Developing and Accrediting Providers of Classroom Assessment Training for Teachers

A commissioned paper for the
Council for the Accreditation of Educator Preparation

Rick Stiggins

Assessment Consultant
5526 Yorkshire Place
Lake Oswego OR 97035
rickstiggins@gmail.com
503-697-7639

Developing and Certifying the Assessment Literacy of Teachers

Assessment is the process of gathering evidence of student achievement in order to inform instructional decisions. Teachers typically spend a quarter to a third of their available professional time involved in assessment-related activities. Compelling research conducted around the world over the past two decades (detailed later) reveals that, if they do it well, profound gains in student achievement can result, with the largest gains accruing for struggling learners; that is, achievement gaps narrow. If they do it poorly, student learning suffers, especially for low achievers. Yet, we still provide very few teachers with the opportunity to master essential principles of sound classroom assessment practice, either in pre- or in-service contexts. This is especially true of those entering the profession through non-traditional channels such as Teach For America.

And, lest we believe teachers can turn to their building principals or district administrators for training or assistance in the assessment domain, if fact, relevant assessment training remains practically *nonexistent* in administrator preparation programs across the nation. Historically, this has left, and continues to leave, both teachers and school leaders without the professional knowledge and skills needed to fulfill rapidly evolving and increasingly important assessment responsibilities. In effect, it renders assessment practically useless as a school improvement tool, and it is both naïve and dangerous to believe we can address this long-standing and deep-seated *instructional problem* by merely demanding higher annual test scores at an ever-higher level of authority with an ever-louder voice or by attaching educators' salaries to those test scores.

This paper centers on the assessment competencies and beliefs teachers and their local leaders need to master in order to assess productively in their schools and classrooms. By "productively" I mean assess in ways that help students succeed. Having mentioned school leaders, because of the context and purpose for which this paper is written, I will leave them behind and refer predominately to teacher development. I will leave the development of assessment literate school leaders—a critically important school improvement topic to say the least—for another time.

Teachers assess well when they routinely (a) gather *dependable* evidence of student achievement, and (b) use the assessment process and its results effectively *either to support or to certify student learning success*, depending on the context. I will expand on these foundations of teacher effectiveness and describe how teacher preparation programs can and should help candidates develop appropriate levels of assessment literacy in these terms. My presentation will unfold in five parts:

- Part 1 examines society's evolving social and educational priorities as they relate to assessment. It is within this socio-political environment that schools and teachers must plan and carry out their assessment practices. Therefore, in effect, this environment dictates the principles that must guide our thinking about assessment competence for teachers. I will list and describe those principles.

- Part 2 details the specific classroom assessment competencies teachers and their direct supervisors need to master in order to be able to assess productively in schools today. These are the competencies that must be reflected in standards for professional preparation (pre- and in-service), certification, and program accreditation.
- Part 3 provides a status report on opportunities for practitioners to become assessment literate including barriers that have so effectively denied practitioners access to those opportunities.
- Part 4 centers on keys to effective assessment training in teacher education contexts and strategies for evaluating and accrediting programs aimed at promoting assessment literacy.
- Part 5 provides a summary of the paper's basic themes in the form of an analysis the cost or risks and benefits of providing productive assessment literacy training.

Part 1: Understanding the Assessment Implications of a Changing School Mission

Over the past two decades, our society has expanded the social mission of its schools. The new two-part mission requires fundamental changes in our thinking about and approach to assessment. Historically, a social mission of schools has been to begin the process of sorting students into the various segments of our social and economic system. We have accomplished this by fixing the amount of time to learn (one year per grade level) and allowing the amount learned during that time to vary. So some students complete first grade having learned a great deal, while others may have learned much less. Those who prospered in first grade carry that advantage into second grade and continue to grow. Those who lagged behind in first grade learn less in second. Thus, the spread widens. Over 13 years of education, students continue to spread themselves along an ever-wider continuum of achievement, thus permitting the assignment of a (hopefully) dependable rank in class at the end of high school. The teacher's role has been to assess student achievement, assign report card grades, and rank those who remain (i.e., have not dropped out) based on achievement at graduation. Assessment's role has been to provide the evidence for grading, weeding out the unwilling or unable, and, ultimately, ranking.

However, because of the accelerating social and technical evolution of our society over the past two decades, we have come to realize the insufficiency of this unitary mission. We have come to understand that schools must take on an additional responsibility: make sure all students—not just those at the top of the rank order—become sufficiently well-equipped lifelong learners to survive. So now we demand that *all* students, not just a select few, master the fundamental reading, writing, mathematics problem solving, and other proficiencies that they will need to contribute in this increasingly complex and fast-changing digital world. Political forums and media outlets herald the new mission almost daily: all students are to be made ready for college and the workplace. In effect, we are facing the reality that our economy can no longer absorb those who fail to finish high school and don't pursue at least some post-secondary training. So now, educators are being held accountable for delivery of universal competence in these lifelong learner terms with dire consequences hanging in the balance if

they fail. The achievement gap must narrow and dropout rates must decline or dire consequences will ensue. Note, however, that this does not absolve educators from also delivering a dependable rank order based on achievement at the end of high school.   The original mission remains in place too.

If this expanded two-part mission is to be fulfilled, then our assessments are going to have to deliver far more than merely evidence for grading, weeding, and sorting. To begin with, they need to motivate all students to strive for academic success, not just a few at the top of the rank order. Historically, schools have motivated students by relying on the intimidation of competition and accountability.   By demanding high performance and threatening dire consequences for failure, schools traditionally have endeavored to generate the kind of "productive" anxiety that would either (a) drive learners to work harder and thus attain ever-higher levels of achievement, or (b) cause them to drop out of school in hopelessness.   Under the previous unitary mission, either result was acceptable because they both enhance the dependability of the rank order at graduation. The despair that comes with ongoing failure drives low achievers even lower while the exhilaration of success energizes high achievers to even greater highs. The broader the spread in achievement between the lows and highs—that is, the greater the spread between students all along the achievement continuum—the more reliable will be the rank order. Mission fulfilled.

But, as it turns out, this motivational strategy falls prey to a fatal flaw in the presence of the expanded institutional mission: it triggers a critical dilemma in emotional dynamics for many students. If struggling learners who have not yet mastered foundational lifelong learner academic standards give up in hopelessness (which the previous system caused many to do), then part two of the mission cannot be fulfilled.   In other words, the traditional value proposition that has underpinned our thinking about assessment and student motivation has held that, by creating an artificial scarcity of academic success and asking students to compete for it, society derives the greatest value from its investment of education dollars.   Assessment's role was to continuously track each student's achievement and, ultimately, to determine the winners and losers. But again, because of the new expanded mission, we must confront the dilemma of emotional dynamics. If some students decide that academic success is inevitably beyond reach for them, they will give up in hopelessness long before they become lifelong learners. Mission conflict.

From a similar perspective, traditionally, policy makers (most of whom also have never been given the opportunity to become assessment literate either) have endeavored to motivate teachers and school leaders by demanding ever-higher high scores on annual accountability tests, assuming that this would drive them to strive for excellence.   The inadequacy of this approach has revealed itself over the past 60 years in the fact that achievement, as reflected on our National Assessment of Educational Progress, has flat-lined while dropout rates have increased radically in many contexts. The demand for accountability obviously has not done the job. Instead of striving for more effective schools, record numbers of teachers in struggling schools who are at a loss as to how to deal with such low test scores, continue to leave the profession in hopeless frustration in record numbers. Unless and until we find ways to help these teachers experience professional success in the form of greater and more frequent academic success for their students, this exodus will continue. In other words,

teachers don't need to be motivated to try harder. They need access to more productive practices, so they can see more of their students succeeding and thus develop their own sense of professional efficacy.

Thus, it becomes clear that, for all of the good they may have done to help us *identify* our achievement gaps over the decades, our assessment beliefs and processes have done little to help us *narrow* those gaps. Indeed, they have caused major segments of the student and teacher populations to give up in hopelessness. This should not surprise us, given the general lack of assessment literacy among our corps of professional educators. *Practitioners have not been given the opportunity to learn how to assess well or use assessment productively. The bottom line is that our educational infrastructure is not now, nor has ever been, prepared to deal productively with assessment processes or results.*

This need not be so. It is truly paradoxical that, as it turns out, assessment represents one domain of professional practice where productive *instructional practices* are available for teacher use in maximizing student motivation and increasing achievement across the range of achievement. I refer here to day-to-day *classroom level of assessment*, where applications have been developed that do improve student learning. I will describe these practices below and translate them into specific competencies in part two of my presentation.

But first I will complete this retrospective on our evolving social and educational environment as it relates to traditional assessment practices in order to establish the relevance of the assessment competencies I will describe later.

Historically, our vision of excellence in assessment, with its obsessive focus on large-scale accountability testing, has demanded reliance on the great efficiency (meaning relatively low cost) of multiple-choice tests. This has had the effect over the decades of severely restricting the array of achievement targets that could be assessed to mastery of elements of knowledge and the simple patterns of reasoning. What gets tested in high stakes environment is what gets taught. So, in effect, the assessment method has had a profound impact on our beliefs about the valued outcomes of schooling. While mastery of specific elements of knowledge and simple patterns of reasoning dominate, complex patterns of reasoning and problem solving, performance skills, and product development capabilities—all foundations of 21st century lifelong learning but not translatable into the dominant assessment method— typically have been all but ignored. To assess these more complex forms of achievement appropriately in the future, which we know we absolutely must do, every practitioner will have to be ready to bring the full array of assessment methods to bear. Without the ability to develop and use written, performance, and interactive assessments when needed, truly quality assessment will remain beyond reach. Historically, as I have said, practitioners have not been given the opportunity to develop this capability.

Indeed, over the decades the entire matter of assessment quality has been given short shrift in schools, to say the very least. In fact, while measurement experts have developed quality control criteria for judging and enhancing the precision of large-scale tests, a strong case can be made that we have not cared at all about the quality of the other 99.9% of assessments that happen in a student's life—those developed or selected by teachers every day in their classrooms. If we had cared about their quality, would we not have made absolutely sure each and every teacher and school administrator entered her or his classroom or school from day

one with the professional knowledge and skills needed to create quality assessments and to use them effectively to support and verify student learning? We already have established that, clearly, we have not done that. The problem is that the new part of the mission—helping all students become lifelong learners—cannot be fulfilled without day-to-day *high-quality* assessments used to identify and overcome gaps in each student's achievement; that is, teachers and their students need continuous access to assessments designed to track and manage improvement in learning as it unfolds over time in the classroom. Only then can they act to promote continuous progress. Our current levels of assessment literacy do not permit this level of assessment use.

For all of these reasons, the time has come to develop a new vision of excellence in assessment for American schools. This revised vision must:

- Accommodate our wide range of assessment *purposes* (users and uses)
- Accommodate the full range of *learning targets* that underpin success in the adult world
- Assure the *quality of assessments* at all levels for all purposes,
- Assure the *effective communication* of results to the intended user, and
- Assure *productive motivational dynamics* for all learners and their teachers.

These, then, are the keys to assessment quality that my team and I have spent over three decades developing and refining. They can be developed into guiding principles teachers, school leaders, and communities must embrace if our school improvement processes are to be effective (Stiggins, Manifesto). Each implies mastery of a specific set of assessment competencies by teachers and their supervisors. First, I will provide an overview of the principles and then I will use them to provide the structure of the remainder of my presentation.

Guiding Principle 1: Assessors in any context must begin each assessment with a clear sense of purpose; that is, they must know in advance why they are assessing—who will use the assessment results and how will they use them.

If assessment is the process of gathering evidence of student learning to inform instructional decision making, then assessments conducted in schools must meet the information needs of a wide variety of different decision makers (users and uses), including students, parents, teachers, school leaders, district leaders, policy makers and communities. To use assessment effectively, we must understand that these various decision makers, in fact, need access to different kinds of information in different forms at different times to help students learn and to certify and report the extent of student learning success. More specifically, it is helpful to think of assessment users at three different levels of application:

- In the classroom,
- At the level of interim/benchmark levels (standardized assessments used periodically during the school year, say, quarterly to evaluate progress) , and
- At annual accountability testing time.

Each level of use involves different decisions and decision makers and places fundamentally different demands on assessments used.

In addition, it is helpful to understand that, at each level of use, practitioners can take advantage of assessments either *to support* student learning or *to judge its sufficiency* (that is, to certify the learning, as in grading it). Table 1 (see next page) blends these three levels of use and the two purposeful kinds of applications to create a six-cell portrait of a truly balanced local assessment system, each cell of which makes a unique contribution to student well-being through assessment. Local school leaders must be sufficiently assessment literate to be able to create such a system if they are to be sure all important users have their information needs met. Teachers must be sufficiently assessment literate to conduct assessments in their classrooms in a manner consistent with their intended purpose—formative or summative. If the information needs of any of the users in this table is disregarded, or if decision makers are provided with mis-information due to inept assessment, inappropriate instructional decisions will be made resulting in harm to student confidence, motivation, and learning, as well as teacher efficacy.

For this reason, the starting place for the creation of a quality assessment in any context must be a clear sense of purpose—an understanding of the information needs of the intended assessment user/decision maker(s) to be served. The assessment in question must be designed to provide the needed information. Without this, the assessor cannot proceed with productive assessment design, development, or implementation. So, teachers and others who assess must understand (a) how their assessments fit into this big picture provided by Table 1, and (b) how to differentiate among and fulfill the variety of purposes for assessment that are their assigned responsibility. In particular, they must understand the differences between assessment used to support learning and assessment used to certify it. Both are important but they are different.

Guiding Principle 2: Assessors must always start the assessment process with a clear and appropriate vision of the learning target(s)—that is, an understanding of precisely what needs to be assessed by way of student achievement.

The quality of any assessment also depends on how clearly and appropriately one defines the achievement target(s) to be assessed, because these are the targets that must be translated into the exercises and scoring schemes that will make up the assessment. Teachers cannot validly (accurately, dependably) assess academic achievement targets that they have not precisely and completely defined and mastered themselves.

Our academic standards focus on many different kinds of achievement expectations, from mastering *content knowledge,* to using that knowledge for complex *reasoning and solving problems,* to demonstrating a *performance skill* (performing a flute recital or speaking Spanish), to creating *products* that meet certain standards of quality (writing a strong term paper, for example). These diverse kinds of achievement demand that teachers bring the full array of assessment methods into play in the classroom; that is, no single method can reflect all of these. Without a clear vision of the learning target(s) an assessor cannot pick a proper method.

Table 1

Essential Components of a Balanced Local District Assessment System

Level of Assessment and Key Issues	Formative Applications	Summative Applications
Classroom assessment		
o *Key decision(s) to be informed?*	What comes next in the student's learning?	What standards has each student mastered? What grade does each student receive?
o *Who is the decision maker?*	Students and teachers	Teacher
o *What information do they need?*	Evidence of where the student is now on learning progression leading to each standard?	Evidence of each student's mastery of each relevant standard
o *What are the essential assessment conditions?*	• Clear curriculum maps per standard • Accurate assessment results • Descriptive feedback • Results point student and teacher clearly to next steps	• Clear and appropriate standards • Accurate evidence • Focus on achievement only • Evidence well summarized • Grading symbols that carry clear and consistent meaning for all
Interim/benchmark assessment		
o *Key decision to be informed?*	Which standards are our students not mastering? Where can we improve instruction right away? Which students need what specific help?	Did the program of instruction deliver as promised? Should we continue to use it?
o *Who is the decision maker?*	Instructional leaders; Learning teams; Teachers	Instructional leaders
o *What information do they need?*	Any standards students are struggling to master	Evidence of mastery of standards
o *What are the essential conditions?*	• Clear and appropriate standards • Accurate assessment results • Results reveal how *each* student did in mastering *each* standard	Accurate assessments of mastery of program standards aggregated over students
Annual testing		
o *Key decision(s) to be informed?*	Which standards are our students mastering/not? Where and how can we improve instruction over the long term?	Are enough students meeting standards?
o *Who is the decision maker?*	Curriculum & instructional leaders	School and community leaders
o *What information do they need?*	Standards students are struggling to master	Percent of students meeting *each* standard
o *What are the essential assessment conditions?*	Accurate evidence of how *each* student did in mastering *each* standard aggregated over students	Accurate evidence of how *each* student did in mastering *each* standard aggregated over students

Further, without a clear and complete understanding of what it means to demonstrate success at learning—what it means to perform well and that differs from performing poorly—the assessor cannot create or select exercises (items, tasks) or scoring schemes. Without a sense of the final achievement destination and a clear vision of the signposts along the way leading to that target against which to check students' progress, teachers will be unable to assemble assessments in an array which, over time, helps both them and their students see growth so they can feel confident that ultimate success is within reach.

Guiding Principle 3: Assessors must always create or select high-quality assessments that yield dependable evidence of the extent of student mastery of the learning target(s) in question.

This principle rests on a moral and ethical imperative: the right to wittingly or unwittingly mis-measure the achievement of students is NOT an entitlement that comes with academic freedom. No one in any context (K-12 or higher education) is entitled to use poor quality assessment because of the harm it does to learners. To create dependably high quality assessments, assessors need to do four specific things well. They must be prepared to:

- Select a proper assessment method given the learning target(s) to be assessed
- Sample student achievement appropriately
- Develop high-quality exercises and scoring schemes to include in the assessment
- Anticipate sources of bias that can distort results and minimize their effects

Assessments that rely on an improper method, sample poorly, include bad items, are mis-scored, or yield biased results due to factors other than student achievement will lead to incorrect inferences about the level of student learning and poor quality instructional decisions, placing student academic well-being in jeopardy.

The lack of assessment literacy among teachers and, frankly, the historic distrust of teachers at accountability test time have combined to give rise to the sense that it is unwise to rely on teachers to develop their own assessments. One alternative solution to meeting teachers' information needs has been to provide teachers with the tests they need. Published instructional materials often have included their own assessments. But the problem in this case is that these tests most often are developed in the complete absence of any awareness of or concern for assessment quality. Another more recent approach has been to form "data teams" and have teachers analyze annual test data to derive the information they need to make their instructional decisions. This is not inappropriate. It's just that tests given once a year are of little value to those who are making instructional decisions every three to four minutes. Still another option has been to form teachers into collaborative professional learning communities where they can create interim assessments and team up to analyze results. Again, this is not inappropriate. It's just that, most often, the team members lack the assessment literacy needed to do this job well. The result will be poor quality assessments and all that goes with them.

Those who have advocated these alternative approaches to providing the achievement

information teachers need reveal a very high level of naiveté about instructional life in classrooms. These approaches to assessment cannot provide the evidence teachers need to do their jobs of supporting learning while it is happening or certifying learning at report card grading time. Anyone familiar with day-to-day classroom instruction understands that teacher effectiveness and student learning success turn on teachers and students having continuous access to information on where students are now in relation to where we want them to be so students can narrow that gap. Teachers who are unable to generate that evidence *on their own* will not (cannot) be as effective as they need to be in managing the assessment/instruction interaction.

Guiding Principle 4: Assessors must always communicate assessment results effectively to the intended user(s).

Historically, assessment results have taken the form of numerical scores attached to briefly labeled forms of achievement such as *reading, writing, science, math,* and the like. Often, neither the actual learning target(s) hidden under these labels nor the meaning of the scores are understood by those who are to act on results. Truly effective communication of assessment results requires that these things be clear both to the communicator and to the message receiver who is to act on them.

Although many assessments differentiate levels of achievement in ways that can be translated into scores, both common sense and research advise us that numbers (scores) are not the only—or in certain contexts even the best—way to communicate about achievement. Sometimes we can and should use words, pictures, illustrations, examples, and many other forms of feedback to convey meaning. Classroom level decision makers need access to useable *information*, not merely "data".

Educators who are critical consumers of assessment information are constantly asking of the assessment results they receive, "Precisely what is being assessed here, and do I know what the results actually mean in terms of those learning targets?" They do not rest until they have good answers to both questions, and they certainly don't use the results to affect students until they have those answers. They demand clear thinking about and communication of achievement standards and effective communication about mastery of each relevant standard, and they demand these both in their own assessments and those of others.

Further, the attributes of effective communication vary by assessment purpose. Communication of assessment results intended to support learning must differ in form from communication of judgments regarding the sufficiency of learning. School leaders must be aware of these differences and they must build local district and classroom communication policies and regulations that are sensitive to those differences. Teachers too must be prepared to conduct day-to-day classroom communications that do both effectively when needed. Further, and perhaps most importantly, when the purpose is to support learning, well-prepared teachers can help students understand their own assessments by providing descriptive feedback (vs. judgmental) so those students can make smart decisions about how to advance their learning based on interpretation of their own results.

Guideline 5: *(Note: This guideline represents a compilation of key points made in the first four guidelines. I collect them here as a summary guideline because of their importance and because they represent a paradigm shift in our thinking about productive classroom assessment.)* **When appropriate <u>during student learning</u>, competent teachers involve their students in the assessment, record keeping, and communication processes to keep them in touch with, understanding, and in control of the development of their own academic capabilities.**

Historically, our assessment systems have assumed that, if we just get school leaders and teachers the right assessment results at the right time in the right form then all of the important instructional decisions will be made and schools will become as effective as they can be. To be sure, the adults in the system make key decisions and major contributions to the effectiveness of schools. However, this assumption has the effect of leaving out the other half of the data-based instructional decision making team: *students*. Assessment is NOT merely something the adults do to students. We must open our minds to the understanding that many of the most crucial data-based instructional decisions are made by the learners. Students are constantly evaluating their own performance and making key decisions about their own likelihood of success, how to achieve that success, or whether to give up in hopelessness based on their evaluation of their own past and current academic performance and they are acting on those judgments. *The greatest potential value of classroom assessment is realized when we open the process up during their learning and welcome students in as full partners so they can watch themselves grow.*

In fact, the emotions surrounding the assessment experience form the foundation of school success for every student. Successful learners begin with a sense of optimism and hope, anticipating that they are going to succeed. Each time they do succeed, the resulting evidence supports their positive expectations of themselves. This triggers the release of dopamine which not only feels good (is fulfilling) but also immediately enhances actual cognitive functioning enabling further success. So they try for more learning with enthusiasm and vigor, thus succeeding again and laying a foundation for even more effective cognitive engagement. This is how and why success can become a self-fulfilling prophesy. From then on, optimists predict learning success in similar contexts and act productively on those predictions.

On the other hand, problems arise when students expect no success or predict success and fail. No dopamine. No good feeling of fulfillment. Not enhanced learning ability. Depressed energy. Disengagement. Pessimism. This too can begin to feed on itself and result in a losing streak and downward spiral. We know how to prevent this. If teachers are careful and use the proper assessment processes (described below), they can make sure students continuously predict success and deliver on that expectation. Here is how:

Within each of the four guiding principles above—clear purpose, clear target, quality assessment, and effective communication—there are opportunities to bring students into the classroom assessment, record keeping, and communication processes while they are learning and, in doing so, tap a well spring of motivation that cannot be tapped in any other way. The greatest potential value of classroom assessment is realized when we open the process up *during the learning* and welcome students in as full partners so they can watch themselves grow. This is NOT merely about having students trade test papers or homework in order to

score or grade each other's work. I am suggesting something entirely different from that.

Students who participate in the thoughtful analysis of the quality of their work during instruction come to understand the critical element of the valued achievement targets become better performers. When students:

- see what good and poor quality work look like right from the beginning,
- learn to self-assess and track their own success as they grow, and
- play a role in tracking and communicating with others about their success,

their engagement and achievement skyrocket, with the largest achievement gains accruing for low achievers. These are principles of assessment FOR learning (Chappuis, et al., 2012). Everyone wins, but those who have the most to win win the most and achievement gaps narrow. As students learn to confidently and competently evaluate their own work, they travel rapidly and confidently down the road to becoming better performers.
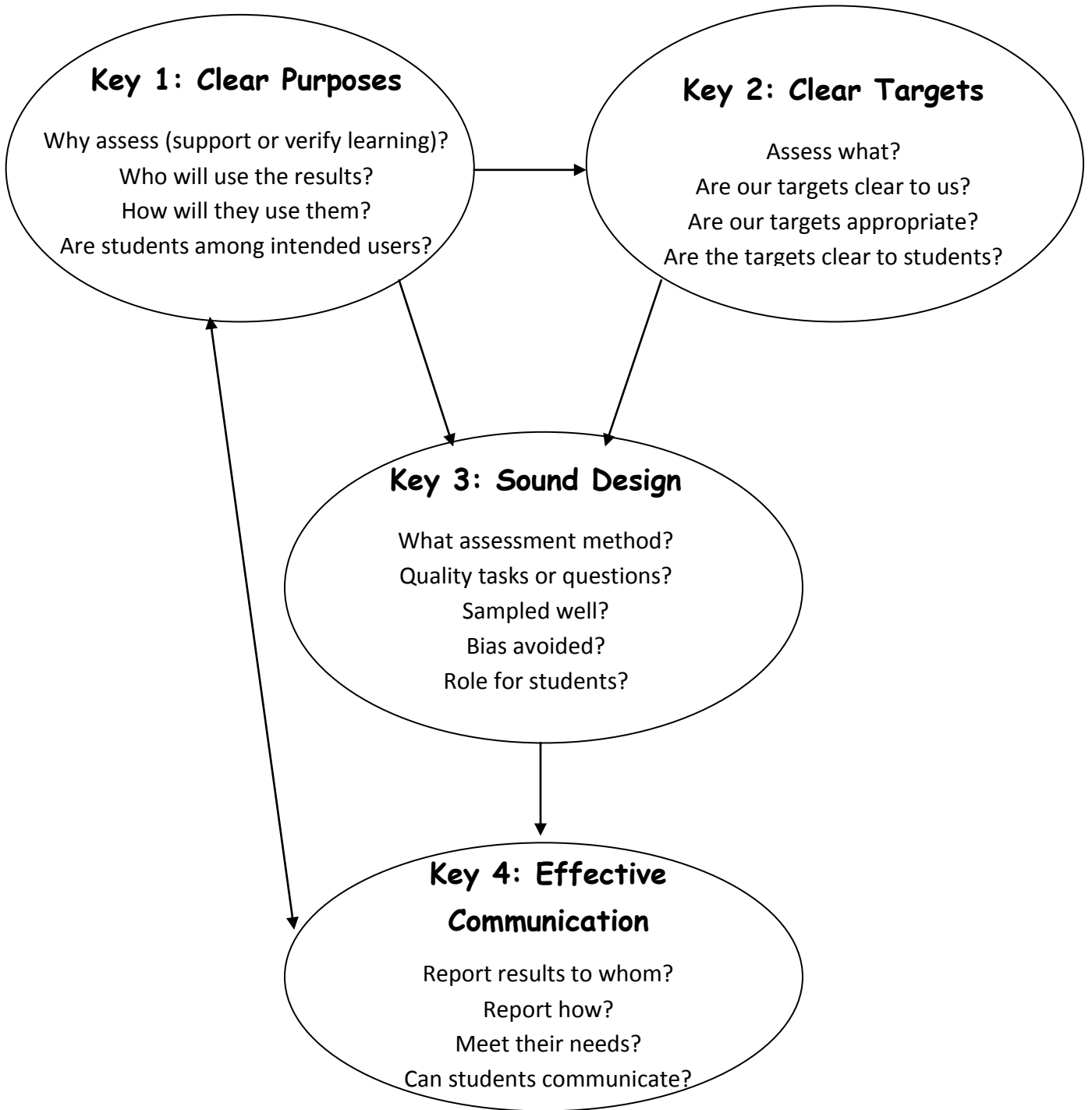
A Summary of Principles

It is instructive to summarize these five guiding principles in graphic form as in Figure 1 (next page). Teachers start the assessment process with a clear understanding of the assessment context: a clear purpose (instructional decision to be made) and a clear and appropriate learning target. Then they design or select an assessment to fit that specific context—to provide the information needed. Having conducted the assessment the next step is to transform the results into useful information which is delivered into the hands of the intended user in understandable terms. If this process unfolds, at least in part, during the learning—to meet the information needs of growing student/decision makers--the
result is likely to be profound gains in achievement and a narrowing of achievement gaps.

Part 2: Therefore, Teachers' Classroom Assessment Competencies

As mentioned previously, teachers typically spend a quarter to a third of their available professional time involved in assessment related activities. According to the above guidelines, this time must be spent evaluating their information needs and those of their students depending on the purpose for the assessment, carefully articulating the learning target(s) to be reflected in the exercises of their assessment, designing and developing quality assessments, administering them, scoring them, and communicating results appropriately. If teachers are to adhere to these guidelines, (a) both they and their direct supervisors must be provided with high-quality opportunities to master the specific professional standards outlined below and (b) the certification process must verify their mastery. Therefore, accreditation--of programs whose goal is to impart these competencies for teacher candidates—can and should require evidence that programs have actually attained that goal.

Figure 1

Keys to Assessment Quality

.

Key 1: Clear Purposes

Why assess (support or verify learning)?
Who will use the results?
How will they use them?
Are students among intended users?

Key 2: Clear Targets

Assess what?
Are our targets clear to us?
Are our targets appropriate?
Are the targets clear to students?

Key 3: Sound Design

What assessment method?
Quality tasks or questions?
Sampled well?
Bias avoided?
Role for students?

Key 4: Effective Communication

Report results to whom?
Report how?
Meet their needs?
Can students communicate?

Syntheses of dozens of studies conducted around 8the world tell us that teachers who master these competencies and apply them routinely in their classrooms, depending on the context, can rely on classroom assessment to raise student achievement between .4 to .8 standard deviations with the largest gains accruing for chronic low achievers (Black & Wiliam,1998; National Research Council, 2001; Hattie & Timperley, 2007; Wiliam, 2011;). The work synthesized therein forges powerful links between classroom assessment processes during learning and ultimate student learning success. These, then, are the competencies that correspond to the guiding principles developed in Part 1**:

1. Assessment literate teachers understand why each assessment event must begin with a clear purpose. That is, assessors must know why they are assessing—what they want their assessment to help them accomplish. They must believe that assessments can be used either to support or verify student learning depending of the context. They must be prepared to:
   a. Use assessments effectively to support learning (formative application) or certify it (summative use) depending on the context
   b. Anticipate the information needs of whose who will use the results in any particular context, paying special attention to decisions students are likely to make based on their interpretation of their own results
   c. Anticipate the instructional decision(s) to be informed by the results in the context of any particular application
   d. Determine what specific results an assessment must yield in order to inform those decisions (that is, serve the intended purpose)

2. Assessment literate teachers begin each assessment as competent, confident masters of the learning targets that are their assigned responsibility and that, therefore, are to be assessed. More specifically, they must be prepared to:
   a. Analyze each of their assigned achievement standards to determine the scaffolding students will climb over time to reach those standards.
   b. Classify learning target(s) as content knowledge, reasoning or problem solving, performance skills, or product development capabilities
   c. Select an appropriate assessment method from among the array of possibilities given the kind of learning target to be assessed—knowledge, reasoning, performance skill or product. (more about this under competence #3)
   d. Transform learning targets when necessary into student- and family-friendly forms when necessary to promote understanding

3. Assessment literate teachers can design and create assessments that provide dependable evidence of student learning. They can:
   a. Select the proper assessment method(s) given the learning target(s) for

_____

each particular classroom assessment context. They must know when to use and when to avoid the use of selected response, written response, performance assessment, and direct personal communication with students

    b. Develop plans for sampling student achievement with an efficient array of exercises or items that will support confidence inferences about student achievement

    c. Build quality assessments using each of the types of assessment methods above, including sound exercises and scoring procedures

    d. Anticipate and minimize distortion of results due to sources of bias relevant in each particular context, including biases due to students' cultural or linguistic background, flaws in the assessment itself, conditions of test administration, or scoring errors

    e. Understand why, when, and how to involve students in the design and development of practice assessments during their learning to advance their learning

4. Assessment literate teachers can communicate assessment results in a manner that fits each classroom assessment context. To do so, they must be prepared to:

    a. Keep accurate achievement records that arise from quality assessments

    b. Keep achievement records separately according to achievement target

    c. Keep achievement records separately according to purpose (formative and summative)

    d. Combine and summarize achievement information gathered over time appropriately to support inferences about student mastery of learning targets

    e. Communicate effectively when the intent is to support student learning by transforming results into descriptive feedback that help students understand what comes next in their learning

    f. Apply principles of effective communication when the purpose is to report judgments of the sufficiency of student learning in relation to expectations by, for example, relying on appropriate practices for transforming results into report card grades

    g. Engage students continuously in the processes of record keeping and communicating with others about changes in their own academic capabilities, such as with student-involved parent/teacher conferences

5. Assessment literate teachers routinely apply principles of assessment FOR learning while their students are growing in order to maximize student confidence, engagement, and motivation to strive for academic success—that is, they understand that students are key assessment users too, and take advantage of the emotional dynamics of student involvement to support learning success. Again, the specific competencies associated with this overarching idea are integrated into the first four competencies above, as in assuring that teachers know when and how to:

    a. Link assessment results directly to instructional planning

b. Share learning targets and standards of quality with students from the beginning of the learning
c. Design and use assessments so that students can self-assess and help to set goals for what comes next in their learning
d. Involve students in tracking, reflecting on, and communicating about their own learning progress

My colleagues and I have framed these competencies in terms of what the research evidence cited above suggests teachers really need to bring with them into their classroom from day one. In other words, regardless of the classroom assessment context, they had better be clear about who they're trying to help and how, the learning target to be assessed, the keys to quality assessment, how to communicate results productively, and when and how to involve their students. Without these, they will be unable to proceed with confidence. However, further deliberations by a team of qualified specialists might well lead to the development of continua for each competence from novice to intermediate to advanced levels of expertise. We have not done this work but remain open to its completion down the road.

Part 3: A Status Report on the Development of Assessment Literacy

All 50 states expect assessment competence as a condition for licensure. The National Association of State Directors of Teacher Education *Certification Manual on the Preparation and Certification of Educational Personnel* records those expectations, and the joint state project, the Interstate Teacher Assessment and Support Consortium (InTASC), makes it explicit in their teacher standards. Appropriate national professional associations hold similar expectations, as do the National Council for Accreditation of Teacher Education (NCATE) and the National Board for Professional Teaching Standards (NBPTS).   However, expectations in all of these contexts vary in their content coverage and tend to be quite general.

I also believe it is safe to assume that commonly used certification examinations focused on pedagogical knowledge also cover assessment knowledge at some level. However, because the detailed specifications of those exams are secure and have not been available for review and evaluation, I cannot comment on the specific aspects of assessment literacy addressed in their actual test items.

However, even with these expectations place, opportunities for teachers to attain appropriate levels of assessment literacy remain rare in pre- and in-service contexts and it has been so for decades. Our recent investigation of this matter attempted an analysis of the ten top-rated teacher preparation programs as rated by *US News and World Report* in 2009 in terms of the extent of their coverage of the five competencies described above (Stiggins and Herrick, 2009). Six programs agreed to participate in the investigation, while four declined. Of those six, only one program requires completion of courses in assessment. The others include assessment content in other courses. Analysis of the detail of those course syllabi revealed that only Competence 3 above, developing quality assessments, receives attention.   The other standards were essentially ignored. These results corroborate prior investigations of teacher education offerings in assessment detailed in our report and elsewhere (summarized in Stiggins and Conklin, 1992)

Given that the effective use of high-quality classroom assessment is so crucial to effective instruction my colleagues and I conducted a series of three meetings with national leaders in teacher preparation to identify the barriers that have so effectively kept assessment training from their preparation programs. Participants included deans of education, teacher education program directors and faculty, and classroom assessment specialists. A detailed list of the results of these deliberations appears in the appendix. It begins with guiding principles that the participants thought should be born in mind as our discussions of possible barriers proceeded. That having been established, we identified *institutional*, *contextual*, and *personal/professional* barriers revealing that, clearly, conditions exist within and around higher education institutions that mitigate against thoughtful assessment preparation for teacher candidates.

For instance, in terms of institutional barriers, there has really never been an internal institutional norm in higher education for quality assessment as a part of instruction. As a result, institutions have no accepted protocols for carrying out sound assessment practices or evaluating the quality of their work in this domain of instruction. Further, there are no institutional incentives for faculty to become assessment literate or practice sound assessment practices within their courses. Teacher education faculties, as a very relevant example, frequently lack members with sufficient assessment literacy to teach this material to their candidates.

Context barriers arise from realities imposed on higher education from the outside. For example, while licensing or certification standards refer to the need for assessment literacy, often they are vague on specifics and thus are subject to diverse institutional interpretations. If those local translations are made by faculties lacking assessment literacy, the result can be inadequate teacher preparation in this domain. Combine this with reality that practicing teachers who supervise field experiences of candidates often lack assessment literacy, and learning opportunities for candidates in assessment literate are likely to be thin in their coverage.

Personal/professional barriers arise within the psyche or the reality of individual faculty members. For instance, if one lacks assessment literacy yet still presumes to teach (or is assigned responsibility for teaching) sound practices to candidates anyway, the result can be either confusion about sound practices on the part of the candidate or the mastery of unsound practices. Or, when faculty have very sophisticated technical measurement knowledge but little knowledge of actual classroom practice, they can set lofty achievement standards that are beyond reach for and irrelevant to candidates. Either way, candidates will be (and historically have been) frustrated by the experience and are likely to remain uneasy about the meaning of sound assessment practices; and obviously, the quality of their assessment practices will suffer. Or there is a more subtle but related and equally troubling professional barrier: it can be risky for faculty to subject their own in-course assessment practices to scrutiny if they are not confident of their own assessment literacy. If a faculty member teaches lessons about sound assessment practices but is not modeling them in their assessments, the hypocrisy will be clear to students. Better, some might assert, to avoid the entire topic. Again, a longer list of the barriers discussed in our meetings appears in the appendix.

I believe that, when assessment training has been offered in pre-service contexts, it has commonly suffered from one of two presentation errors: either it is taught by an assessment specialist who knows little about learning targets or life in K-12 classrooms or schools, or it is taught by a subject matter methods specialist with inadequate assessment background. Either way is out of balance when what is essential is balance: a blending of what is to be assessed with how to assess it productively in a human environment.

Part 4: Assuring Quality Assessment Training in Teacher Preparation

In this part of the presentation, we consider a path past these barriers resulting in the development of assessment literate candidates, improved overall program quality, the collection of evidence of teacher competence for certification and program accreditation purposes.

I offer these ideas without prejudice in the sense that I have no way of knowing the current state of affairs in any particular institution. I have not studied methods of teacher preparation beyond those applied in the domain of assessment literacy. Further, I intend no prior judgments about the assessment capabilities or practices of faculty members in any particular institution. Rather, my offerings represent criteria against which those who lead teacher preparation might judge the efficacy of their faculties and programs as regards the development of candidate assessment literacy and program quality in assessment. I also acknowledge that assessment is only one of a number of domains of instructional competence programs must help candidates develop. However, the procedures I offer for developing and certifying assessment competence will generalize easily to other instructional competencies as well. In this sense, they can provide evidence of total candidate preparedness to teach that can be summarized across candidates to assist in evaluating program impact for accreditation.

However, for this evidence to be brought to bear in reaching these goals, certain institutional (faculty) conditions must be satisfied; that is, the guiding principles from Part 1 must be brought to the fore within programs of teacher preparation. The accreditation process could seek evidence of and make judgment about the extent to which these standards have been met:

- Relevant assessment *purposes* must be articulated in the context of teacher preparation, identifying the decisions to be made on the basis of assessment results, who will make them, and what information they need in order to support candidate learning or certify their competence depending on the context
- Relevant *learning targets* must be spelled out in learning progressions of achievement targets, including versions to be shared with candidates from the beginning of the learning
- All faculty members must be sufficiently assessment literate to devise *quality assessments* of candidate achievement for use in formative and summative contexts
- A record keeping and *communication* system is needed to deliver assessment results to intended users in a timely and understandable form, whether the purpose is to support learning or certify mastery of key standards

- Candidates can be involved deeply in self-assessment, record keeping, and communication processes both while they are learning and as they build their own case for certification with summative evidence that they are ready to teach

The barriers identified in the previous section will (can) only be removed if these conditions are satisfied. I detail each of our guiding principles below as they can play out in teacher preparation programs.

But before doing so, let me foreshadow how I would address the matter of candidate certification, a responsibility of state education agencies. In the end, the plan as I lay it out would place responsibility for building and presenting their case for the professional competence of candidates with the candidate. This would be akin to the process used by the National Board for Professional Teaching Standards and Alverno College in Milwaukee, WI in which the candidate assembles and presents a portfolio of evidence that all standards have been met. This approach is far superior to certifying based on a transcript of courses completed or a score on any one-shot certification test, because it addresses (certifies) all relevant professional standards by relying on the full array of assessment methods as needed to produce a rich portrait of readiness to teach. In addition and as importantly, the very process of building such a portfolio would represent a powerful learning experience for candidates as they prepare to make their case. And finally, it models the application of sound assessment practices within the context of candidates' experience as learners that they can, in turn, bring to bear in their own classrooms with their learners.

In addition, as mentioned above, while my focus has been on the development of classroom assessment competence, the process I spell out below can be applied to development and certification of all of the professional competencies needed to teach effectively. I propose that evidence gathered in this manner can be summarized across a sample of candidates to supply information relevant to program effectiveness and therefore qualification for accreditation.

While in Parts 1 and 2 I highlighted student-involved assessment FOR learning as its own guiding principle (#5) for special emphasis, in this section I will weave this idea continuously into a vision of its potential power in developing new teachers. In effect, I recommend teaching assessment FOR learning, in part, by modeling it.

The Matter of Clarifying Assessment Purposes

In the context of teacher preparation, at least four sets of decisions can be informed by assessment results, each associated with a different decision maker—all of whom impact candidate learning success. Each set of information needs must be accounted for.

On the formative side, *candidates* themselves 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. In order to use assessment FOR learning, they need to know the levels of achievement they are expected to attain, where they are currently at any point in time in relation to those expectations, and what they can do to close the gap and to achieve success. Only then can they plan and effectively manage their studies. As they travel the 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, on the summative side, *candidates* need evidence of their ultimate success— they will need access to a reservoir of information that tells them, ultimately, that they have mastered all of the materials required to become a teacher. This is the evidence they need to build their case that they deserve certification.

On the formative side, *professors* need access to information about the achievement of their students in order to diagnose the instructional needs of individual students and groups. Some assessment results may be unique to an individual student and some comparable across 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.

On the summative side, *faculty members* and *field experience supervisors* need dependable evidence of student mastery of relevant achievement standards in order to assign course grades and in order to provide candidates with the evidence of mastery of those standards in order to build their portfolio case for ultimate qualification to teach.

Program administrators must judge the quality of their program as a whole, determining if the teachers being graduated can put all of the pieces of good teaching 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 can 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* must gather evidence of program efficacy in order to determine if the program under scrutiny measures up to their standards. That body of evidence can and should include information about the quality of the assessment practices used during instruction, as well as specific evidence of candidate achievement. This means program evaluators must be sufficiently assessment literate to judge the quality of the in-course or in-context assessments reported by faculty. And, in this same vein, they must be qualified to sample candidates' final report portfolios and judge the quality of the evidence of candidate's preparedness to teach contained herein in order to make judgments about overall program quality. I will address this again later.

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.

Implications for Candidate Development and Ultimate Program Success

Ultimate candidate and program success turns on the quality of the instructional decisions that guide learning. Assessment is the process of gathering evidence of learning to inform those decisions. One key to candidate and program success is the quality of that evidence (of those assessments) that informs those decisions. If important evidence is missing or the evidence is of poor quality, candidate learning and ultimate disposition will suffer. One key to success, then, resides in understanding the information needs of the full array of decision makers who influence candidate and program success. During the learning, both students and professors make critically important instructional decisions that determine outcomes. Recent breakthroughs in formative applications of assessment have helped us see the power of a student/professor partnership that shifts much of the assessment during the learning to the learner's shoulders, acknowledging the importance of the decisions candidates make as they grow. This deep student involvement in their own assessment can set the stage for similarly deep involvement in the preparing the evidence of their own competence that will serve as the basis for their ultimate certification to teach. All of this models for candidates the very instructional decision making partnership they must form with their own students as they launch their careers.

The Matter of Establishing Achievement Expectations

The very foundation of a quality assessment, whether it is to be used to support or certify learning, we have established, is the definition of the learning target(s) to be assessed. In this regard, faculties of teacher education might productively address the following questions as they articulate their program's professional standards and prepare to assess formatively and summatively for certification and accreditation:

- 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?

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 ways one brings their knowledge to bear to reason and solve problems. 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 moment's 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 observable interactive behaviors to help students learn. These are the kinds of things one can see in teaching performance if one directly observes a teacher in action (live or on video). 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, teachers must be able to create examples of those products illustrating 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.

Implications for Candidate Development and Ultimate Program Success

By modeling good practice in their own clear and public up-front 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 teachers and students, increased sense of efficacy for both, and, of course, the potential for 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 diverse array of learning targets that underpin the academic subjects and pedagogy that represent their professional responsibilities. Part of the instructional and assessment task faced by teacher education faculty is to be sure they are. But, at the end of the preparation program, it should be the candidates' task to build and present their case on behalf of having mastered all relevant learning targets. This is how they will be judged.

The Matter of Quality Assessments

We already have established that the development of quality assessments for use in any particular context requires the careful:

- Selection of proper assessment methods to assure a match to the intended target(s),
- Assembly of high-quality assessment exercises and scoring schemes
- 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 engage in the development and use of formative assessments, all must understand and be able to apply these 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.

Here once again is an opportunity for teacher education faculty members to model for their candidates the application of sound practices. They can consistently point out to their students how their particular achievement expectations fit into the total curriculum that leads to ultimate success for teacher candidate and illustrate for them how to devise assessment plans and actual assessments that will yield dependable information for them and their

candidates revealing how each is progressing on the journey to excellence, as well as when they have arrived. This will require the regular matching of assessment methods to learning targets, along with the development of quality assessments that fairly sample candidate achievement and minimize the impact of key sources of bias that can distort results.

Implications for Candidate Development and Ultimate Success

It is worthy of note, once again at this point, that the list of barriers to development of assessment literate teachers included the stipulation that there has been no institutional norm in higher education demanding the accurate assessment of student achievement, let alone an acknowledgement of the power of student involvement in the assessment process to promote greater learning. As a result, there has been little modeling of sound practice in higher education instructions. If this can be turned around, benefits will accrue to candidates, faculty and institutions.

Further, one of the most productive assessment FOR learning strategies, teachers have discovered, is to engage students as partners during their learning in the *creation* and use of practice assessments like those for which they will be held accountable down the road. The acts of striving to understand keys to academic success sufficiently well enough to translate them into sample exercises and scoring guides, taking the assessment, scoring it, teaming up to analyze and interpret the results, and then acting on those results to overcome weaknesses all have the effect of promoting major advances in student learning. If new teachers come into the classroom with knowledge of and proficiency in implementing student-involved assessment procedures like this, 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, but only if they are qualified to offer this instruction and model these practices.

If prospective teachers do not learn these lessons about standards for quality 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 mis-measurement of their achievement, along with all of the predictable consequences of that. As stipulated previously, this represents a pedagogical, ethical, and indeed moral imperative in teacher education.

The Matter of Effective Communication of Assessment Results

With the information needs 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.

Effective Communication in a Formative Context

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 on the standard that is the focus of instruction. 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 for the sake of candidates' emerging academic and professional self-efficacy. This will provide the motivation needed to continue or to change directions if necessary.

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. Each would contain evidence of mastery of each relevant standard in the domain of study.

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.

Effective Communication in a Summative Context

As an 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, it is a status report portfolio. Change over time has no relevance in this collection of work. It contains only the most dependable and compelling evidence of mastery of pre-set professional standards, with that evidence taking the form of assessments successfully completed by the candidate as attested to by faculty members reflective of ultimate mastery of key standards. 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. Relevant 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 their portfolios for final evaluation. It would be incumbent upon the faculty to be sure 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.

As mentioned previously, this same approach could guide candidate certification by state education agencies. Given the diversity of content and pedagogical knowledge and skills

required for success across grade levels, content domains, and classroom contexts, it is not possible to conduct a single standardized evaluation of competence that will be sufficient in its coverage to verify the professional preparedness of all candidates. The certification portfolio approach addresses this problem head on by permitting each candidate to paint a rich and complete portrait of themselves in the context in which they intend to teach.

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 and candidates immense time and energy savings by permitting the entry, storage, and retrieval of achievement standards, actual assessments of their results as evidence of achievement (including samples of student work), and student self-reflections.

Implications for Candidate Development and Ultimate Program Success

As above, faculties of teacher education can help candidates learn about effective communication in an assessment context, in part, by modeling them. Modeling on the formative side takes the form of student involvement in self-assessment, record keeping and regular communication with others about their continuing progress. Modeling on the summative side takes the form of the building over time and ultimately communicating to others the evidence that they have mastered the standards of teaching excellence held as important by the faculty. In both contexts, the lessons learned will leave candidates understanding that effective communication of learning success begins with clear targets, includes quality assessment of each target, and ends with effective communication of evidence of mastery of each relevant target. This is true regardless of the content domain of the learning.

Part 5: A Summary Costs and Benefits

There is no question about the need for assessment literate teachers, nor is there any question about what that means in competence terms or what the impact of that competence will be on student learning. Yet, clearly, what has been happening (or not) over the decades in the development of classroom assessment competence has not worked. So the time has come, I suggest, to experiment with alternative approaches. This journey forward will take an investment in facing risks. But the evidence suggests there also is the promise of a fine return on that investment. As we move forward, it is important to weigh these potential costs and benefits for candidates and institutions.

The Costs and Risks

Clearly this plan depends on faculty agreement to a set of achievement standards that are reflective of the knowledge, reasoning, skills, products and dispositions that underpin effective teaching. And, it calls for a collaborative division of these standards across the courses and field experiences that comprise a total program of study. I have spelled out only

the classroom assessment standards in this paper. This plan calls for the development of a curriculum for each key professional standard, or perhaps set of standards, describing how candidates will grow from totally unprepared to highly qualified teachers over the span of a program. The emotional and time costs of doing this work of collaborative integration may be considerable. A reward system that has tended to acknowledge individual versus team achievement will be asked to value collaboration.

This system calls for the ongoing assessment of candidate success along the journey to becoming good teachers, when the faculty may lack assessment literacy. Such assessment competence 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 higher education assessment traditions prepares the faculty to lead candidates though this. Assessment has always been the job of the professor. Faculty will need time to shift of values and to learn how to do this.

This assessment plan requires systematic and continuous record keeping and communication about candidate achievement cast, not in terms of course grades, but mastery of each specific professional competence standard. This stands in stark contrast to an environment that traditionally delivers a transcript of grades 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.

This way of thinking about assessment opens the door to the evaluation of faculty and overall programs in terms of the actual achievements of their teacher candidates. Only those programs that model sound practices, offer productive instruction in assessment, and, ultimately, produce competent teachers will stand muster before established standards of professional competence in assessment practice. Those who review programs for accreditation will need to be trained to be sufficiently assessment literate themselves to judge quality.

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 centers on standards that truly do underpin excellence in teaching, we will receive in return highly-qualified teacher candidates. The students of those new teachers, in turn, will benefit from greater learning.

The evidence of candidate competence in teaching will be of higher quality. Both formative and summative assessments will be both more valid and reliable because those developing and implementing them will be sufficiently assessment literate to make it so.

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—those of candidates and faculty.

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 his or her 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-study 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:

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

Those who evaluate programs for accreditation find a clear path to the completion of their task in these questions too. Can program leaders and faculty provide evidence of opportunity to learn principles of sound assessment? Is there evidence in learning experiences observed and as reflected in the summative portfolios of candidates that sound practices have been modeled? Do summative portfolios provide compelling evidence of appropriate levels of assessment literacy in artifacts and student self-reflections? Is there testimony offered by faculty that students leave the program feeling that they have mastered the principles of sound assessment practice?

References

Black, P. & William, D. (1998) Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.

Chappuis, J. (2009) *Seven Strategies of Assessment FOR Learning*. Boston, MA: Pearson Education.

Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2012) *Classroom Assessment FOR Student Learning: Doing It Right—Using It Well,* 2nd Ed., Boston, MA: Pearson Education

Hattie, J. & Timperley, H. (2007) The power of feedback. *Review of Research in Education*. Retrieved October 9, 2007 from https://rer.sagepub.com/content/77/1/81.full.pdf+html

National Research Council. (2001). *Knowing What Students Know: The science and design of educational assessment,* Committee on the foundations of Assessment, Pelligrino, J. Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education, Washington, DC: National Academy Press.

Stiggins, R.J. & Herrick, M. (2007) Status Report of Teacher Training In Classroom Assessment. Unpublished research report, ETS Assessment Training Institute, Portland OR.

Stiggins, R.J. & Conklin, N.F. (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. *Educational Measurement: Issues and Practice*. 18 (1), 23-27.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation.* 37 (1): 3-14.

Appendix

Overarching conditions

Before deliberating about barriers to teacher preparation in classroom assessment, the committee of leaders in teacher education described in the text of the paper on page 16 chose to clarify several conditions or factors that we must bear in mind as we strive to remove the barriers:

- It is important for higher education institutions that support the development of effective schools (such as the organizations represented at this meeting) to create and advocate visions of excellence in assessment practice for teachers and for school leaders, even though it may be very challenging to make these visions a reality.
- "Pre-service" assessment training is not an appropriate frame of reference for teachers because practitioners are arriving in classrooms from so many different directions and with diverse backgrounds. As a result many are beginning practice before being trained to do so; so better to refer to "early career" training in assessment.
- A paradigm shift is at hand: Traditionally, we have seen assessment as the index of our effect on student learning; now we must help practitioners at all levels see it as part of the cause of our effect—that they can use assessment both to support and to verify learning.
- It will be important to think about training in sound assessment practice in terms of a developmental continuum where the vision extends from novice to expert; early career preparation must begin at the novice level for teachers and school leaders.
- We should continue to investigate special pedagogies for developing assessment literacy in ways tailored to the unique needs of the novices. For example, traditional course work may need to be supplemented with experiences in which teacher education or educational leadership faculty members teach sound assessment practices by modeling them or candidates take responsibility for developing their own assessment literacy through guided practice and team work.
- Two factors influence a new teacher's ability to practice sound assessment: knowledge of sound practice (assessment literacy) and the presence of enabling conditions in schools that permit or allow good practice; that is, context may dictate where one starts in developing assessment literacy and in implementing sound practices.
- Faculties cannot overlook how crucial it is for candidates to be masters of the content to be taught, learned, and assessed; that is, to understand the learning targets as an essential foundation for quality assessment.
- Ultimately, school leaders need a framework to use in the process of auditing, supervising and promoting the development of a teacher's classroom assessment practices.

Barriers to Productive Early Career Preparation in Assessment

Institutional Barriers—those that may arise from the realities of higher education institutional culture

- Colleges will evolve very slowly in this domain in part because of faculty territoriality—there are those in charge of the assessment wisdom and those who are not; so collaboration in the service of sound assessment practice has been and may continue to be challenging.
- Quality assessment has never really been an institutional priority in any internal context; the institution has no accepted protocols for carrying out sound assessment practices or evaluating their appropriateness.
- There are no institutional incentives for faculty to become assessment literate or to practice sound assessment practices within their courses.
- As a result, there has been no universally accepted set of criteria by which to judge the quality of an assessment or the appropriateness of its use in higher education.
- As a result, very often, teacher education or school leadership faculties lack someone with sufficient assessment literacy to teach it to candidates.
- Often program requirements such as in educational administration and teacher education are dictated by state law and don't include assessment; time with candidates is very brief and is filled with other required course work priorities.
- Traditionally, higher education faculties have not had to be concerned about struggling learners—those who can't hack it shouldn't be in college and should be weeded out; but teachers and school leaders do need to help those who struggle and need to learn how.
- Often, the measurement community in academia has been indifferent to, and sometimes cynical about, the possibility of sound classroom assessment for teachers; this community has been rigid about the meaning of sound assessments and who can access that wisdom; therefore, it has been uninvolved in, and sometimes opposed to, better assessment training for teachers and school leaders.

Contextual Barriers—those that may arise from the educational context outside of higher education in the governance of schools or from the reality of practitioners' life in schools

- The learning and performance demands of teachers and school leaders early in their careers can be overwhelming; assessment is merely one part of that which is to be mastered.
- Historically, teachers have not been trusted to assess either well or honestly; there is a pervasive fear that they will manipulate the evidence in their own best interest.
- Licensing or certification standards refer to the need for assessment literacy but are vague on specifics and thus are subject to diverse local interpretations; this means those who set such standards may not have sufficient understanding of sound assessment practice.

- Those practicing teachers who supervise field experiences often lack assessment literacy.
- Some of those experienced teachers who supervise field experiences may get it (about sound assessment practices) and do it well but not be able to teach the meta-cognitive keys to doing it well.
- The peer pressure in schools can demand compliance to existing norms of assessment behavior; so candidates or new hires who bring new ideas can be pressured into submission, thus preventing change in assessment practices.
- The dominance of the testing culture in the USA (that is, dominance of the testing industry) commands such a high proportion of the available resources that little is left to support balanced assessment or quality classroom assessment.
- The textbook industry has promulgated text-embedded tests as the solutions—thus trying to teacher-proof classroom assessment.
- There is a pervasive sense that teachers cannot assess well and so we need to "fix the problem" for them by providing the assessments or assessment results they need rather than teaching them to generate and use their own—so assessment literacy is unnecessary for new teachers.
- There has been a pervasive stereotypic view of assessment merely as an accountability tool, particularly among the political elite and in the measurement community, that has inhibited the development of the capacity to use assessment to support learning.

Personal/Professional Barriers—those that may arise from within the psyche of faculty members

- When one lacks assessment literacy yet still presumes to teach (or is assigned responsibility for teaching) sound practices to candidates anyway, the result can be either confusion on the part of the candidate or the mastery of unsound practices. Either way, candidates will be frustrated by the experience and are likely to remain uneasy about the meaning of sound assessment practices; and obviously, the quality of their assessment practices will suffer.
- Often, practitioners simply do not either believe or trust those who bring messages about the nature and importance of sound assessment practice; the measurement community often struggles to establish the relevance of its domain in the classroom—K-12 or in higher education; better, some might assert, to keep the entire topic at arm's length.
- Faculty members can regard it as risky to expose themselves to the evaluation of their teaching effectiveness through the evaluation of the learning outcomes of their students; it can be similarly risky to subject one's coursework assessment practices to scrutiny if one is not assessment literate; personal professional accountability in these senses can prevent higher education faculty from confronting their classroom assessment realities, let alone their ability to teach sound assessment practices to their teacher or school leaders candidates either didactically or through modeling. In other

words, if one teaches sound assessment practices but fails to model them, the hypocrisy will be clear to students; better, some might assert, to avoid the entire question.

- Often both teachers and administrators limit their focus on assessment simply because practical trade-offs must be made—there is more to do now than time will permit and this has been a low priority.   In other words, they have yet to learn that time invested in sound assessment can make their professional lives both more efficient and more productive.