The challenge for the teacher education community is to avoid the generalizing dualisms and to appeal to the aims of the particular assessment process. An assessment process in and of itself is neither "good" nor "bad," valid or invalid. Rather, an assessment's worth and merit is ultimately determined by its actual use and by the subsequent claims made.

Thus far, a substantial discussion has not occurred about the assumptions underlying both authentic and traditional assessment processes. In effect, the teacher education community is in the process of determining whether different kinds of assessment are even possible given constraints imposed by institutional structures, systems, and processes. Whereas it is not advisable to organize the discussion within the context of the dualism between authentic/traditional or even between authentic/alternative, disagreements about whether assessments are authentic or inauthentic are not wholly undesirable. Discussions of the underlying perspectives may in fact provide opportunities for redefining terms.

Teacher evaluation at the in-service level has clearly become a political instrument at times, yet the discussion has not centered on the level of authenticity suggested by the various forms of assessment. Rather, teachers and policymakers have argued about the qualifications and credentials of the assessor, the frequency of assessments, and their stated purpose and intended use.

The Concern Over Validity

Consider several of the features Poster and Poster (1991), whose work in Great Britain represents a shift away from traditional teacher assessment, suggest are present in a well-run appraisal system:

- Integrates the individual and the organization.
- Provides the opportunity to initiate problem-solving and counseling interviews.
- Encourages self-development.
- Provides the basis for an institutional audit.
- Provides for the dissemination of career development advice.
- Gives individuals greater clarity of purpose through the provision of clear objectives, while allowing for autonomy of method.
- Helps build collective morale.
- Encourages and inspires individuals and enhances their self-esteem and self-confidence.

- Reduces alienation and removes resentment.
- Facilitates the identification of potential talent.
- Enhances the communication of organizational aims to all staff and facilitates the coordination of effort.
- Channels individual effort into organizational goals.
- Provides a mechanism whereby individual effort can be recognized even if no financial rewards can be offered.
- Provides a mechanism whereby individuals can influence the organization.

Poster and Poster admit that few appraisal systems can achieve all of these goals; however, many practicing educators typically find that the assessment and evaluation of their teaching meet few of these goals.

Authentic assessment focuses attention on the value of the experience for participants (as evidenced in the central questions that opened the chapter) and on the interpretations of practice as gathered from teachers' experiences. The questions of what constitutes the "data," who uses them, and how are of primary importance. Validity, within an authentic assessment climate, is associated with the thoughtful consideration of teachers' needs and the value, held for those involved, of processes (including decision making), documentation, and representation. Validity, within a traditional assessment context, is commonly understood in terms of whether or not the process measures what it claims to measure. However, by focusing only on the assessment itself, many educators miss an important feature of validity. As Cronbach (1971) noted, "One validates not a test, but an interpretation of data arising from a specified procedure" (p. 447). The ritual and ceremony in social sciences include techniques designed to substantiate the validity of "tests" of teaching. These procedures may establish that an assessment achieves a level of criterion validity and performs well on tests of reliability among experts, but they may not appeal to Cronbach's standard. Validity is established both in the interpretation and use of the data produced by an assessment technique. For instance, a classroom observation instrument may yield data that appear on the surface to "measure" what it claims to "measure," but validity may be compromised if the data are then used to rank teachers for the purpose of merit pay.

Authentic assessment focuses attention on the use and interpretation of information. It articulates teachers' understandings about the contexts of their experiences as well as understandings of those experiences. Again, the questions of who uses these interpretations, why, and how are of primary importance. Generally, alternative assessments have not contributed to the understandings and development of those being assessed. Just as in the public school context traditional forms of assessment have focused on ranking, sorting, and grading students rather than on explicitly promoting their development as learners, so it is in the context of teacher education and evaluation. Traditional forms of teacher evaluation likewise are less concerned with teacher development per se and more focused on ranking, sorting, and grading teachers according to reward or merit structures.

As the educational community has grown in its interest in authentic assessments for student evaluation, validation experts have rushed to caution educators on the widespread use of

such assessments (Messick, 1994). Others have claimed that gauging the validity of authentic assessments can follow validation patterns similar to those used in traditional types of assessments, although a marked interpretive stance must also be taken (Moss et al., 1992).

FOUNDATIONS OF AUTHENTIC ASSESSMENT PERSPECTIVES

Reappraising Phenomenology

Like education in general, teacher education since the 1970s has undergone a transformation. That transformation could be described as a movement away from behaviorism, with its emphasis on external evidence of learning and observable learning "objectives," to constructivism or cognitivism. Constructivism, for example, emphasizes the learner's prior knowledge as critical to the learning process, something which behaviorism ignored or downplayed considerably. More broadly, constructivism assumes that learning is private and that the evidence of learning is often hidden from the view of the onlooker.

In teacher education one might recognize this shift as one that began with competency-based teacher education (Houston & Howsam, 1972) in which the objectives for teaching were clearly defined and measurable, to a "reflective" (Zeichner & Liston, 1987) or "constructivist" (Fosnot, 1989) teacher education that suggests that teachers are responsible for building meaning into their own pedagogies. However, these shorthand descriptions represent a larger philosophical movement that anticipated and, to some extent, defined these shifts.

In the most general terms, education has moved "indoors" to where individuals' experiences are viewed as central to understanding learning and teaching. This change in educational discourse may be considered the result of the attack on positivism brought on by phenomenology, one of several schools of thought questioning the claims of the logical positivists. Phenomenology, as articulated by Husserl (1962), maintains that each individual's experience is crucial to understanding the nature of reality, crucial to understanding consciousness, and therefore central to understanding learning. Phenomenology owes much to Socrates' disillusion with the methods of science, which led him to study not the physical realities of the world but, instead, the mind and the products of the mind. Edie (1987) describes the phenomenological experience this way:

Concepts are not things or substances or forces at all; they are rather meaning or structures forged by the mind in its experience of things. "In itself" the world is neither true nor false, nor is it meaningful or valuable, it takes on meaning only in relation to a mind which orders and relates its parts, which thus institutes objects of thought and, by thinking the world, introduces into it the relationship of knowledge, of possible truth and falsity. What the mind creates is a tissue of possibilities. (p. 4)

In contrast to a strictly positivistic view of reality, phenomenologists assert that the subjective is vitally important (Husserl, 1962; Merleau-Ponty, 1962). This stance places value squarely on the self; however, it must undergo a modification or extension

within the context of teaching. Teaching is about persons-in-relation, primarily the relation between teachers and students. The professional development of teachers places weight on a much broader range of possibilities with respect to self and others. What Husserl and other phenomenologists make clear is that so-called subjective experiences are central to understanding, whether those experiences are derived alone or with others.

Grumet (1992) recognizes the value of contextual understandings, the "impact of milieu," for and on the self: "autobiographical methods are rooted in context" (p. 40). Context, in this view, is embedded in "metatheory" that honors interpretations of human experience and educational (research) endeavors. For teachers, the path of inquiry extends inward and outward, in lived practices and in implicit and explicit theories.

Because authentic assessment seeks to throw teachers back on to themselves, teacher educators with a desire to help teachers think more clearly about their world may begin with a dialog. Rather than meeting a particular teaching competency, as judged by an outside observer who, by implication, represents a world independent of a teacher's consciousness, questions can shift to the teacher. "What happened?" comes out of the central question that heads this chapter: "What am I doing?" As another example, the question, "Did you feel that you met the intents of the 'lesson?'" asked of a teacher might come out of a second and third fundamental question, "How am I doing?" and "How can I do better?" or "How can I enhance my practice?" Such questions may provide direction for dialog between teachers and observers and for teachers in their more private moments. For the phenomenologist, the question, "What am I doing?" is of central importance.

Phenomenology plays a central role for those interested in authentic assessment. Authentic assessment seeks to help "the learner" become more aware of the learning process; therefore, the phenomenological perspective emerges as genuine. Because authentic assessment seeks to involve the teacher-as-learner in the assessment process and in a control role, the process is a phenomenological one. A low-inference teaching evaluation instrument (such as a teaching "checklist") as a tool to gauge the quality of a teacher's lesson that further represents the sum of his or her teaching fails to consider the perceptions of the teacher during the observed lesson. For instance, teachers are centrally concerned with the welfare of children and youth, and, from the phenomenological outlook, such concerns must be taken into account. Indeed, the private experiences of the teacher are what have the greatest validity or, more appropriately, most value. Therefore, the process of evaluating teaching becomes one of individual exploration, but as embodied in action and in sensitivity to context. Grumet (1992) made the argument that a

phenomenology of educational experience examines the impact of acculturation on the shaping of one's cognitive lens. Existentialism recognizes culture as the given situation . . . through which the individual expresses his [sic] subjectivity, embodied in acts in the world. Awareness of self develops not in hermetic introspection, but in the response of subjectivity to objectivity. (p. 40)

Barber (1990), for example, points out the value of self-evaluation and suggests that self-assessment when used in a

threatening environment can greatly enhance professional development. Elliot (1989) recognizes the phenomenological in teacher assessment, suggesting that when teaching is considered a reflective practice within the boundaries of a professional ethic, "it constitutes a form of moral science in which teachers' self-evaluations play a central role in the development of professional knowledge" (p. 256).

Other researchers also point out indirectly that teachers' perceptions are crucial in their own evaluations of their work. In a study that directly examined teachers' perspectives on their own evaluations, Peterson and Comeaux (1990) found that an alternative evaluation procedure, in which teachers watched a videotape of their lessons and were asked to respond to series of questions about their lessons, was rated highest among several other evaluation systems (such as teacher behavior checklist) by both experienced and emerging teachers working in Florida and Wisconsin. McLaughlin and Pfeifer's (1988) statement that, "Any teacher evaluation system depends finally on the responses of those being evaluated, the teachers" (p. 4), suggests that teachers must value and respect the evaluation process and product if the assessment is to be effective. Although reliance on such evidence is not unproblematic, the essential point is that teachers must not only have a central role in the assessment of their own teaching but also must see it as a valuable and valid process.

The focus on the phenomenological perspective taken here should not suggest that all those working in the authentic assessment of teaching agree with the assumptions of phenomenology. The recent attention to poststructuralism in education may also offer a vehicle for advancing authentic teacher assessments. For instance, Delandshere and Petrosky (1994) argue that the interpretive narratives written by teachers may play an important role in the assessment of teaching. Within a poststructuralist framework, teacher narratives become stories that can then be interpreted much like works of fiction. Such a reading, they argue, allows teacher educators, for example, to read emerging teachers' personal narratives by using codes and systems of text interpretation. The poststructuralist view of teacher narratives also suggests that teachers create new knowledge (not simply describe their experiences) as they construct a narrative. These features are honored within the phenomenological perspective as well.

In addition, educators working toward authentic assessment may also base their work in cognitive science, particularly those studies focused on the nature of consciousness (Dennett, 1991). If authentic assessment is based on teachers' views of their own experience, then the study of what we know (consciousness) takes its rightful place in the discipline of teacher assessment. However, whether the study of consciousness is the domain strictly of philosophy or of science is a matter of great dispute.

Teacher educators working toward authentic assessment of teaching may rely on a range of theoretical and epistemological perspectives as they approach their work. Clearly, the innovations under way in the authentic assessment of teaching have had little time to become firmly rooted in any one theoretical viewpoint and are unlikely to do so. Phenomenology may not emerge as the primary philosophical perspective of authentic assessment, and this chapter makes no such prediction. However, phenomenology might be considered as the philosophy

at the heart of many recent theories of learning and teaching that focus on the experiences of learners (e.g., constructivism).

Also important are the types of "data" that teachers may use in creating a picture of their teaching. What evidence might a teacher or teacher educator use in assessing teaching? And how might these data be considered authentic? These questions are examined in following sections.

ROLES IN AND APPROACHES TO AUTHENTIC ASSESSMENT

Roles of the Self

Internal Information Gathering: Autobiographical Writing and Other Forms of Self-Assessment The authentic assessment of the self in professional school settings can involve autobiographical writings and explorations that place the individuals and their experiences at the center. The primary value of autobiographical writing within teacher education is "rooted in the process of coming to terms with oneself" (Knowles & Holt-Reynolds, 1991, p. 106).

Conceptualization of the autobiographical method as a vehicle for reflecting on and assessing practice in teacher education rests on the creation and development of texts about such practice. Personal or life history accounts; journals of various kinds; explorations of personal metaphors; reflective accounts of practice; professional development summaries (representing many such reflective accounts); and other formal and informal records and writings provide the basis (Holly, 1989; Knowles & Cole, in press).

These multiple text approaches to self-reflexivity aim to enhance emerging teachers' self-understandings, as well as their thinking about teaching and learning. Personal history accounts enable writers to understand their present inclinations to practice in light of meanings associated with earlier experiences. For example, such accounts may include elements of artistic expressions of various kinds that focus on the primacy of stories about early learning experiences in educational settings. These accounts are also likely to focus on and illuminate the implicit theories, values, and beliefs that underpin emerging teachers' orientations to becoming education professionals and the subsequent development of their practices. Within such mental landscapes of emerging ideas about practice the notion of authenticity can be established because such personal documents contain representations of formative experiences in relation to more immediate "professional" thinking.

Autobiographical writers generate stories of experience and accounts of practice that, in turn, can become the basis for continuing conversations with others about the process of becoming a teacher (Clandinin, Davies, Hogan, & Kennard, 1993; Knowles, 1993). Sensitive teacher educators can, moreover, facilitate emerging teachers' reflexivity through engaging in ongoing conversations and writing about their practices. This kind of personal exploration in relation with others is the hallmark of the principles of experiential education (Dewey, 1938). Such explorations into the inquiring self do not have to take place in isolation.

One of the roles of university-based teacher educators, for example, may be to facilitate the professional development of preservice teachers by facilitating collaboration with their peers. Among the many ways in which preservice teachers might work together is sharing autobiographical writing and developing collective accounts that arise about early experiences in schools, teaching together, or out of other joint or group work associated with being teachers in preparation. Emerging teachers' "horizontal evaluations" (Gitlin & Smyth, 1989) of each other's teaching practices, for example, could take place within several contexts and could encourage collegial collaboration (Knowles & Cole, with Presswood, 1994).

In keeping with the proposed definition of authentic, forms of self-assessment are not viewed as unproblematic, perhaps partly because of the freedom of writers to construct less accurate and trustworthy, and even reliable, accounts of their thinking and life. Crites (1979) reminded us that "experience is an imaginative construction" and that "in our experiencing we employ the same imaginative forms that appear, highly refined, in artistic expression" (p. 107). Assuming that the constructive process is integral in experiencing, it follows, Crites claims, that self-delusion is rooted in this very process. However, goals for critical self-inquiry are not intended to be clearly defined and "measurable," nor is a single version of pedagogy to be encouraged (Eisner, 1991). From this perspective, individuals' interpretations of experience are considered central to authentic assessments. It is the autobiographical presence itself in writing that is at the heart of authenticity, the search for justification, and the assessment of teaching practices. Some authors (Connelly & Clandinin, 1991; Eisner, 1991; Geertz, 1988) consider the question of signature as a significant criterion of authenticity in a writer's identity. As Eisner (1991) writes:

each person's history, and hence world, is unlike anyone else's. This means that the way in which we see and respond to a situation, and how we interpret what we see, will bear our own signature. This unique signature is not a liability but a way of providing individual insight into a situation. (p. 34)

When teachers ask "what" and "how" questions about their practice, they might desire to know more about how they can personally go about both engaging and constructing versions of their own pedagogies and professional development. At such points they are at the doors of self-inquiry, constructivism, and phenomenological reflexivity. Their signatures and voices must be validated in order for the process of authentic teacher assessment to be handled sensitively, ethically, and meaningfully. In a broader sense, even this criterion of authentic assessment must be continually reviewed and revisited by preservice teachers and teacher educators alike. Authentic assessments, then, could be considered as those that emerging teachers experience as genuine and that actually rescript their thinking about practice.

In the process of assessing their practices, emerging teachers might also learn more about the theoretical views of experience. Such perspectives may validate the efforts of inquiring selves to make sense of experiences. University teacher educators might also benefit from engaging emerging teachers' narratives; the narratives can provide a window into the context of preservice teacher education itself and its relevance for personal

inquiry and formal research (Knowles, 1993; Knowles & Cole, with Presswood, 1994). Finally, emerging teachers might also learn more about the primacy of stories as a vehicle for both understanding and constructing authentic practice.

External Information Gathering

Persons associated with teaching and learning as an inquiry-based phenomenon gather external information to inform their thinking about contexts and practices. This process is implemented in various educational environments and by various professionals. Observations, interviews, and artifacts are examples of external information-gathering techniques that typically provide opportunities for the development of sensitivities to classroom environments and understandings about teaching practices. Like internal forms of information gathering or assessment, external approaches to exploring teaching can promote personal and professional understanding of the central questions, "What am I doing?," "How am I doing?," and "How can I do better?"

External forms of information gathering and assessment, like internal forms of assessment, involve a similar set of processes. Whether information is being generated or collected about the environment; its contexts, processes, events, and people; or about the self, both demand a sensitivity to how moments, events, and circumstances can be heard and seen, documented and understood, and then finally visited and revisited. Some qualitative researchers, for example, struggle with the distinction between the "internal" and the "external" as sources of experience and information. Traditional practices of doing ethnographies promote distance, models, fixed and stable realities, and detachable conclusions (Rose, 1990), but for teachers involved in both the internal and external representation of practices, this difficulty may be less evident.

Alternative ways of assessing teaching must be rendered credible through increasingly introspective methods of engagement and analysis. Like ethnographers, reflexive teachers struggle to bring the self and the environment into intelligible relationships. Geertz (1988), for example, recommended journal writing as a channel for ensuring that ethnographic accounts are both reliable and personal. Rose (1990), another ethnographer, experimented "until [he] . . . broke with the old categories and inaugurated a new narrative responsiveness to changing world cultural relations" (p. 15). Such progressive forms of ethnography can be appreciated in relation to work being produced within the teacher education community. Contributors to this autoethnographic approach include Diamond (1992), Middleton (1993), and Mullen (1994). To continue this tradition into the arena of classroom work is not altogether a new suggestion. For example, Gitlin et al. (1992) suggested that explorations of both personal and institutional histories are prerequisites for "educational research" activities, the bases for initiating enhanced practices in schools.

The internal and external modes of practice that allow teachers to construe understandings of their work as a basis for assessing teaching have been highlighted. Potentially meaningful information-gathering approaches, such as participant observation, interviewing, or collecting documents or artifacts, and analytical practices such as ethnography, are at the

heart of authentic documentation of teachers' work and inquiries. The value of these information-gathering and -documenting processes is rooted in their sensitivity to the central questions of assessment that arise out of inquiry into educational contexts.

Changes in Participants' Roles

The incorporation of alternative assessments in preservice teacher education will no doubt change the way in which the various participants carry out their respective roles. For instance, the questions posed at the beginning of this chapter must become a habit to teachers thinking about their practice. These same questions presume that teacher educators facilitate the development of reflective inquiry-oriented teachers with habits, attitudes, and practices of self-assessment and development. Authentic assessments take into account information about the teacher in relation to students, colleagues, school- and university-based teacher educators, and others. In inquiry-oriented preservice teacher education (see, e.g., Clandinin et al., 1993; Knowles & Cole, in press; Knowles & Cole, with Presswood, 1994; Zeichner & Liston, 1987), preservice teachers learn to engage in inquiry into and to assess the multiplicities of roles, contexts, and relationships that define their emerging practice. As an example, portfolio assessment of preservice teachers may invite school-based teacher educators to become the preservice teachers' advocates. A teacher educator who assists in the development of an emerging teacher's portfolio functions to best represent abilities and achievements of the new teacher. Instead of the gatekeeper function typically held, the school-based teacher educator emerges as a portfolio advisor. Such school-based teacher educators prepared for such activity would likely see their role differently. But, even more so, portfolio development places greater responsibility in the hands of emerging teachers who have to make difficult decisions about the form and focus of their public representations.

The Role of University-Based Teacher Educators

The authentic assessment of emerging teachers suggests new roles for university-based teacher educators. Quite possibly their most important function is to work toward understanding field experiences from the perspectives of preservice teachers. Much understanding can come about from engaging with written and spoken narrative accounts of preservice teachers' experiences (Clandinin et al., 1993; Knowles & Cole, with Presswood, 1994). The writing on autobiographical and other forms of self-assessment advocated in this chapter reflect such a perspective.

A final issue of competing importance is the evidence that university-based teacher educators are inquiring into and formalizing their own developing perspectives as teachers/researchers (e.g., Cole & Knowles, in press; Diamond, 1992; Hunt, 1987; Knowles, 1992; Middleton, 1993). For example, as an experienced outdoors educator Knowles (1992) wrote about a peak teaching experience involving students from a New Zealand high school. After canoeing to a glow worm grotto that inspired students to marvel at the "power of beauty and

... the organization and design of Nature's panoramic night display" (p. 7), Knowles was silent during his "greatest teaching moment," one that offered an opportunity for personal assessment and professional writing embedded in the site of a special event. Bullough (1994) provided another example. As a teacher educator grappling with aspects of curriculum development and his own pedagogy, he explores the meanings of particular teaching experiences through the exploration of aspects of his personal history. Similarly, Rafferty (1994) traces the development of her particular approach to portfolio development in teacher education.

The Role of the School-Based Teacher Educators

Processes associated with the authentic assessment of emerging teachers suggest a new role for cooperating teachers and other school-based teacher educators. In considering teacher involvement in authentic assessment, Herman et al. (1992) suggest that new instructional and other roles for both teacher educators and emerging teachers need to be discovered. In the typical student teaching arrangement, for example, school-based teacher educators serve an important evaluative role, although several studies show that college supervisors provide emerging teachers with more substantial evaluative feedback (Guyton & McIntyre, 1990). However, emerging teachers tend to place more value on their cooperating teachers' evaluations than on those of their college supervisors (Yates, 1981). Typically, school-based teacher educators are asked to provide one or two formal evaluations of the emerging teacher. The format of these evaluations is generally provided and is sometimes prescribed by the teacher preparation institution; however, the many informal but formative evaluations of emerging teachers by the school-based teacher educators are intended to improve practice. Authentic assessment strategies may serve to capture the subtleties of these ongoing "suggestions" for improved teaching because of the expanded range of legitimate information about practice that can be potentially drawn.

Authentic assessment may also change the role of school-based teacher educators. One example of the changing role of school-based teacher educators is provided by the University of Houston's Pedagogy for Urban and Multicultural Action (PUMA) program. Student teaching "interns" present their exit portfolios to a group of educators and peers in the professional development school where their teaching experience took place. In this reflective conversation, school-based teacher educators are asked to report in what area they provided the greatest assistance. In addition, each member of the portfolio review committee reports what he or she learned from the student intern, thus blurring the lines between teacher and student. The focus of the portfolio presentation in the PUMA program is on the emerging teacher's experience; the committee offers additions and interpretations of the emerging teacher's thoughts.

In portfolio assessment stakeholders in the evaluation process (such as university faculty and a school district) can often shift their role from "gatekeeper" to advocate for those being evaluated. The role of advocate may be a proper fit for the school-based teacher educator. Instead of assisting in the gatekeeping function of evaluation, school-based teacher educators help emerging teachers develop substantial portfolios that rep-

resent their own professional development. As school districts in Pittsburgh, Houston, and Tucson, for instance, experiment with and refine their own evaluation procedures, many are making use of portfolio assessment in a variety of contexts.

The Role of Peers

Working in isolation has been a norm of school culture and of teachers' work (Hargreaves, 1990; Johnson, 1990; Rosenholtz, 1989). Increasingly, however, widespread efforts are being made to challenge the isolationist mode of teaching and working in schools by encouraging teachers to engage in collaborative or joint work. Like their experienced counterparts, preservice teachers also have traditionally been very much alone in their work:

Student teaching can be characterized somewhat as the teaching profession has been—as a lonely profession. In the same way that practicing teachers tend to be isolated from their colleagues by the organization of schools and by the ways in which schools are designed, the student teacher-cooperating teacher dyad appears to be isolated from other dyads and, indeed, from other cooperating teachers and student teachers. (Griffin, 1989, p. 362)

Preservice teacher education programs typically do little to change this pattern. Placing preservice teachers in individual classrooms and keeping their attention focused within the confines of that classroom and on the more technical aspects of teaching (on which they are typically evaluated) foster the perpetuation of norms of isolation and a relatively narrow conception of teachers' work.

If teachers are expected to change the way they think about and carry out their work and to learn from collaborative assessment of their practices, the experience of isolation and its subsequent focus on individual development within teacher preparation programs must change. Preservice programs hold particular promise for challenging traditional norms because old socialization patterns stand the chance of being interrupted by a new generation of teachers who conceptualize and carry out their work in more collaborative ways. Authentic assessment opportunities in which preservice teachers work with their peers are likely to enhance and encourage collaborative practices and careerlong professional development. Examples of collaborative self-assessment practices include developing collaboratively constructed group and individual portfolios; sharing autobiographical writing; group activities and discussions; and peer observations of practice.

The Role of Students

Whereas the teacher's self-evaluation is critical in the implementation of authentic assessment, all thoughtful teachers, at some point, turn their attention to the students' experiences in their classrooms. Some teachers may even invite students to provide their own views of the educative experience the teacher has provided. The question to be raised is, "What role does student assessment of teaching play in the authentic assessment of teaching?"

Students in colleges and universities are regularly invited to provide their own assessment of the teacher's instruction. Although the assessment of teaching in higher education is slow at moving toward alternatives, let alone authentic assessments, student evaluations of teaching may hold a promise of authenticity. However, as they are currently designed, most course and instructor evaluations resemble the teaching checklists that have been demonstrated to be quite inauthentic, in addition to other fundamental flaws. In spite of the obvious shortcomings of teaching evaluations in higher education, the results of such evaluations are often used in high stakes decisions (e.g., tenure) (Marsh, 1987). In order to create more authentic student assessments of teaching, teachers themselves must have a voice in the type of evaluation methods and materials used. For instance, typical assessments of teaching in higher education are predetermined forms that may or may not reflect the goals of an individual instructor. It seems, too, that the students providing the ratings must be given the opportunity to respond to more authentic questions about their teacher's work.

Student evaluations have been generally ignored by K-12 teachers primarily because of the perceived unreliability of students' observations of teachers and the propensity of younger children to acquiesce to those in leadership roles. Nevertheless, student evaluations of teaching effectiveness may be considered as an alternative element to traditional assessments of teaching, especially if such assessments can be integrated into a broader self-reflective evaluation scheme in which teachers view the students' assessments as a form of "internal" data.

The important issue is whether student evaluations can be considered a form of authentic assessment of teachers' work. Young children are less likely to provide "reliable" assessments of their teachers' professional work. This criticism is worthy only if we assume that student assessments remain inauthentic themselves. Typical student evaluations at the university level do not represent authentic assessments of their instructors. Students are given many opportunities over the course of a school year to critique their teachers' work in relatively free form. Such input might become well suited to improve teaching. Shor (1987) suggested that students provide teachers with critiques of provided educative experiences in a forum of open and honest communications. The blurring of the teacher/student distinction, Shor claims, is demonstrated by student critiques, essential for liberatory teaching. To include student evaluations as elements within authentic assessments of teaching is appropriate when the form of such contributions is less confining, less "traditional," and more authentic.

The Role of School Administrators

School administrators play a major role in teacher evaluation. Hickox and Musella (1992) outlined the typical performance evaluation conducted by school principals—a ritual familiar to most practicing teachers. But perhaps the greatest impact of principals' performance appraisals is felt during the first few years of teachers' careers. Typical beginning teacher evaluation procedures illustrate the critical role of school administrators (Peterson, 1990).

The practices of many emerging teachers are closely scrutinized. The average of three formal evaluations that administra-

tors—usually the school principal or assistant principal—impose on emerging teachers is stressful. By treating the assessment of emerging teachers in a sensitive manner, administrators can have a great influence on how these teachers perceive their professional selves. Thus, the organizational demands that require more intensive evaluation of emerging teachers can either enhance or disrupt initial socialization into the profession; new teachers may fear and even resent administrators' obligatory evaluations or may instead view them as an opportunity for professional growth.

Those who evaluate emerging teachers are often bound by standards imposed by school districts and state bureaucracies. A highly structured assessment system, based on bureaucratic scrutiny designed to remove incompetent teachers, can actually work against an institution's best interests. The application of "minimum" teaching standards can frustrate good teachers to the point of quitting (McLaughlin & Pfeifer, 1988). Nearly all emerging teachers believe they will be treated as professionals, yet when administrators impose strict minimum competencies, many may feel apprehensive about inquiries into their practices and official perceptions of them. The implied denial of professional status that results from imposed assessments of practice may account for why many teachers leave teaching soon after they begin (Schlechty & Vance, 1983).

EXAMPLES OF EFFORTS IN AUTHENTIC TEACHER ASSESSMENT

Self-Assessment as Narrative

Personal Knowledge in Teaching Personal knowledge is a critical component of teacher assessment. It is a powerful vehicle for enhancing learning and approaching writing as a problem-solving or thinking-through process (Knowles & Cole, with Presswood, 1994). Personal knowledge can be constructed through life history accounts and other forms of autobiographical writing. Such accounts are intended to bring forward stories of experience of learning in formal and informal settings and the meanings attributed to those experiences. Personal history accounts, for example, provide a medium for preservice teachers to access their private mental worlds and to assist their professional development.

The self, if approached as an invaluable research "instrument" (Glesne & Peshkin, 1992), is central to genuine teacher assessment. Personal history accounts, for example, sometimes draw attention to preconceptions of teaching, implicit beliefs about good teachers, appropriate learning contexts, family values, sources of inspiration that influenced the decision to teach, and more (Knowles, 1993). Such topics can be pursued as chapters in more fully developed narrative accounts of personal and professional practice.

The notion of assessment does not sit easily within formal educational studies of personal knowledge. Indeed, the concept of "authentic assessment" would probably be viewed as somehow out of whack by those whose exclusive focus is on the development of personal knowledge. The impetus for personal knowers is to address issues and phenomena of greatest human concern, not of technical interest and accuracy. Having ac-

knowledged this divergence, it is also important to clarify that positions on "assessment" do exist in the literature on personal knowledge. As indicated in the section on "Internal Information Gathering," the issue of narrative criteria is rooted in a researcher's autobiographical presence. In other words, the more visible the writer and his or her self, the more reliable the basis for assessing the research. As Connelly and Clandinin (1994) write, "a text written as if the researcher had no autobiographical presence would constitute a deception about the epistemological status of the research. Such a study lacks validity" (p. 11).

Conceptions of autobiographical presence, together with signature and voice, underscore issues of integrity *and* rigor. Efforts to develop authentic assessments strive to address these overlapping dimensions. Integrity can be thought of as that which involves the "personal participation of the knower in the knowledge he [sic] believes himself to possess," that which "takes place within a flow of passion" (Polanyi, 1962, p. 300). Like integrity, rigor can be understood as inextricably linked to one's convictions and deepest passions. Another dimension of rigor draws attention to the criteria that guide the text and by which it may be read and assessed.

Teachers as researchers of their own practices strive to clarify such matters while also liberating themselves from "objective" measures of knowledge. This intellectual tension is at the heart of the issue of narrative assessment for qualitative researchers. The process of authentic assessment for researchers can therefore be mapped according to a cyclical movement: the realization is that "we can voice our ultimate convictions only from within our convictions." The vision is to "aim at discovering what [we] truly believe in and at formulating the convictions which [we] find [ourselves] holding" (Polanyi, 1962, p. 267).

As preservice teachers prepare themselves to begin the "real" work of teaching in schools, they assert that context for learning about teaching is the classroom itself. This section was created with the intention of honoring this conviction. It was also created with the intention of honoring an authentic approach to assessment. This approach empowers teachers and researchers alike to become engaged in reflection on their own practices and narratives of observation. In turn, they can equip themselves more fully to respond to pressing concerns that grow out of the core questions, "How am I doing?" and "How can I do better?"

Portfolio Assessment

The popularity of implementing portfolio assessment in both in-service and preservice education contexts is growing, but there are very few noteworthy examples in the available research literature. Although this lack of documentation is of concern, it may simply represent the lag between design and implementation of practice and lacked opportunities for researching and research reporting (Bird, 1990).

Collins (1991) outlined the use of portfolios among experienced secondary biology teachers. This portfolio assessment, Biology Teacher Assessment Project (BioTAP), calls on teachers to document their instruction in several broad areas. In one element of the portfolio teachers were asked to provide background information, including a professional biography and a profile of the school and community setting and the internal

school environment, to provide a context for the remainder of the portfolio. Collins reports that teachers regarded this work and the portfolio content as being relatively unimportant, perhaps because those who participated in this pilot project already knew one another.

In a second element of the portfolio teachers were asked to show evidence of their planning and preparation by developing a unit of instruction. The teachers were asked to document the activities of the unit, to complete a daily lesson log, and to reflect on the implementation of the unit in specific ways. A third element invited teachers to submit evidence, including a videotape of a lesson that used either alternate materials or an innovative laboratory activity. A fourth element, designed to document teachers' assessment skills, asked that they maintain a 6-week-long journal of their responses to the various forms of evaluation. A fifth element asked teachers to submit evidence of their work in the larger educational community, recognizing that professional teachers, besides practicing in classrooms, also engage in interchanges with local school and community leaders. This element was less well received by the teachers primarily because they believed that their primary responsibility was teaching students subject matter, not serving on committees or local professional organizations, for instance. A sixth element, an "open" category, invited teachers to submit any other evidence they deemed valuable.

The elements of the BioTAP assessment that received the most positive ratings from the teachers were highly student-centered. Notwithstanding the incongruity between the research methodology and the intention of the assessment process in postproject interviews, the teachers indicated that the experience embedded in the portfolio process, although not enjoyable, was valuable for three reasons: (1) the devoted interest and concern toward their profession from "outsiders"; (2) the "face validity" of the portfolio (it looked like their teaching); and (3) the fact that the portfolio development process "impelled them to clarify their intentions and beliefs about students, about biology, and about teaching" (Collins, 1991, p. 164). The teachers reported that they would not engage in portfolio development if rewards were not forthcoming.

Miller and Tellez (1993) outlined the use of teaching portfolios for emerging teachers working in a professional development school. In such school contexts, they argued, portfolio assessment can reach its full potential. For example, at the completion of the student teaching period, emerging teachers "present" their portfolio to university- and school-based faculty members, some of whom have observed emerging teachers' practices. A peer of the emerging teacher also participates with faculty in the evaluation process. The contents of the portfolio are similar to those used in the BioTAP example. What seems most powerful in this model is the presentation of the portfolio to a committee of experts. The portfolio presentation serves both as a professional growth activity (for both preservice and in-service teachers) and as an entree into the profession. In interviews regarding the portfolio process, the emerging teachers initially regarded the process as a program requirement but partway through the process felt ownership in what they were doing, viewing the portfolio as theirs.

Bird (1990) articulated both the problems and possibilities of using portfolios of teaching. He argues for their use but notes

that research evidence of their value does not yet exist. By outlining a typology for the contents of a teacher's portfolio, Bird suggests that the contents be defined by the participants (e.g., the teacher alone, the teacher in concert with other educators, or someone outside the classroom) and by the degree of formality required of the documents (e.g., ranging from notes from a parent to diploma and licenses).

Bird's primary emphasis was on the potential collaboration with other educators, which portfolio assessment may encourage. As an assessment issue, collaboration deserves considerable attention as a vehicle for improving teacher assessment and evaluation processes generally. Bird (1990) claimed that,

Schoolteachers would spend considerable time working with their portfolios and with their colleagues to examine, refine, and share a growing stock of strategies, practices, plans, activities, and materials—a body of increasingly refined solutions to the concrete problems of school teaching. (p. 254)

King (1991) reported on portfolio assessment use in the Teacher Assessment Program (TAP) at Stanford University, as did Wolf (1991) who described the use of portfolio assessments. After considering the issues associated with their use, Wolf concludes that the primary benefit of teaching portfolios is their contextual sensitivity to teaching and consideration of the personal histories of teachers.

Many teacher educators advocate teaching portfolios for the purpose of assessment. Portfolios are accepted by both teacher education faculty and emerging teachers. But there remains an important distinction between the creation of portfolios and their evaluation. Furthermore, there is the issue that the contents of portfolios need to articulate the spirit and work of teachers' practices. Teacher educators need to be thoughtful about the essential difference between the creation of teaching portfolios themselves and related assessment issues. Few may argue about the value of a teaching portfolio as a tool for self-improvement, but when the portfolio is tied to an evaluation system used to determine certification, potential consequences may then carry high stakes.

Performance-Based Assessment

Assessment in teacher education is typically performance-based, but the implementation and articulation of performance assessment are typically very specific and highly selective. Rarely does a teacher evaluation system not rely on some form of classroom performance criteria embedded in university classrooms and courses or in field experience classrooms and schools. In their review of research on the assessment of teaching, Andrews and Barnes (1990) described six teacher programs in detail, and all but one use performance-based assessment. Over half of these evaluation programs were developed in state departments of education for use in the assessment of experienced teachers. Many university preservice programs emulate the evaluation design of local school districts or the state, aware that their students will soon be measured by such systems.

It is important to point out that performance-based assessment of teaching is not necessarily new (this point is discussed

later in this section) and that its implementation in no way guarantees authenticity. The studies explored here may indeed remind one of the distinction between alternative and authentic assessment made earlier in this chapter. Readers may discover that recent investigations into performance-based assessments reflect the alternative assessment perspective but fall short of authenticity.

Performance-based assessments may invite teachers into "assessment centers," sites removed from their own classrooms. The sites presume that the conditions under which assessment take place can be better controlled in a center. The best-known assessment center was developed at Stanford University in association with the TAP whose primary mission was the development of assessments for potential use by the NBPTS. One recent research example of performance assessment developed at the TAP, and which purported to be authentic, invited two secondary history teachers—one an experienced man, the other a beginning woman—to demonstrate their practice in three exercises: (1) evaluation of student papers, (2) use of documentary materials, and (3) textbook analysis (Wilson & Wineburg, 1993). As might be expected, the two teachers' performances differed substantially. For example, when asked to assess the "historical soundness of the text, particularly as it applied to the history of women and minorities" (p. 750), the veteran teacher claimed the sample text to be sound because it devoted attention to women and minorities. However, the younger teacher reported dissatisfaction with the sample text because it neglected important historical references to women and minorities. Wilson and Wineburg avoided comparing the teachers' performance but admitted that the younger teacher's responses fell more in line with their views on teaching and learning. Even with respect to the divergent views on the treatment of women and minorities, Wilson and Wineburg pointed out that the experienced teacher's thoughtful but not radical views are to be valued as those of a generalist. They also noted that the youthful idealism and critical orientation of the younger teacher has "showed little staying power" in schools, perhaps best predicting burnout and early exit from the profession. Are those evaluating teachers any closer to understanding and assessing teachers' complex work? The answer to this question depends largely on who is asking the question. The basis for questioning the authentic nature of this performance assessment lies in the less-than-comprehensive approach to understanding practice.

Haertel outlined a performance test that is completed outside the teachers' classrooms in a "quasi-laboratory" environment. Noting the lack of any systematic research into teacher performance/assessment activities completed outside classrooms, Haertel outlines several prototype exercises developed by the TAP (Shulman, Haertel, & Bird, 1988). Again, such assessments may reflect alternatives to typical classroom performance assessments, but do they reflect the authentic perspective?

Linn, Baker, and Dunbar (1991) noted that the call for authentic assessment is not new, pointing out that Linquist (1991) made explicit the tasks of achievement test authors: "to make the elements of his test series as nearly equivalent to, or as much like, the elements of the criterion series as consequences of efficiency, comparability, economy, and expediency will permit" (p. 52). Linquist's recommendations, made over four decades ago, suggest that in the clamor to develop achievement

tests, the educational community somehow took a wrong turn. The majority of achievement tests, including those designed to assess knowledge of teaching, bear little resemblance to the elements of the established criteria. The current attention to authentic assessment, then, can be conceived not as an entirely new direction but, rather, as the rediscovery of a once-illuminated path.

In keeping with the spirit of Linquist's work, Linn et al. (1991) suggested that to achieve adequate levels of validity (their term) authentic assessments must meet several criteria. Authentic assessments must include evidence "regarding the intended and unintended consequences, the degree to which performance on specific assessment task transfers, and the fairness of the assessments" (p. 20). They must also be explicit about the cognitive complexity required to "solve" problems posed. The meaningfulness of such assessments for teachers (and students) must be addressed. Required also is an appropriate basis for judging the content quality and comprehensives.

CONCLUSION

It is one matter to conclude that it is of fundamental importance that assessment methods and procedures be reconstrued, but it is quite another to implement authentic forms of assessment in institutional contexts. The difficulty of using such forms rests in the possibility that they may be viewed as exceedingly complex, unfeasible, unmanageable, or all three. However, a focus on the nonmeasurable, nonquantifiable aspects of teaching gives importance to the subjective and idiosyncratic elements of practice that are usually lost in institutional and bureaucratic environments. As Eisner (1991) wrote: "The cultivation of productive idiosyncrasy in the art of teaching is as important as in the art of painting" (p. 79).

This chapter argues the need for and the place of authentic assessment of teachers' practices in classrooms and in schools. Although the focus of this chapter has been on emerging teachers with respect to the value of their assessments of teaching, teachers at more advanced stages in their careers benefit from authentic explorations of their practices. Because the innovations are both promising and necessary, authentic teacher assessment has a place in the reinvention of teacher education. It is critical that the educational community think about how further research on authentic forms of assessment can be carried out in preservice teacher preparation programs in universities, schools, and classrooms. Research-based theory and technique need to become more culturally and contextually sensitive than previously conceptualized. One struggle encountered in writing this chapter grows out of the deep entrenchment of traditional concepts of assessment within discussions of authentic and alternative assessments. Nonetheless, researchers and practitioners must be sensitive to the understandings that configure integrated interpretations of teachers' work. Consequently, it is hoped that this chapter promotes conversations in the educational community.

In the third edition of the *Handbook of Research on Teaching*, Shavelson, Webb, and Burstein (1986) aligned the measurement of teaching to applications of generalizability theory, the view that sources of error in measurement can be reduced

by explicitly assessing each source of error and reducing its contribution to overall error in the measurement. One of the studies on which they drew to discuss generalizability theory assessed the effect of error introduced by different observers, occasions, and books used in a teacher's lesson. The different aspects of this study are too numerous to describe here. However, the lesson on which teachers were observed in collecting the data is illustrative. In that study the teachers were observed teaching lessons from "Books A, B, and C of the Distar Language I program" (p. 67). The measurement of teaching was whether or not the "teacher followed the Distar format in group activities and individual activities" (p. 67). After assessing teachers' instruction using this lesson format, the researchers were unable to reduce a significant portion of the error in these assessments from overall error present in assessing teaching. They concluded that more research must be conducted to improve the applications of generalizability theory to the assessment of teaching.

The intent here is not to be critical of earlier work in teacher assessment. On the contrary, all research is conducted in its own time and sensitive to existing constraints and opportunities. The type of research just described likely led to improved pedagogical skills, and whether such assessments have a role in contemporary teacher evaluation is yet undecided. The author recognizes—and hopes—that the current emphasis on authentic assessment will give way to more comprehensive assessments in the future. However, the preceding example alone illustrates the need for more authentic ways of assessing teachers. If psychometricians have been unable to measure teachers' ability to instruct using a preprogrammed, highly behavioral-based educational model such as Distar, how can teacher educators maintain that the traditional measurement methods will work when a teacher uses indeterminate instructional strategies such as reader-response instruction or readers' workshop approaches (e.g., Graves, 1983)? The pedagogical evaluation of current innovations does not respond to traditional measures—for either students or their teachers. The assessment of teaching must recognize the phenomenological nature of educational life while recognizing earlier efforts to assess the complex act known simply as teaching.

Much as Schön (1987) described the way in which easy problems can be solved using research-based theories and techniques, those aspects of teaching that can be evaluated using traditional approaches tend not to be highly important to the work of teachers. As the analogy of the swamp captures the complexities of teaching practices, so too does its richness offer opportunities for authentic assessments of teaching practices.

Far from reviewing a large corpus of research in authentic teacher assessment, this chapter examines a small but growing number of studies in teacher education. It is difficult to avoid relying on the common (and annoying) habit among social scientists: suggestions that further research must be conducted before much conclusive can be said about the phenomena under investigation. However, in this case, there may be more justification than is normally encountered.

Future work in the area of authentic assessment should engage in efforts to transform monologic, unidimensional rating scales, measurement, and teaching outcomes. In developing a

coherent, internally consistent vision of authentic assessment—an aesthetic of assessment—researchers and practitioners will need to acknowledge, describe, record, and utilize the tensions that exist between the “high ground” of theory and the “low ground” of practice, between critical, theoretical spaces and those that are contextually sensitive and relational.

REMAINING QUESTIONS

Because the innovations are promising and because I believe that authentic teacher assessment has a place in the reinvention of teacher education, I offer, for consideration, several questions about authentic assessment. Some of these questions represent suggestions for future research; others encourage critical involvement in practice:

- What are the multiple ways in which authentic assessment is cast by teacher education practitioners?
- If authentic assessment of teachers is an answer, what is the question? On what epistemological and moral ship does one embark on the authentic assessment journey?
- What are the questions that might best steer improvements in the assessment of practice?

- What are the influences of authentic teacher assessment on emerging teachers? On experienced in-service teachers? On school and university teacher educators? On school administrators? On emerging and beginning teacher peers? On others?
- Can authentic assessment processes be encouraged and developed in climates of accountability?
- How can current research methodologies (both qualitative and quantitative) help in studying the authentic assessment of teaching? Does the teacher education community need an altogether novel approach to study authentic teacher assessment?
- Can the authentic assessment of teaching realistically be fostered in highly politicized, interest group-driven, educational climates? What influences might shape the development of authentic assessment in ways other than those offered in this chapter?
- What are the ethical dimensions of authentic assessment? Who is responsible for ensuring ethical practices?
- To what extent are the underlying assumptions of authentic assessment incompatible with the structures and perspectives underlying traditionally oriented schools and classrooms? What might be the status of these competing but evolving perspectives in the future?

Addressing such questions constitutes the future of authentic assessment.

References

- Andrews, T. E., & Barnes, S. (1990). Assessment of teaching. In W. R. Houston (Ed.), *The handbook of research on teacher education* (pp. 569–598). New York: Macmillan.
- Baratz-Snowden, J. (1993). Assessment of teachers: A view from the national board for professional teaching standards. *Theory Into Practice, 32*(2), 82–85.
- Barber, L. W. (1990). Self-assessment. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 216–228). Newbury Park, CA: Sage.
- Bird, T. (1990). The schoolteacher's portfolio. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 241–256). Newbury Park, CA: Sage.
- Bullough, R. V. (1994). Personal history and teaching metaphors. A self-study of teaching as conversation. *Teacher Education Quarterly, 21*(1), 107–120.
- Cizek, G. J. (1991). Innovation or enervation?: Performance assessment in perspective. *PPhi Delta Kappan, 72*(9), 695–699.
- Clandinin, D. J., Davies, A., Hogan, P., & Kennard, B. (1993). *Learning to teach, teaching to learn*. New York: Teachers College Press.
- Collins, A. (1991). Portfolios for biology teacher assessment. *Journal of Personnel Evaluation in Education, 5*(2), 147–168.
- Connelly, F. M., & Clandinin, D. J. (1994). Narrative inquiry. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.). Oxford: Pergamon Press.
- Crites, S. (1979). The aesthetics of self-deception. *Sounding, 62*, 107–129.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 443–507). Washington, DC: American Council on Measurement.
- Darling-Hammond, L. (1986). A proposal for evaluation in the teaching profession. *Elementary School Journal, 86*(4), 531–551.
- Delandshere, G., & Petrosky, A. R. (1994). Capturing teachers' knowledge: Performance assessment a) and post-structuralist epistemology, b) from a post-structuralist perspective, c) and post-structuralism, d) none of the above. *Educational Researcher, 23*(5), 11–18.
- Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown.
- Dewey, J. (1938). *Experience and education*. New York: Macmillan.
- Diamond, C. T. P. (1992). Accounting for our accounts: Autoethnographic approaches to teacher voice and vision. *Curriculum Inquiry, 22*(1), 67–81.
- Dwyer, C. A. (1993). Teaching and diversity: Meeting the challenges for innovative teacher assessments. *Journal of Teacher Education, 44*(2), 119–129.
- Edie, J. M. (1987). *Edmund Husserl's phenomenology*. Bloomington: Indiana University Press.
- Educational Testing Service. (1992). *The Praxis series: Professional assessments for beginning teachers*. Princeton, NJ: Author.
- Eisner, E. W. (1991). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. New York: Macmillan.
- Elliot, J. (1989). Teacher evaluation and teaching as a moral science. In M. L. Holly & C. S. McLoughlin (Eds.), *Perspectives on teacher professional development* (pp. 239–258). London: Falmer.
- Fosnot, C. T. (1989). *Enquiring teachers, enquiring learners: A constructivist approach for teaching*. New York: Teachers College Press.
- Geertz, C. (1988). *Works and lives: The anthropologist as author*. Stanford, CA: Stanford University Press.
- Gitlin, A., Bringham, K., Burns, M., Cooley, V., Myers, B., Price, K., Russell, R., & Tiess, P. (1992). *Teachers' voices for school change: An introduction to educative research*. New York: Teachers College Press.
- Gitlin, A., & Goldstein, S. (1987). A dialogical approach to understanding: Horizontal evaluation. *Educational Theory, 37*(1), 17–27.

- Gitlin, A., & Smyth, J. (1989). *Teacher evaluation: Educative alternatives*. New York: Falmer.
- Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. White Plains, NY: Longman.
- Graves, D. (1983). *Writing: Teachers and children at work*. Portsmouth, NH: Heinemann.
- Griffin, G. (1989). A descriptive study of student teaching. *The Elementary School Journal*, 89(3), 343-364.
- Grover, B. W. (1991). The teacher assessment dilemma: What is versus what ought to be! *Journal of Personnel Evaluation in Education*, 5(2), 103-119.
- Grumet, M. R. (1992). Existential and phenomenological foundations of autobiographical methods. In W. S. Pinar & W. M. Reynolds (Eds.), *Understanding curriculum as phenomenological and deconstructed text* (pp. 28-43). New York: Teachers College Press.
- Guyton, E., & McIntyre, D. J. (1990). Student teaching and school experiences. In W. R. Houston (Ed.), *The handbook of research on teacher education* (pp. 514-534). New York: Macmillan.
- Hanson, F. A. (1993). *Testing testing*. Berkeley: University of California Press.
- Hargreaves, A. (1990). Cultures of teaching. In I. Goodson & S. Ball (Eds.), *Teachers' lives*. New York: Routledge.
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hickox, E. S., & Musella, D. F. (1992). Teacher performance appraisal and staff development. In M. Fullan & A. Hargreaves (Eds.), *Teacher development and educational change* (pp. 156-169). London: Falmer.
- Holly, M. L. (1989). *Writing to grow: Keeping a personal professional journal*. Portsmouth, NH: Heinemann.
- Houston, W. R., & Howsam, R. B. (1972). *Competency-based teacher education: Progress, problems, and prospects*. Chicago: Science Research Associates.
- Howe, K. R. (1992). Getting over the qualitative-quantitative debate. *American Journal of Education*, 100(2), 236-256.
- Hunt, D. (1987). *Beginning with ourselves: In theory, practice and human affairs*. Cambridge, MA: Brookline Books/OISE Press.
- Husserl, E. (1962). *Ideas: General introduction to pure phenomenology*. Translated by W. R. Boyce Gibson. New York: Collier.
- Johnson, S. M. (1990). *Teachers at work: Achieving success in our schools*. New York: Basic Books/HarperCollins Publishers.
- Kerchner, C. K., & Mitchell, D. (1983). Unionization and the shaping of teachers' work. *Teacher Education Quarterly*, 10(4), 71-88.
- King, B. (1991). Thinking about linking portfolios with assessment center exercises: Exercises from the Teacher Assessment Project. *Teacher Education Quarterly*, 18(3), 31-48.
- Knowles, J. G. (1992). Geopietry, the concept of sacred place: Reflections on an outdoor education experience. *Journal of Experiential Education*, 15(1), 6-12.
- Knowles, J. G. (1993). Life history accounts as mirrors: A practical avenue for the conceptualization of reflection in teacher education. In J. Calderhead & P. Gates (Eds.), *Conceptualizing reflection in teacher development* (pp. 70-92). London: Falmer.
- Knowles, J. G., & Cole, A. L. (In Press). Developing practice through field experience. In *A knowledge base for teacher educators*. San Francisco, CA: Jossey-Bass, Committee on the Professional Knowledge Base, American Association of Colleges for Teacher Education.
- Knowles, J. G., & Cole, A. L., with Presswood, C. S. (1994). *Through preservice teachers' eyes: Exploring field experiences through narrative and inquiry*. New York: Merrill.
- Knowles, J. G., & Holt-Reynolds, D. (1991). Shaping pedagogies through personal histories in preservice teacher education. *Teachers College Record*, 93(1), 89-113.
- Lindquist, E. F. (1991). Preliminary considerations in objective test construction. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 119-184). Washington, DC: American Council on Education.
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20(8), 15-21.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253-388.
- Marzano, R. J., & Kendall, J. S. (1991). *A model continuum of authentic tasks and their assessment*. Aurora, CO: Mid-continent Regional Educational Laboratory.
- McLaughlin, M. W., & Pfeifer, R. S. (1988). *Teacher evaluation: Improvement, accountability, and effective learning*. New York: Teachers College Press.
- Merleau-Ponty, M. (1962). *Phenomenology of perception*. New York: Humanities Press.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Middleton, S. (1993). *Educating feminists: Life histories and pedagogy*. New York: Teachers College Press.
- Miller, A., & Tellez, K. (1993). *Innovative use of teaching portfolios in a professional development school*. Paper presented at the annual meeting of the Association for Teacher Educators, Los Angeles.
- Miller, M. D., & Legg, S. M. (1993). Alternative assessment in high stakes environment. *Educational Measurement: Issues and Practice*, 12(2), 9-15.
- Moss, P. A., Beck, J. S., Ebbs, C., Matson, B., Muchmore, J., Steele, D., Taylor, C., & Henter, R. (1992). Portfolios, accountability, and an interpretive approach to validity. *Educational Measurement: Issues and Practice*, 11(3), 12-21.
- Mullen, C. (1994). *Imprisoned selves: A narrative inquiry into incarceration and education*. Unpublished doctoral dissertation, University of Toronto, Canada.
- Murphy, R. J., Singer, J., Willet, J., Kemple, J., & Olsen, R. (1991). *Who will teach? Policies that matter*. Cambridge: Harvard University Press.
- National Board for Professional Teaching Standards. (1986). *Toward high and rigorous standards for the teaching profession*. Detroit, MI: Author.
- National Evaluation Systems. (1990). *Study guide: Reading test*. Amherst, MA: Author.
- National Evaluation Systems. (1993). *Preparation manual: Elementary comprehensive test*. Amherst, MA: Author.
- Peterson, K. D. (1990). Assistance and assessment for beginning teachers. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 104-115). Newbury Park, CA: Sage.
- Peterson, K. D., & Chenoweth, T. (1992). School teachers' control and involvement in their own evaluation. *Journal of Personnel Evaluation in Education*, 6(2), 177-190.
- Peterson, P. L., & Comeaux, M. A. (1990). Evaluating the systems: Teacher's perspectives on teacher evaluation. *Educational Evaluation and Policy Analysis*, 12(1), 3-24.
- Polanyi, M. (1962). *Personal knowledge: Towards a post-critical philosophy*. Chicago: The University of Chicago Press.
- Poster, C., & Poster, D. (1991). *Teacher appraisal: A guide to training*. London: Routledge.
- Rafferty, C. D. (1994, February). *Portfolio assessment and secondary methods classes: What happens when the twain meet?* Paper presented at the annual meeting of the Association of Teacher Educators, Atlanta.
- Rose, D. (1990). *Living the ethnographic life*. Newbury Park, CA: Sage.

- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. White Plains, NY: Longman.
- Schlechty, P., & Vance, V. (1983). Recruitment, selection, and retention: The shape of the teaching force. *The Elementary School Journal*, 83(4), 469-487.
- Schön, D. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.
- Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 50-91). New York: Macmillan.
- Shepard, L. A. (1993). The place of testing reform in educational reform: A reply to Cizek. *Educational Researcher*, 22(4), 4-9.
- Shor, I. (1987). *Critical teaching and everyday life*. Chicago: The University of Chicago Press.
- Shulman, L. S., Haertel, E. H., & Bird, T. (1988). *Toward alternative assessment for teachers: A report on a work in progress*. Stanford: Teacher Assessment Project, Stanford University.
- Sikorski, M. F., Niemiec, R. P., & Walberg, H. J. (1994). Best teaching practices: A checklist for observations. *NASSP Bulletin*, 78(561), 50-54.
- Sikula, J. (1990). National commission reports of the 1980s. In W. R. Houston (Ed.), *The handbook of research on teacher education* (pp. 72-82). New York: Macmillan.
- Smith, K. E. (1990). Developmentally appropriate education or the Hunter teachers assessment model: Mutually incompatible alternatives. *Young Children*, 45(2), 12-13.
- Wilson, S. M., & Wineburg, S. S. (1993). Wrinkles in time and place: Using performance assessments to understand the knowledge of teachers. *American Education Research Journal*, 30(4), 729-770.
- Wiggins, G. (1991). A response to Cizek. *Pbi Delta Kappan*, 72(9), 700-703.
- Wolf, K. (1991). The schoolteacher's portfolio: Issues in design, implementation, and evaluation. *Pbi Delta Kappan*, 73(2), 129-136.
- Yates, J. W. (1981). Student teaching in England: Results of a recent survey. *Journal of Teacher Education*, 32(5), 44-46.
- Zeichner, K., & Liston, D. P. (1987). Teaching student teachers to reflect. *Harvard Educational Review*, 57(1), 23-48.
- Zessoules, R., & Gardner, H. (1991). Authentic assessment: Beyond the buzzword and into the classroom. In V. Perrone (Ed.), *Expanding student assessment* (pp. 47-71). Alexandria, VA: ASCD.