

Specifications for a Performance-based Assessment System for Teacher Preparation, Richard J. Stiggins, Assessment Training Institute

Dr. Stiggins is the President of the Assessment Training Institute, Portland, Oregon, an organization he created in 1992 to build "greater student motivation and achievement through effective day-to-day classroom assessment." The Institute conducts seminars, conferences and district training programs, and produces books, interactive training videos, and assessment planning packages, to help educators understand how to use sound assessment for instruction. Dr. Stiggins has coined the phrase "assessment literacy" to represent the standards of professional excellence to be attained by practicing educators. He has served as a district assessment director, director of test development for ACT, and on the faculties of Michigan State, the University of Minnesota and Lewis & Clark College in Portland.

In this paper, Dr. Stiggins notes that, even though typical teachers will spend a quarter to a third of their professional time involved in assessment-related activities, very few states explicitly require competence in assessment as a condition of teacher licensure. As a result teacher preparation programs typically neglect this facet of training. So he shifts his focal point from in-service training to the initial preparation of teachers, still maintaining his perspective that classroom assessment can be more than an index of learning—it also can be means to greater achievement for students.

Taking the NCATE performance-based unit accreditation standards adopted in 2000, he selects three as the focus for his assessment system "specifications." These are: Standard 1, *Candidate knowledge, skills and dispositions*; Standard 2, *Program assessment and unit evaluation*; and Standard 5, *Faculty performance and development*. The assessment system of the teacher preparation unit must verify candidate competence while at the same time documenting the impact of the program and the faculty on that competence. But, in addition, the unique feature of the system proposed by Dr. Stiggins is his expectation that faculties of education will weave into their courses concrete and specific ways to model for their candidates the effective use of classroom assessment as a teaching tool. In other words, they will teach sound assessment practices by demonstrating them in their own instruction.

Thus, the paper develops key ideas about the place of assessment training in teacher preparation (it is essential), the critically important roles of faculty (in modeling the effective use of classroom assessment), and of teacher candidates themselves (who become "partners" in gathering and reporting evidence of their growing competence).

Dr. Stiggins' conclusion summarizes the challenges and the benefits of this vision of assessment in teacher preparation. First, on challenges, his system expects that:

- Faculty will work as a team to agree upon a set of achievement standards for candidates; this runs counter to a tradition in which each faculty member decides upon the content priorities of her or his own courses and the means by which teacher candidates will be measured. Thus, a reward system that has

traditionally centered only on individual faculty achievement will be asked to value collaboration among faculty.

- Ongoing assessments will document candidate success with quality classroom assessments developed and used by an *assessment literate* faculty, when we know that this level of assessment literacy is lacking throughout our educational system, including in higher education.
- There will be heavy reliance on student-involved assessment, record keeping and communication processes; nothing in our assessment traditions prepares the faculty for this.
- The assessments will feed into an information management system that fosters continuous record keeping and communication about candidate progress and ultimate success; this stands in stark contrast to a traditional environment that relies on difficult to interpret grades delivered to the registrar and to a transcript every quarter or semester.
- This kind of assessment system will permit the evaluation of faculty in terms of actual impact on the achievements of their teacher candidates; traditionally, accountability for learning has not been part of our tradition.

But on the positive side, Dr. Stiggins' paper also builds a strong case for the benefits of his proposed assessment system:

- Faculty clearly articulate standards that underpin excellence in teaching; this sharpens the focus of instruction and assessment.
- Evidence of candidate competence in teaching will be of higher quality; both formative and summative assessments will be more valid and reliable.
- The faculty's assessment task becomes easier to manage because it will be limited to what is important for candidates to know and be able to do, and faculty will develop an understanding of how to select the most efficient assessments for each context. Besides, with candidate involvement, the workload of assessment will be spread across more shoulders. Candidates take lead responsibility for assembling the evidence of their preparedness to teach.
- The program environment will set candidates up to succeed and will continuously reveal to them that they are in control of that success.
- Individual faculty members' sense of efficacy will increase as candidates demonstrate regular improvement; all faculty members will see the effect of their own unique contribution to the total teacher preparation program.

SPECIFICATIONS FOR A PERFORMANCE-BASED ASSESSMENT SYSTEM FOR TEACHER PREPARATION

Richard J. Stiggins
Assessment Training Institute
50 SW 2nd Ave., Suite 300
Portland OR 97204
800-480-3060
ati@assessmentinst.com

July 5, 2000

The NCATE 2000 unit standards describe a vision of excellence in teacher preparation that include three requirements calling for systematic assessment and evaluation of candidate and faculty performance. “Standard 1, Candidate Knowledge, Skills and Dispositions,” stipulates that:

Candidates preparing to work in schools as teachers or other school personnel know the content of their fields, demonstrate professional and pedagogical knowledge, skills and dispositions and apply them so that all students learn. Assessments indicate that candidates meet professional, state, and institutional standards.

“Standard 2, Program Assessment and Unit Evaluation,” holds that:

The unit has an assessment system that collects data on the qualifications of applicants, and performance of candidates and graduates. These performance data and other information are used to evaluate and improve programs.

And “Standard 5, Faculty Performance and Development,” establishes similar expectations for teacher education faculty members:

The unit’s faculty model best professional practices in scholarship, service and teaching, including self assessment of their effectiveness and effect on candidate performance; they also collaborate with colleagues in the disciplines and schools. The unit systematically evaluates faculty performance and facilitates professional development.

These standards bring with them a new emphasis in teacher preparation programs on the collection of evidence to document candidate’s competence in teaching. That evidence will bear on very important decisions, such as the certification of candidates to enter the profession and the accreditation of programs that prepare them. Therefore, accurate assessments of student performance is essential, as is the effective use of the resulting information in the decision making process.

The purpose of this paper is to describe an integrated, instructionally relevant assessment system capable of delivering dependable information about candidate achievement and faculty performance. It assures that assessments meet accepted standards of quality and that the accompanying information management system delivers evidence of performance into the hands of key decision makers in a timely and understandable form. The presentation that follows details those standards of sound assessment and information management and then describes their application in the teacher preparation context.

A TWO-PART MISSION

Given the NCATE Standards listed above, it becomes clear that one purpose for this assessment system is to assist teacher preparation programs in meeting requirements for unit accreditation. The assessment system must document the impact of instruction on the specific competence of teacher candidates, certifying their preparedness to teach.

However, a less obvious but perhaps even more important purpose for this assessment system is to provide faculties of education with a concrete and specific way to model for their candidates the process of weaving assessment effectively into the teaching and learning process. In other words, with such a system in place, professors can teach their prospective teachers about the development and application of sound assessment practices by demonstrably adhering to the principles of sound practice themselves. There can be no better way to teach the teachers than by having professors showcase best practices, in this case in assessment.

This represents an important second purpose for the NCATE assessment system for two related reasons. First, the typical teacher will spend a quarter to a third of her or his available professional time involved in assessment-related activities (Stiggins and Conklin, 1992). All acknowledge that, if they do it well, their students will prosper and, if they do it poorly, their students will suffer. Effective instruction requires the continuous collection and use of information about student achievement via classroom assessments.

Yet, over the decades, teacher preparation programs have been repeatedly cited for their failure to provide teachers-in-training with the depth of assessment training required to meet their professional responsibilities. For the most recent summary of that research, refer to Black and Wiliam (1998). Even today, very few states explicitly require competence in assessment as a condition of teacher licensure (Stiggins, 1999) and virtually no states require that principals be assessment literate to be certified (Trevisan, 1999). As a result, we remain a national faculty unschooled in the principles of sound and effective classroom assessment.

By designing and implementing the assessment system envisioned in the NCATE Standards in the manner described below, faculties of teacher education can address this problem once and for all.

A UNIQUE ASSESSMENT ORIENTATION

Thus, in describing the active ingredients in an assessment system, I will suggest how teacher education faculty can model both the processes of gathering dependable evidence of achievement and of using classroom assessments as motivators that maximize the academic success of their candidates. But to accomplish this in one system, I must redefine the relationship between assessment, student motivation and achievement. The system described below does this by relying heavily on student involvement in the assessment process; that is, by turning classroom assessment into both an instructional intervention and a source of evidence of competence. The reader will recognize at once that this represents a significant departure from our assessment traditions.

Compelling evidence synthesized from dozens of rigorous investigations conducted around the world by Black and Wiliam (1998) reveals that deep student involvement in classroom assessment, record keeping and communication processes yields statistically and educationally significant, indeed almost unprecedented, gains in student achievement. For this reason, I will describe how these tactics can be woven into the assessment systems that NCATE envisions for teacher preparation programs.

Student-involved classroom assessment opens the assessment development process and brings students in as partners (Stiggins, 2001). Under the careful management of their teachers (who begin with a clear and appropriate vision of what they want their students to learn), students are invited to practice articulating and then applying the criteria by which their learning will be judged. For example, students might work with their teachers to analyze samples of work to devise the performance criteria that ultimately will serve as the basis for judging their performance. This kind of involvement during practice accomplishes two things. First, it provides students with a clear vision of the meaning of academic success. While they may not yet be able to hit the target, they can learn to see it. Second, their involvement in assessment design early in the practice phase of learning reveals to them where they are now in relation to that vision of excellence. The result is a clear sense in their minds of the path that they must follow to succeed. The effect is classrooms in which there are no surprises.

Student-involved record keeping brings them into the process of monitoring improvements in their performance through repeated self-assessment with performance criteria held constant. One way to accomplish this is by having students build portfolios of evidence of their success over time and by requiring periodic student self reflections about the changes they see. In effect, teachers use repeated student-involved classroom assessments to permit students to watch themselves grow—to help them chart and thus feel in control of their own success. The result is a continuing belief in the minds of learners that the target is within reach for them.

Student-involved communication brings them into the process of presenting the evidence of their own success in mastering the required knowledge and skills. One way to do this is through the use of student-led portfolio conferences. In this case, students know from the beginning of the learning process that, ultimately, they will be expected to tell the story of their own success (or failure). When they prepare over the long haul to tell that story, they experience a fundamental deep-seated internal shift in their internal sense of responsibility for their own learning. The pride in accomplishment that students feel when they have a positive story to tell and can tell it convincingly with compelling evidence can be immensely reinforcing and motivational.

By taking advantage of these tactics, faculties of teacher education can maximize, and then certify, the competence of their graduates. They can help their students see and understand what it takes to become a good teacher, find and follow the most efficient and effective path to that vision of excellence, and feel in charge of (rather than victimized by) the assessment process during their journey. Student-involved assessment, record keeping and communication can help teacher candidates:

- Build the self-confidence they need to master the knowledge and skills needed to teach,
- Build a convincing case with credible evidence that they have met program achievement standards and, indeed, are ready to teach, and
- Learn valuable lessons about how to take advantage of student-involved assessment in their own classrooms as they teach.

Further, not only does this kind of involvement lead to greater achievement for teacher candidates, but it spreads the assessment, record keeping and communication workload across many more shoulders. This can have the effect of making the faculty's assessment job much easier—that is, of making assessment in general far more manageable from the point of view of teacher education faculty. In the ultimate win/win situation, candidates learn more and faculty members save time.

With this orientation in mind, then, let's consider the components of an assessment system that meets the information needs of faculty and candidates.

THE ATTRIBUTES OF AN EFFECTIVE ASSESSMENT SYSTEM

Effective assessment systems, regardless of their context, meet four standards of excellence. They are specifically designed to (1) serve clearly articulated purposes, (2) arise out of clear and appropriate achievement expectations, (3) generate accurate information about achievement by using quality assessments, and (4) rely on dependable record keeping to produce effective communication about student achievement. I will describe these four characteristics and then discuss how they apply in the context of teacher preparation.

Attribute 1: Sound assessment systems are specifically designed to serve carefully articulated purposes.

We cannot design assessment, record keeping and communication systems to satisfy specific information requirements unless and until we have determined who will use the assessment results and how. Figure 1 lists the important users of assessment information in typical educational systems. Note that different users have different information needs at different times to answer different questions during the teaching and learning process. To provide quality information for faculty and candidates at the classroom level, we need to provide sound evidence of student achievement day-to-day by means of classroom assessment. To provide useful information for use in program evaluation contexts, to set policy or for instructional support, we need high-quality information that is comparable across classrooms and summarized periodically. This calls for standardized assessment procedures. Because of these differences in information needs, we must begin each assessment event with a clear sense of whose needs we intend to meet. So one criterion by which to judge assessment system quality is its sensitivity to user needs.

Figure 1
Sample Users and Uses of Assessment

Assessment User Sample Questions

Student	<p>Am I succeeding?</p> <p>Am I improving over time?</p> <p>Do I know what it means to succeed here?</p> <p>What should I do next to succeed?</p> <p>What help do I need to succeed?</p> <p>Do I feel in control of my own success?</p> <p>Does my teacher think I'm capable of success?</p> <p>Do I think I'm capable of success?</p> <p>Is the learning worth the effort?</p> <p>How am I doing in relation to my classmates?</p> <p>Where do I want all of this to take me?</p>
Teacher	<p>Are my students improving?</p> <p>Is it because of me?</p> <p>What does this student need?</p> <p>Is this student capable of learning this?</p> <p>What do these students need?</p> <p>What are their strengths that we can build on?</p> <p>How should I group my students?</p> <p>Am I going too fast, too slow, too far, not far enough?</p> <p>Am I improving as a teacher?</p> <p>How can I improve?</p> <p>Did that teaching strategy work?</p> <p>What do I say at parent/teacher conferences?</p> <p>What grade do I put on the report card?</p>
Administrator	<p>How do we define success in terms of student learning?</p> <p>Is this teacher producing results in the form of student learning?</p> <p>How can I help this teacher improve?</p> <p>Is instruction in our building producing results?</p>

Is instruction at each grade level producing results?
 Are our students qualifying for college?
 Are our students prepared for the workplace?
 Do we need professional development as a faculty to improve?
 How shall we allocate building resources to achieve success?
 Is our program of instruction working?
 What adjustments do we need to make in our curriculum?

Accrediting
 Agency

Do graduates of this program meet standards?
 Is the faculty of this institution performing up to standards?

Attribute 2: Quality assessment systems provide evidence of candidate mastery of clearly articulated and appropriate achievement expectations.

Knowing precisely what we are asking candidates to master is important because different forms of achievement require the application of different assessment methods to obtain dependable evidence. Thus, in any assessment context, we must begin the assessment development process by defining what it means to succeed. Does the faculty expect candidates to:

- Master subject matter content *knowledge*? Are there things that we expect them to know and understand? If so, precisely what are our expectations? Does this mean they must know it outright? Or does it mean they must know where and how to find it if and when they need it?
- Use their knowledge and understanding to *reason* and solve problems? If so, what specific patterns of reasoning must they master?
- Demonstrate mastery of *performance skills*, where it's the doing that is important? If so, precisely what processes or behaviors must they demonstrate to what level of proficiency? What specific skills must be in their repertoire?
- Use their knowledge, reasoning and skills to create *products* that meet standards of quality? If so, what products and precisely what does a good one look like?
- Develop certain motivational *dispositions*, attitudes, values? If so, what affective characteristics are they expected to develop?

Because no single assessment method is capable of assessing all of these, assessors cannot select proper methods for use in a particular context unless and until they know which or what combination of these expectations is to be assessed. Further, if a certification decision hangs in the balance, then the assessor must determine the level of proficiency required for the examinee to be judged sufficiently competent to be certified.

Attribute 3: Quality assessment systems rely on assessment methods capable of accurately reflecting the intended achievement target(s) and serve the intended purpose(s).

Since most assessment contexts call for student mastery of several different kinds of achievement and since no single assessment method can reflect them all, the assessor is always faced with the challenge of selecting from among a variety of methods. The available options include *selected response tests and quizzes* (multiple choice, true/false, matching and fill in), *essay assessments*, *performance assessments* (based on observation and judgment), and direct *personal communication* with the student. Part of the challenge of building sound assessment systems is to consistently match methods to the intended targets. Strong and weak matches are depicted in Figure 2. Our challenge is to be sure that all who will design and use assessments to develop and then certify the competence of teacher candidates are sufficiently assessment literate to know and understand what methods to use when and how to use them well.

Faculties of education also must bear in mind that their graduates must leave their program and enter their own classroom having met this same assessment literacy standard.

Figure 2
Aligning Achievement Targets to Assessment Methods

TARGET TO BE ASSESSED	ASSESSMENT METHOD			
	SELECTED RESPONSE	ESSAY	PERFORMANCE ASSESSMENT	PERSONAL COMMUNICATION
KNOWLEDGE MASTERY	Multiple choice, true/false, matching, and fill-in can sample mastery of elements of knowledge	Essay exercises can tap understanding of relationships among elements of knowledge	Not a good choice for this target--Three other options preferred	Can ask questions, evaluate answers and infer mastery--but a time-consuming option
REASONING PROFICIENCY	Can assess understanding of basic patterns of reasoning	Written descriptions of complex problem solutions can provide a window into reasoning proficiency	Can watch students solve some problems and infer about reasoning proficiency	Can ask student to "think aloud" or can ask follow up questions to probe reasoning
SKILLS	Can assess mastery of the knowledge prerequisites to skillful performance--but cannot rely on these to tap the skill itself		Can observe and evaluate skills as they are being performed	Strong match when skill is oral communication proficiency; also can assess mastery of

			knowledge prerequisite to skillful performance
ABILITY TO CREATE PRODUCTS	Can assess mastery of knowledge prerequisite to the ability to create quality products--but cannot use these to assess the quality of products themselves	A strong match can assess; (a) proficiency in carrying out steps in product development, and (b) attributes of the product itself	Can probe procedural knowledge and knowledge of attributes of quality products--but not product quality
DISPOSITIONS	Selected response questionnaire items can tap student feelings	Open-ended questionnaire items can probe dispositions	Can infer dispositions from behavior and products Can talk with students about their feelings

From Stiggins, Richard J., *Student-Involved Classroom Assessment*, 3rd ed. (Columbus, Ohio: Merrill Education, 2001)

But match is not the only dimension of accuracy that deserves our attention. To be accurate, assessments also must generate a *representative sample* of student performance that is sufficient in its scope to permit confident conclusions about candidate achievement. All assessments rely on a relatively small number of exercises to permit the user to draw inferences about a student's mastery of larger domains of achievement. A sound assessment offers a representative sample of all those possibilities that is large enough to yield dependable inferences about how the respondent would have done if given all possible exercises. Each assessment context places its own special constraints on our sampling procedures. Our quality control challenge is to know how to adjust the sampling strategies to produce results of maximum quality at minimum cost in time and effort.

Further with respect to accuracy, sound assessments are developed and used in ways that control for the various *sources of bias* that can distort results. Even if we devise clear achievement targets, transform them into proper assessment methods, and sample student performance appropriately, there are still factors that can cause a candidate's score on a test to misrepresent his or her real achievement. Problems can arise from the test, the candidate or the environment where the test is administered.

For example, tests can consist of poorly worded questions, place verbal or linguistic demands on respondents that are confounded with mastery of the material being tested, have more than one correct response, be incorrectly scored, or contain racial or ethnic bias. The candidate can experience extreme evaluation anxiety or interpret test items differently from the author's intent, as well as cheat, guess or lack motivation. Any of these could give rise to inaccurate results. Or the assessment environment could be uncomfortable, poorly lighted, noisy or otherwise distracting. Part of the challenge of building effective assessment systems is to anticipate the potential sources of bias and to be sure all who are involved in devising assessments, preparing candidates, and managing assessment environments are able to anticipate and prevent these problems before they ever impact results. A partial list of sources of potential bias is provided in Figure 3.

Figure 3
Examples of Common Sources of Bias in Classroom Assessment

1. *Sources common to all assessment methods:*
 - A. Potential problems that can occur within the student:
 - Lack of reading proficiency
 - Language barriers
 - Emotional upset
 - Poor health
 - Physical handicap
 - Peer pressure to mislead assessor
 - Lack of motivation at time of assessment
 - Lack of testwiseness (understanding how to take tests)
 - Lack of personal confidence leading to evaluation anxiety
 - B. Possible problems that can occur within the assessment context:
 - Noise distractions
 - Poor lighting
 - Discomfort
 - Lack of rapport with assessor
 - Cultural insensitivity in assessor or assessment
 - Lack of proper equipment
 - C. Examples of problems that arise from the assessment itself (regardless of method)
 - Directions missing or vague
 - Poorly worded questions
 - Poor reproduction reduces readability
2. *Sources unique to each format:*
 - A. Possible problems with multiple choice tests:
 - More than one correct response
 - Incorrect scoring key
 - Incorrect bubbling on answer sheet
 - Clues to the answer in the item or in other items
 - B. Problems with essay assessments
 - Students lack writing skill
 - No scoring criteria
 - Inappropriate scoring criteria
 - Evaluator untrained in applying scoring criteria
 - Bias due to stereotypic thinking or knowledge of prior performance
 - Insufficient time or patience to read and score carefully
 - C. Potential problems with performance assessment
 - No scoring criteria
 - Inappropriate scoring criteria
 - Evaluator untrained in applying scoring criteria

Bias due to stereotypic thinking or knowledge of prior performance
 Insufficient time or patience to observe and score carefully

D. Possible difficulties when using personal communication

Insufficient sample per student
 Inaccurate record keeping
 Distortions in memory of performance
 Bias due to stereotypic thinking or knowledge of prior performance

Attribute 4: Effective assessment systems provide for the efficient and effective storage, retrieval and communication of information.

To communicate effectively, regardless of the assessment method(s) used, certain conditions must be satisfied within the total assessment system. For example, if we are to communicate effectively about achievement, as mentioned previously, we must first identify the achievement targets we wish to communicate about. Both message sender and message receiver must be aware of these achievement expectations. If we can't agree on the targets that are to be the focus of our assessment and communication systems, we're sure to have difficulty arriving at a mutually agreeable set of symbols to convey information about candidate success. If message receivers don't know what knowledge, reasoning, skill, product or disposition targets underpin our assessments, then grades or any other symbols that are based on those assessments will be devoid of meaning for them.

Given a common vision of expectations, the next essential foundation for effective communication, again as mentioned, is the transformation of our targets into quality assessments, assessments capable of producing accurate information about candidate achievement. Inaccurate information remains inaccurate regardless of how we transform it for communication.

The result of our administration and scoring of these high-quality assessments will be a reservoir of accurate information about the extent to which each individual candidate has met the established achievement expectations. If the time period over which we gather information is relatively brief and the target relatively narrow, this reservoir may include just one or two tests or projects. This will be relatively easy to store and retrieve. But if it spans a longer period covering broader targets, such as a quarter, semester, full year, or entire educational program, our records may include numerous pieces of information we must consider together to draw our conclusions. It is this accumulated academic record of candidate success that we must summarize and share with those who must consider it in their decision making. To accomplish this, storage and retrieval systems must be thorough, dependable and efficient. In other words, they must be capable of accumulating a complete and accurate record of performance over time, taking advantage of modern information management technologies whenever possible.

However, effective communication requires more than accurate assessment and efficient storage. It also requires the thoughtful preparation of the interpersonal communication

environment. Message sender and message receiver(s) must be ready to share information in an environment conducive to hearing and understanding each other. Such an environment is possible only if *all participants assign the same meanings to the symbols used to convey information*, whether they are words, pictures, examples, grades, scores, graphs, charts, or something else. As teachers transform state standards into classroom achievement targets, then transform those targets into assessment exercises and scoring criteria, and then transform assessment results into evaluative judgments about proficiency (such as grades), there is ample opportunity along the way for the true meaning of achievement to be lost. Accurate communication about levels of achievement is possible only if we prevent that loss of meaning by (a) carefully explaining to users the specifics of all of the above transformations, and making sure message receivers understand the intended meaning of the symbols we used to communicate before we ever use them.

Further, if the faculty is to communicate effectively, there must be a designated time, place, and set of circumstances where message sender and message receiver(s) can attend, without distraction, to the information being shared. This might take the form of a conference between sender and receiver, a written report of achievement, a public presentation of achievement information, the delivery of an anticipated report card, or other communication method. To work effectively, these opportunities must permit a time when all involved can suspend other activities and attend to information being conveyed. To be most effective, they also should include time to interpret the meaning of the message, check for understanding, and devise action plans if needed. For the record, transcripts of grades and grade point averages have never satisfied these conditions.

APPLICATIONS IN TEACHER PREPARATION

Now consider how these criteria might be applied in teacher preparation programs striving to meet NCATE 2000 standards. Remember, the assessment system will need to satisfy the following necessary conditions:

- All relevant purposes must be articulated, identifying the decisions to be made on the basis of assessment results.
- All relevant achievement expectations must be spelled out.
- All faculty will need to be sufficiently assessment literate to devise quality assessments of candidate achievement.
- A record keeping and communication process will need to be developed to deliver results to intended users in a timely and understandable form.
- Candidates will need to be involved deeply in assessment, record keeping and communication processes.

Attribute 1: Articulating Purposes

In this context, at least four sets of decisions can be informed by assessment results, each associated with a different decision maker.

Candidates need information about their own achievement in order to make decisions about their own needs and to plan their use of the resources available to them. They need to know the levels of achievement they are expected to attain, where they are now in relation to those expectations, and what it will take to achieve success. Only then can they plan and effectively manage their studies. As they make that journey to competence, they need access to dependable and understandable information about their rate of progress, so they can retain the belief that the effort is worth the learning. They must believe this if they are to keep trying. Further, students need evidence of ultimate success—they need access to a reservoir of information that tells them that they have mastered all of the materials required to become a teacher.

Faculty members need access to information about the achievement of their students in order to diagnose the instructional needs of individual students and groups. That means that some assessment results can be unique to an individual student while some will need to be comparable for all students. They need these kinds of evidence to evaluate the efficacy of the instructional interventions that they plan and carry out. Only then can they make the adjustments needed to accommodate the diverse needs of their students. Indeed, they need evidence of the achievement of their students in order to judge their own efficacy, to know their strengths, areas in need of improvement and thus professional development goals for the future.

Program administrators must judge the quality of their program as a whole, determining if the teachers being graduated can put all of the pieces together in the classroom. They need to know that resources are being used effectively, that professors are competent, and that students are satisfied with the instruction they are receiving. Typically, these data will include evidence of student achievement and that evidence must be comparable across students.

In addition, program administrators and faculty together must certify the competence of their graduates. They need a complete set of indicators of competence for each individual student so as to be able to compare each student's evidence with pre-established standards to determine if they have mastered what it takes to be a teacher. Typically, their decision is backed up by a state licensing board that also must certify competence. So they need access to evidence of that competence.

Finally, *accreditation agencies*, such as NCATE, must gather evidence of program efficacy in order to determine if the program under scrutiny measures up to their standards. That body of evidence must include information about the specifics of candidate achievement.

Thus, different decision makers need access to different information about different facets of candidate achievement at different times and in different forms to do their jobs.

No single assessment can do it all—an integrated, multifaceted assessment system is needed.

Attribute 2: Establishing Achievement Expectations

NCATE Standards 1 and 2 require the collection of evidence of teacher candidate competence. In order to gather such evidence, those competencies must be defined:

- What does a teacher need to know and understand to be effective in the classroom?
- What specific patterns of reasoning must a teacher be prepared to apply to be effective?
- What performance skills must teachers master to teach effectively?
- What products must teachers be able to create as a matter of routine to promote successful learning in their classrooms?
- What dispositions do effective teachers need to bring to the classroom?

In all of these cases, the teacher education faculty must first establish ultimate standards of performance to be attained by teacher candidates by the end of the program in order to certify competence. But then the faculty also must develop a teacher training curriculum that will unfold over time to bring their candidates to ultimate competence. Effective assessment systems will serve the formative purpose of helping candidates and faculty track progress toward ultimate success and the summative purpose of proving that they arrived there. This standard asks the faculty to lay a foundation for the systematic and accurate assessment of candidate achievement by defining precisely where “there” is in terms outlined below. Note that no assertions are made about what theory of learning or what candidate selection criteria or instructional interventions a faculty might use to help their enrollees become competent. Rather, these deal only with how they define progress and success.

Knowledge Mastery

Several domains of knowledge form the foundation of the ability to teach effectively. To begin with, one must be a confident, competent master of the academic discipline(s) or content arenas that they plan to teach. In addition, there is specific knowledge of pedagogy and classroom assessment to be mastered during teacher preparation. This subdivides, for example, into such categories as knowledge of theories of learning, child development, teaching methods, the principles of sound assessment, and classroom management techniques, to mention a few. In each case, there are enduring theories, generalizations or principles, as well as concepts that teachers need to understand outright. In addition, there may be specific facts that they can learn to look up later if and when they need them. The point is that there is an important knowledge base that underpins success as a teacher.

Therefore, assessment of prospective teachers’ mastery (knowledge and understanding) of this content becomes important in three contexts. First, in the context of classroom

assessment, it is incumbent upon faculty members to help candidates understand that they are making progress in mastering this important material. Second, from a certification point of view, the program must certify that each graduate has mastered this essential material. And finally, from a program evaluation perspective, faculties of teacher education must evaluate the overall impact of their instruction to be sure that all candidates are leaving with this essential foundation of knowledge in place.

Patterns of Reasoning

Each profession defines itself, at least in part, in terms of certain patterns of reasoning. This is true of physicians, auto mechanics, attorneys, artists; and it is true of teachers too. That is, knowledge of teaching methods is useless, unless graduates can figure out on their own when and how to apply that knowledge to benefit student learning. For example, one must be proficient at analytical reasoning to be able to conduct task analyses of learning requirements—that is, to break learning down for students into manageable chunks. One must be able to reason comparatively, draw inductive and deductive inferences and synthesize concepts and ideas if they are to promote understanding. Without doubt, effective teachers are proficient critical thinkers—they can make and defend judgments through the effective application of appropriate criteria. Further, they can assemble this variety of reasoning patterns in any context and, at a moments notice, generate solutions to complex classroom problems.

Further, and from a different perspective, effective teachers understand the reasoning processes that form the foundation of whatever subjects they teach. They know, for example, that reading is a reasoning process happening as readers build their own comprehension. They understand that the process of writing, of composing original text, represents the outward manifestation of a teachable, learnable and assessable reasoning process. They understand the algorithms of math problem solving and the scientific method they teach in those domains.

Thus, these patterns of reasoning underpin success in teaching also. So teacher preparation programs must provide opportunities for candidates to learn these things. And if they do represent important prerequisites to effective teaching, then faculties of education must be in a position to clearly define those learning targets, blend them into instruction, and assess student mastery of them, for all of the same reasons mentioned in the previous section.

Performance Skills

Obviously, effective teachers rely on certain interactive behaviors to help students learn. These are the kinds of things one can see in teaching performance if one videotapes a teacher in action. They include verbal skills, interpersonal interaction skills, uses of various forms of media during the teaching process. To be sure, skillful performance requires access to the relevant knowledge base that underpins good teaching (as above) and proficiency at reasoning on one's feet to find solutions to teaching and learning

challenges (again, as described above). But the indicator of proficiency is the behavior that the evaluator sees being demonstrated in the classroom.

If these performance skills are indeed essential for good teaching, then the faculty of education must be prepared to assess their quality and use assessment results to help candidates grow, certify ultimate competence and evaluate program effectiveness.

Product Development Capabilities

From time to time in all professions, proficiency reveals itself in the form of products created by the professional that meet certain standards of quality. For the surgeon, it is an appropriate repair. For the auto mechanic, it is a smoothly running engine. For the author, it is a finished manuscript. For the teacher, there also are teaching-related products that underpin good teaching and that must therefore meet certain standards of quality. These represent tangible products that are created by the teacher, but that exist independently of that teacher, that provide evidence of proficiency.

For example, one must be able to create products such as written lesson plans, assessments of student learning, and communications with families about student performance and classroom life, all of which meet their own unique standards of quality. Further, if there are products that they expect their students to learn to create, they (the teacher) must be able to create examples of those products that illustrate the differences among high, middle and low-level performance.

So in this case, it is the responsibility of the teacher education faculty to identify those key products and to be prepared to determine through rigorous assessment that their graduates can create them. That is, they themselves must know and understand the performance criteria by which one judges the quality of a lesson plan, assessment of student learning or communication. Only then are they ready to help teacher candidates become competent, certify that competence and verify the impact of their program.

Dispositions

Successful teachers develop certain attitudes, values, interests, preferences, and motivational pre-dispositions that prepare them to fulfill the responsibilities of a teacher. Dispositions vary in their object, direction and intensity. That is, one develops attitudes about such things as students, colleagues, school subjects. One values certain kinds of learning or kinds of student behavior. In all cases, dispositions are directed at someone or something and are directional. We can have positive or negative attitudes along a continuum. We can have strong or weak values, interests or preferences. Finally, dispositions can vary in their intensity from very positive to somewhat positive to somewhat negative or very negative.

It may be that a teacher education faculty would establish certain dispositions that they expect their graduates to develop and demonstrate in order to be confident that they are

ready to teach. If that is the case, then the systematic assessment of those dispositions would be needed, both for formative and summative purposes.

Teaching Teachers

By modeling good practice in their own clear specifications of appropriate achievement targets, faculties of education reveal to their students the benefits of carefully articulating their knowledge, reasoning, skill, product and disposition expectations. Those benefits include increased efficiency for teacher and student, and increased sense of efficacy for both, and, of course, the potential for the development of quality assessments that accurately reflect student achievement and serve their intended purposes. Obviously, to gain access to these benefits, each teacher must be a master of the knowledge, reasoning, skills, products and dispositions that underpin the academic subjects that represent their assigned teaching responsibilities. Part of the instructional and assessment task faced by teacher education faculty is to be sure they are.

Further, it is worthy of note that NCATE 2000 Standards, articulating what candidates are expected to know and be able to do, represent the very kinds of achievement expectations that an assessment system would be designed to evaluate.

Attribute 3: Accurate Assessments of Achievement

With the above achievement destinations in hand, along with road maps to success, the foundation is in place for the formative assessment of candidate progress and the summative assessment of qualifications to teach. The next step in assessment systems development is the transformation of those achievement expectations into the exercises and scoring schemes that will comprise the assessments of achievement. As mentioned previously, the development of accurate assessments requires the careful

- Selection of proper assessment methods to assure a match to the intended target(s),
- Assembly of assessment exercises into an array that representatively samples the relevant domains of student achievement, and
- Development, administration and interpretation of assessments so as to minimize bias.

Since all teacher education faculty members will be involved in the development and use of formative assessments, all must be assessment literate; that is, they must understand and be able to apply standards of sound assessment practice. More specifically, they must know what assessment methods to use when, how to devise assessment exercises and scoring criteria that meet specific standards of quality, and how to communicate effectively about student achievement.

In effect, faculty members must understand, for each course they teach, how their particular achievement expectations fit into the total curriculum that leads to ultimate success for teacher candidates. What specific enduring understandings, reasoning

proficiencies, performance skills, product development capabilities or dispositions are their candidates to master? Then that instructor's responsibility is to devise an assessment plan for that course that will yield dependable information for them and their candidates on how each is progressing on the journey to excellence, as well as dependable evidence that they arrived.

This will require the regular application of the target-by-assessment method match information depicted in Figure 2, along with the principles of sound sampling that accompany each of the four basic assessment methods and the avoidance of bias in assessment through the thoughtful control of sources listed in Figure 3. All faculty members who conduct assessments that feed into the assessment systems must be able to build and use instruments and procedures that score high on these scales.

In addition, all of this will be most easily and effectively managed by faculty members and their candidates with the application of assessment procedures that provide for deep student involvement in the assessment, record keeping and communication processes.

Teaching Teachers

If prospective teachers do not learn these lessons about standards of excellence in assessment, then faculties of teacher education place the students of their graduates directly in harm's way. Those students and their families face the prospect of the ongoing mismeasurement of their achievement, along with all of the predictable consequences of that. Thus, this represents a pedagogical, ethical and indeed moral imperative in teacher education.

If new teachers come into the classroom with knowledge of and proficiency in implementing student-involved assessment procedures, they carry with them tools that have been proven to maximize both the confidence and achievement of their students. To reiterate, these can be taught and learned through modeling by teacher education faculty.

Attribute 4: Information Management

With the information of all assessment users identified, achievement expectations in place, and accurate assessments being used, key standards are in place for the implementation of an effective assessment system. However, all of this is wasted if procedures are not also in place to deliver information about candidate achievement into the proper hands in a timely and understandable form.

To reiterate briefly, this means both message sender and receiver understand precisely what it is that they need to communicate about—both are aware of the important achievement expectations. Any lack of understanding in this regard will serve as a barrier to effective communication. It also means that a storage place must hold a reservoir of accurate information about each candidate's achievement. Inaccurate or inaccessible information serves no one well. In order to assure the effective movement of information from sender to receiver, both must understand the symbols used to mean the

same thing. If they do not, incorrect interpretations are assured. And finally, an opportunity must be created for message sender and receiver to share information in a focused environment free of distractions. Without this, the connection will be missed.

In teacher preparation programs, two interrelated information management systems are needed. One facilitates communication during the learning process. This one connects candidate and faculty member. The second arises from the first, relying on evidence accumulated late in the learning process to facilitate summative decisions—that is, to certify competence to teach. To work effectively in the teacher preparation context, *I suggest both systems must be managed by the candidate, not the professor.* The faculty's job is to be sure each candidate has access to dependable information about her or his mastery of important achievement targets.

Formative Information Management

As the candidate journeys through the curriculum learning to master knowledge, reasoning, skill, product and disposition expectations, ongoing classroom assessments will provide evidence of success. Early in the learning process, they might not be able to demonstrate a high level of proficiency. But over time and with guided practice, if the program of instruction is working, proficiency will increase. An information management system is needed to document that change.

This need can be satisfied with a growth portfolio built by each candidate individually. This is best conceived as a series of growth portfolios, each providing candidates with the opportunity to accumulate, in collaboration with their professors, evidence of increasing proficiency in one particular domain of performance, such as content mastery, classroom management, classroom assessments, etc. The candidate might build one such growth portfolio per course or field experience.

The nature of the evidence will vary as a function of the valued achievement expectations. For mastery of foundational understandings, evidence might take the form of performance on selected response assessments. For complex reasoning, skill or product targets, records of success on essay or performance assessment would be better. During an internship or field experience, products created by the teacher candidate might be collected. In any case, the accumulation of such evidence over time, with standards held constant, will tell both student and teacher that improvement is apparent. Regular self-reflections on that improvement will fit well into this kind of portfolio.

For example, a growth portfolio developed during a course on classroom assessment methods might begin with the candidate developing a sample assessment. Then that candidate could apply standards of assessment quality in evaluating that product. Subsequently, as instruction unfolds and the candidate becomes more proficient as an assessor, additional evaluated samples might be added to the portfolio. In addition, the candidate might reflect on apparent improvements in the quality. Indeed the teacher candidate might also reflect on improvements in the richness of her or his critiques. Both

candidate and faculty member might then participate in a concluding interview, reviewing the evidence of growth and ending proficiency.

At the end of this course, if instruction has been effective and the candidate has mastered essential assessment understandings, this growth portfolio will conclude with some samples of high quality work. One or more of those samples might then be transferred to a status report portfolio, where evidence of ultimate competence resides. Let's consider that next.

Summative Information Management

As the entire teacher training program unfolds, including course work and field-based learning experiences, each candidate will acquire evidence of having met the program's competence requirements for certification as a teacher. This evidence, initially collected in the growth portfolios, would then be transferred to a portfolio in which each candidate builds his or her argument over time for successful program completion and graduation.

This is not a growth portfolio. Change over time has no relevance in this collection of work. It contains only the most dependable and compelling evidence of competence, with that evidence taking the form of assessments successfully completed by the candidate as attested to by faculty members. Again, depending on which part of their case the candidate is building, the evidence might come from selected response, essay, performance or some other form of assessment. Each relevant piece of evidence would be accompanied by the candidate's written reflection on why that piece of evidence is relevant.

Essentially, the responsibility for presenting arguments in support of having successfully completed the program would reside with the candidate. The faculty's classroom assessment responsibilities are to be sure each candidate understands the assessment tasks and then has access to dependable information about his or her attainment of program requirements, as defined by that faculty.

As a culminating experience, all candidates might be called upon to submit their portfolio for final review by the faculty. As a part of that experience, candidates might be asked to conduct a meeting with the faculty, presenting his or her portfolio for final evaluation. It would be incumbent upon the faculty to be sure that criteria are developed for the evaluation of these summative portfolios and that candidates learn those criteria so they can account for them in assembling their presentation. Further, faculty members would need to be sufficiently well schooled in the application of those criteria to be able to demonstrate high inter-rater agreement.

Technology Can Help

Electronic portfolio information management software systems are available in the market place to assist in both the formative and summative applications of these portfolios in the context of K-12 schools. These systems offer faculty immense time and

energy savings by permitting the entry, storage and retrieval of achievement standards, actual assessments or their results as evidence of achievement (including samples of student work), and student self reflections on success. It may be that these can be adapted for use in higher education or that parallel systems already have been developed for college use. For profiles of currently available systems, faculties of teacher education can consult the following web site: www.wested.org/wcmt.

SUMMARY OF COSTS AND BENEFITS

If this system makes so much sense, the reader might ask, why haven't we developed in before? What will be the costs of such a departure from our teacher education traditions? What are the downside risks?

The Costs and Risks

Clearly this plan represents a marked departure from tradition in that it calls for faculty to agree upon a set of achievement standards, reflective of the knowledge, reasoning, skills, products and dispositions that underpin effective teaching. And it calls for a division of these teaching assignments across the courses that comprise a program of study. This runs counter to a tradition in which each faculty member decides upon the content priorities of that course and the standards by which teacher candidates will be measured. The emotional and time costs of doing this work will be considerable.

Similarly, this plan calls for the development of a curriculum for each standard (or perhaps set of standards), describing how candidates will grow from totally unprepared to highly qualified teachers over the span of a program. This again runs counter to the traditions of academic freedom that permit each faculty member to structure her or his own learning sequences. Again, we confront the need for faculty time to formulate such a mutually agreeable curriculum.

In short, the system described above calls for teamwork among faculty members who traditionally conceive of themselves as independent entrepreneurs. A reward system that acknowledges only individual achievement will now be asked to value collaboration.

This system calls for the ongoing assessment of candidate success along the journey to becoming good teachers by an assessment literate faculty, when we know that assessment literacy is lacking throughout our educational system, including in higher education. In fact, current practice typically involves as little assessment of candidate achievement as is possible. Such assessment wisdom will take time to develop and implement.

The system calls for heavy reliance on student-involved assessment, record keeping and communication processes, placing responsibility for evidence gathering and presentation squarely on the shoulders of the candidate. Nothing in our assessment traditions prepares the faculty for this. Assessment has always been the job of the professor. Faculty will need to take time to undergo a shift of values to learn how to do this.

This assessment plan requires systematic and continuous record keeping and communication about candidate achievement. This stands in stark contrast to an environment that traditionally delivers grades to the registrar and then to the transcript every quarter or semester. The development of an information management system capable of delivering greater detail about candidate achievement for both formative and summative purposes will cost money and take time to develop.

And finally, this kind of assessment system will permit the evaluation of faculty in terms of the actual achievements of their teacher candidates. Only those programs that produce competent teachers will stand muster before NCATE standards.

These, then, are the costs. But what do we get in return for this investment? Who benefits and how?

The Benefits

To begin with, if the faculty clearly articulates standards that truly do underpin excellence in teaching, we will receive in return more highly qualified teachers. The students of those new teachers, in turn, will benefit from greater learning.

The evidence of competence in teaching will be of higher quality. Both formative and summative assessments will be more valid and reliable because those developing and implementing them will know what they are doing.

In addition, after development, this system ultimately saves time and energy. Everyone's assessment workload becomes easier to manage for several reasons. First of all, by placing limits on the achievement targets that teacher candidates are to hit, we limit the sheer amount of assessment to be conducted. The faculty assesses only that which they determine is important, given what we know about the attributes of effective teachers. Second, assessment literate faculty will be able to select the most efficient method of assessment for use in each specific context. Third, the actual workload associated with assessment, record keeping and communication becomes easier to bear because it is spread across more shoulders.

Further, candidates will be more highly motivated to work hard and learn as much as they can. This will occur for two reasons. First, the environment within which they learn will set them up to succeed and will continuously reveal to them that they are in control of that success. Besides, they will know from the start of the program that a lack of specific evidence of competence in the end will deprive them of the right to graduate. In other words, grades and a grade point average will not suffice.

And finally, if their candidates demonstrate regular improvement in terms valued by the teacher education community, each faculty members' sense of efficacy will increase. Clear connections between what each faculty member teaches and what their faculty teammates teach will establish the importance of each professor's contributions to the total program.

In another context, I have published guidelines for a faculty self-assessment of the quality of the classroom assessment training that they provide to their graduates (Stiggins, 1999). After listing several subcategories of teacher competence in classroom assessment, I ask that faculties ask themselves four questions:

- Are we teaching it, such that our teacher candidates have the opportunity to learn it?
- Do we model its application in our own teaching?
- Do we certify competence through rigorous assessment?
- And perhaps most importantly, do our graduates feel that they have mastered it?

Not only will the development and implementation of the assessment systems envisioned here help teacher preparation faculties meet NCATE standards for accreditation, but it will provide affirmative answers to these four key questions.

References

Black, P. & Wiliam, D. (1998). Inside the black box. *Phi Delta Kappan*, 80, 139-147.

National Council for Accreditation of Teacher Education (2000). *Performance-based Accreditation Standards*, Washington DC: Author.

Stiggins, R. J. & Conklin N. J. (1992) *In Teachers' Hands: Investigating the Practice of Classroom Assessment*. Albany NY: SUNY Press.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27.

Stiggins, R. J. (2001). *Student-Involved Classroom Assessment*, 3rd ed. Columbus OH: Merrill Education, an imprint of Prentice Hall.

Trevisan, M. (1999). Administrator certification requirements for student assessment competence. *Applied Measurement in Education*, 12(1), 1-11.